Publikationen der Deutschen
Gesellschaft für Photogrammetrie,
Fernerkundung und Geoinformation e.V.

Band 27 **2018**

Beiträge

37. Wissenschaftlich-Technische Jahrestagung der DGPF e.V.
5. Münchner GI-Runde Runder Tisch GIS e.V.
66. Deutscher Kartographie Kongress der DGfK e.V.

7. – 9. März 2018 in München

Photogrammetrie - Fernerkundung - Geoinformatik - Kartographie - 2018

ISSN 0942-2870

Thomas P. Kersten, Eberhard Gülch,
Jochen Schiewe, Thomas H. Kolbe, Uwe Stilla (Hrsg.)

Publikationen
der Deutschen Gesellschaft für
Photogrammetrie, Fernerkundung
und Geoinformation e.V.

Band 27 **2018**

Beiträge

37. Wissenschaftlich-Technische
Jahrestagung der DGPF e.V.

5. Münchner GI-Runde
Runder Tisch GIS e.V.

66. Deutscher Kartographie
Kongress der DGfK e.V.

7. – 9. März 2018
in München

Photogrammetrie - Fernerkundung - Geoinformatik - Kartographie - 2018

ISSN 0942-2870

Thomas P. Kersten, Eberhard Gülch,
Jochen Schiewe, Thomas H. Kolbe, Uwe Stilla
(Hrsg.)

ISSN 0942-2870

Publikationen der Deutschen Gesellschaft für
Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e.V.
Band 27, 931 S., Hamburg 2018
Hrsg.: Thomas P. Kersten, Eberhard Gülch, Jochen Schiewe, Thomas H. Kolbe &
Uwe Stilla

© Deutsche Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
München 2018

Zu beziehen durch:

Geschäftsstelle der DGPF
c/o Technische Universität München
Institut für Geodäsie, GIS und Landmanagement
Lehrstuhl für Geoinformatik
Arcisstraße 21
D-80333 München
Tel.: 089 289-22578, E-Mail: geschaeftsstelle@dgpf.de

Redaktion:

Thomas P. Kersten
HafenCity Universität Hamburg
Labor für Photogrammetrie & Laserscanning
Überseeallee 16, 20457 Hamburg
E-Mail: Thomas.Kersten@hcu-hamburg.de

VORWORT

Die drei Fachgesellschaften Deutsche Gesellschaft für Kartographie e.V. (DGfK), Runder Tisch GIS e.V. (RT GIS) und Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V. (DGPF) laden zu ihren Jahrestagungen 2018 ein, die kooperativ unter der Gesamtüberschrift

PFGK18 (Photogrammetrie - Fernerkundung - Geoinformatik - Kartographie – 2018)

an der Technischen Universität München vom 7. bis zum 9. März 2018 stattfinden. Die PFGK18-Tagung wird von einer Fachfirmenausstellung begleitet und umfasst im Rahmen der drei Veranstaltungen

66. Deutscher Kartographie Kongress der DGfK

5. Münchner GI-Runde des Runden Tisch GIS

38. Wissenschaftlich-Technische Jahrestagung der DGPF

eine breite Palette von Sitzungen und Foren mit verschiedenen Konzepten, wie Fachsitzungen, Posterpräsentationen, Studentenforum, Anwenderforum, Wissenschaftsforum, Praxisforum, Innovationsforum, Preisverleihungen, Exkursionen und Mitgliederversammlungen.

Im Vorfeld der Tagung findet am Dienstag, den 6. März 2018, ein Studentenforum statt, das Studierenden – auch aus den unteren Semestern – die Gelegenheit gibt, Studierende aus anderen Hochschulen sowie namhafte Vertreter von Firmen, Behörden und Hochschulen aus den Bereichen Photogrammetrie, Fernerkundung, Geoinformatik und Kartographie kennen zu lernen und mit ihnen Kontakt aufzunehmen. Die Teilnehmer erhalten einen Einblick in mögliche Tätigkeitsfelder nach dem Studienabschluss und können die typischen Anforderungsprofile an junge Absolventen aus praktischer Sicht kennen lernen. Für die studentischen Teilnehmer ist am Abend ein eigenes Treffen vorgesehen, bei dem die Kontakte zu Studierenden von anderen Hochschulen vertieft werden können. Besonders freuen wir uns über die Teilnahme der Studentengruppen von der Fachhochschule Nordwestschweiz, Jade Hochschule Oldenburg, Universität zu Köln, Hochschule Karlsruhe, Universität Augsburg, Hochschule München und der TU München. An diesem Vortag werden auch Exkursionen zum Galileo-Kontrollzentrum des DLR und der nahe gelegenen weltbekannten Alten Pinakothek angeboten.

Der Mittwoch (7. März 2018) ist ab morgens der Fachfirmenausstellung, dem Weiterbildungsforum und den Anwenderforen gewidmet. Um 11:00 Uhr wird die Jahrestagung mit den Grußworten der Präsidenten, der Verleihung des Hansa-Luftbild-Preises und einem Eröffnungsvortrag „The Science of Where“ von Herrn Michael Goodchild, Emeritus Professor of Geography der University of California, Santa Barbara, offiziell eröffnet. Danach beginnen die Fachsitzungen und die Weiterführung verschiedener Foren. Flankiert werden diese Veranstaltungen mit Exkursionen zum Steinkeller des LDBV München.

Am Donnerstag (8. März 2018) beginnen die Sitzungen mit dem Plenarvortag „How should I drive here? Semantic Mapping for Automated Automobiles“ von Professor Christoph Stiller des Karlsruher Instituts für Technologie (KIT). Zwischen die Fachsitzungen und Foren sind die Präsentationen der Finalisten des Förderpreises Geoinformatik und des Karl-Kraus-Nachwuchsförderpreises eingebettet. Die festliche Abendveranstaltung findet mit Preisverleihungen im Hofgarten und Studentensaal der Traditionsgaststätte „Zum Franziskaner“ statt.

Am Freitag (9. März 2018) werden die Sitzungen und Foren in vier parallelen Strängen fortgeführt und durch den eingebettetem Plenarvortag „Faszination Raumfahrt – Perspektiven der robotischen und astronautischen Raumfahrt in Europa“ vom ESA Astronauten Thomas Reiter zusammengeführt. Die Plenarsitzung am Nachmittag schließt mit einem Resümee und weiteren Ankündigungen die PFGK18-Tagung ab.

Die Präsidenten und Vorstände der drei Trägervereine begrüßen das hohe Interesse, das sich durch die mehr als 500 angemeldeten Teilnehmer aus Deutschland, Österreich, Schweiz, Finnland, Frankreich, Kenia, Niederlande, Südafrika und den Vereinigten Staaten sowie durch die Beteiligung von 28 Firmen und Organisationen zeigt, die als Austeller und Vortragende in den Anwenderforen auf der Tagung vertreten sind. Neben den drei Sprechern für die Plenarvorträge wurden 29 Vorträge eingeladen. Aus den eingereichten wissenschaftlichen Beiträgen wurden 111 Beiträge akzeptiert, die mit 73 Vorträgen und 38 interaktiven Posterpräsentationen vorgestellt werden. Zehn der Posterpräsentationen werden im Wettbewerb zum Karl-Kraus-Nachwuchsförderpreis vorgestellt. In den Tagungsband wurden 95 schriftliche Beiträge aufgenommen.

Die lokalen Organisatoren freuen sich, die PFGK18 im Jubiläumsjahr der Technischen Universität München zu beherbergen. Vor 150 Jahren unterzeichnete Ludwig II. König von Bayern am 12. April 1868 den Gründungserlass für die "Königlich bayerische polytechnische Schule in München", die heutige Technische Universität München. Eineinhalb Jahrhunderte später ist daraus eine Universität von Weltrang geworden. Mit über 20 zentralen Jubiläumsveranstaltungen, beginnend mit dem Festakt im Herkulessaal der Residenz in München, stellt sich die Technische Universität München bis zum Jahresende der Öffentlichkeit vor. Daneben werden rund 120 Veranstaltungen der Fakultäten, Institute, Lehrstühle und sonstigen Einrichtungen einen Einblick in die Mannigfaltigkeit der wissenschaftlichen Forschung und des universitären Lebens an der TUM geben.

Die Präsidenten der drei Vereinigungen bedanken sich bei allen Gutachtern und Mitwirkenden in der Organisation und wünschen den Teilnehmern eine interessante und erfolgreiche Veranstaltung.

Prof. Dr.-Ing. Uwe M. Stilla
Präsident DGPF

Prof. Dr. Thomas H. Kolbe
1. Vorsitzender Runder Tisch GIS

Prof. Dr.-Ing. Manfred Weisensee
Präsident DGfK

INHALTSVERZEICHNIS

GEOPORTALE

P. KUNZ:
Kartographische Herausforderungen bei der Herstellung der TopPlus-Web-Open. 1

C. GLOCK & H. BIRKENBEUL:
ETRS89/UTM Umstellungsstrategie in Bayern für das Liegenschafts-kataster - Umformungsverfahren und Möglichkeiten der Umsetzung. 20

FERNERKUNDUNG I

S. KELLER, F. M. RIESE, N. ALLROGGEN, C. JACKISCH & S. HINZ:
Modeling Subsurface Soil Moisture Based On Hyperspectral Data - First Results Of A Multilateral Field Campaign. 34

P. M. MAIER, S. HINZ & S. KELLER:
Estimation of Chlorophyll a, Diatoms and Green Algae Based on Hyperspectral Data with Machine Learning Approaches. 49

M. WEINMANN & U. WEIDNER:
Erste Erfahrungen mit der Nutzung von Sentinel-2 Daten zur multispektralen und multitemporalen Analyse mit verschiedenen Klassifizierungsverfahren. 58

W. GLATZ & T. BAHR:
A Service Architecture for Processing Big Earth Data in the Cloud with Geospatial Analytics and Machine Learning. 71

PHOTOGRAMMETRIE I

R. BOERNER, Y. XU, L. HOEGNER & U. STILLA:
Registrierung von Flugzeuggetragenen Kameraaufnahmen und UAV-Aufnahmen zur Anreicherung von 3D Daten. 76

S. BUSCH:
Smart Phone Accuracy of Multi-Camera Pedestrian Tracking in Overlapping Fields of View. 86

F. MAIWALD, D. SCHNEIDER & F. HENZE:
Bestimmung der Korrespondenz zwischen historischen Gebäude-aufnahmen basierend auf der Zuordnung geometrischer Merkmale. 103

BILDUNG, STANDARDS UND RECHT

C. LINDNER, H. HODAM, A. ORTWEIN, J. SCHULTZ, F. SELG & A. RIENOW:
Sentinel-Daten für digitale und interaktive Anwendungen
im Schulunterricht. 117

J. SCHULTZ, H. HODAM, C. LINDNER, A. ORTWEIN, F. SELG & A. RIENOW:
Ableitung von 3D-Modellen aus Daten des High Definition Earth
Viewing-Experiments (ISS) – Anwendungen für den Schulunterricht. 129

ATLANTEN

S. BLEISCH, A. ZANDA & S. KORKUT:
Geschichten erzählen mit verlinkten Visualisierungen –
ein Review des Alters-Atlas Konzepts. 141

FERNERKUNDUNG ANWENDUNG - FORST UND AGRAR I

M. RÖDER , E. GÜLCH & H. LATIFI:
Eignungsprüfung einer UAV-basierten Forstinventur als Ersatz
zu traditionellen Feldverfahren in Verjüngungsbeständen. 151

GEOINFORMATIK I

T. BRINKHOFF:
Bestimmung von Ortspositionen unter Nutzung
von freien globalen Daten. 164

F. THIEMANN, S. CAKIR & F. POLITZ:
Generalisierung mittels Deep Learning (CNN) am Beispiel
der Straßenextraktion aus GPS-Trajektorien. 177

A. MEINERT, S. HAHMANN, S. KOPF & P. BRAUSEWETTER:
Simulation von Menschenmengen im urbanen Umfeld. 187

M. ELFOULY, A. DONAUBAUER & T. H. KOLBE:
Prediction of Feed-in Power from Photovoltaics
in Local Distribution Networks. 199

KARTOGRAPHISCHE ALGORITHMEN

J. CHANG & J. SCHIEWE:
Task-oriented Data Classification of Choropleth Maps for
Preserving Local Extreme Values. 213

P. FRECKMANN:
Hot-Spot-Analyse zur Unterstützung der Datenklassifizierung –
ein Versuch zur Minimierung der Subjektivität in thematischen Karten
am Beispiel des Projekts AlpES (Alpine Ecosystem Services). 219

FERNERKUNDUNG ANWENDUNG - FORST UND AGRAR II

K. KUNZ, Y. HU & U. SCHMIDHALTER:
Hochdurchsatz-Phänotypisierung von Weizen unter Hitze- und Trockenstress mit Hilfe von UAVs. 232

FERNERKUNDUNG ANWENDUNG - GEOLOGIE

A. RIEDEL, B. RIEDEL, C. SCHOTTMÜLLER, W. NIEMEIER & M. GERKE:
Bestimmung der Erdoberflächenbewegung mittels Radarinterferometrie nahe der Agua Blanca Störung, Baja California. 242

FERNERKUNDUNG ANWENDUNG - FORST UND AGRAR III

L. PREY, M. WEIGEL & U. SCHMIDHALTER:
Validierung einer UAV-basierten Multispektralkamera zur Erfassung von Bestandesmerkmalen in Weizen. 258

N. WILKE, B. SIEGMANN, A. BURKART, L. KLINGBEIL & U. RASCHER:
Quantifizierung der Ährenanzahl von Gerste mittels UAV-gestützter Bildaufnahmen. 274

U. LUSSEM & G. BARETH:
Introducing a New Concept for Grassland Monitoring: the Multi-temporal Grassland Index (MtGI). 285

WEBKARTOGRAPHIE

S. HENNIG:
Inklusive Web-Kartographie: Ausgewählte Methoden und Ansätze zur Erfüllung (gesetzlicher) Forderungen nach Barrierefreiheit. 294

T. HURTIG & H. SCHARLACH:
NLGA-Map: Web-gestützte thematische Karten mit JavaScript und Leaflet. 306

GEOINFORMATIK II

I. KARUT, A. WICHMANN, Y. FILIPPOVSKA & M. KADA:
Ein Konzept zur integrierten Modellierung von Detailstufen für 3D-Gebäudemodelle. 317

M. CHRISTEN, U. CLEMENT & A. MEYER:
Mixed Reality Anwendungen mit 3D-Stadtmodellen. 328

J. GEHRUNG, M. HEBEL, M. ARENS & U. STILLA:
Eine voxelbasierte Metadatenstruktur zur Änderungsdetektion in Punktwolken großflächiger urbaner Gebiete. 341

SENSOREN UND PLATTFORMEN

A. HANEL & U. STILLA:
Verkehrszeichendetektion und -erkennung in Bildern mittels faltenden neuronalen Netzen. 355

B. BORGMANN, M. HEBEL, M. ARENS & U. STILLA:
Fußgängerbezogene Informationsgewinnung zur Situationsanalyse mit einem mobilen Multisensorsystem. 363

S. BUSCH, J. QUEHL & C. BRENNER:
High Definition Mapping Using LiDAR Traced Trajectories. 376

H. SARDEMANN, A. ELTNER & H.-G. MAAS:
Erfassung von Geometriedaten kleiner Flüsse mit einem unbemannten Wasserfahrzeug als Multisensor-Plattform. 389

KARTOGRAPHISCHE VISUALISIERUNG

A. KELER:
Modeling and visualizing the spatial uncertainty of moving transport hubs in urban spaces - a case study in NYC with taxi and boro taxi trip data. 397

PHOTOGRAMMETRIE II

M. WEINMANN, R. BLOMLEY, M. WEINMANN & B. JUTZI:
Investigations on the Potential of Binary and Multi-class Classification for Object Extraction from Airborne Laser Scanning Point Clouds. 408

K. JACOBSEN & U. SEFERCIK:
Dense Matching mit WorldView-4 und Kompsat-3 Bildern. 422

P. SCHÄR, S. CAVEGN, D. NOVAK, B. LOESCH, H. EUGSTER & S. NEBIKER:
Ein systematischer Vergleich verschiedener Multi-View Stereo-Lösungen für die luftbildgestützte dreidimensionale Infrastrukturkartierung. 431

AUS- UND WEITERBILDUNG

A. BRUNN:
Blended Learning in Photogrammetrie und Bildverarbeitung für regionale KMUs – Ein Erfahrungsbericht. 450

C. TRAUN:
Impulse aus der Praxis – Bereicherung von Forschung und Lehre aus berufsbegleitenden GI-Bildungsangeboten. 461

M. TEUCHER, C. GLÄßER & V. SCHREINER:
E-Learning und Learning Analytics in der universitären Ausbildung der Geowissenschaften - Blended Learning mit ESRI ArcGIS Online im Masterstudium. 465

F. TSCHIRSCHWITZ, M. LINDSTAEDT & T. KERSTEN:
Immersive 3D-Visualisierung von Bauwerken mit Virtual Reality in der Hochschulausbildung. 471

FERNERKUNDUNG II

N. H. ISYA, A. RIEDEL, B. RIEDEL & W. NIEMEIER:
Comparison of Power Law Tropospheric Correction for Time Series InSAR Application. 483

M. F. MEYER, I. PFEFFER & C. JÜRGENS:
Methoden zur automatisierten Suche von Bodendenkmälern in LiDAR-Daten – neues aus Westfalen. 508

F. BEYER & G. GRENZDÖRFFER:
Klassifikation von Vegetationstypen auf Moorstandorten unter Verwendung von multisensoralen Drohnendaten. 521

D. GÜNTHER-DIRINGER & D. BRAUN:
Bilanzierung von Landnutzungsdaten 2007 und 2017 – Ist ein Vergleich möglich? 534

POSTER

N. PREUß:
Untersuchungen zur kartographischen Qualität touristischer Stadtpläne. 536

T. HAHMANN & E. DIEDRICH:
Three years of operations of the Sentinel-1 and Sentinel-3-OLCI PAC at German Aerospace Center. 547

S. SCHNEIDER & V. COORS:
Automatische Extraktion von Fenstern in 3D Punktwolken mittels einer hierarchischen Methode. 559

A. ELTNER, H. SARDEMANN, M. KRÖHNERT & H.-G. MAAS:
Entwicklung eines low-cost Kamerapegels zur Erfassung hydrologischer Extremereignisse. 573

M. MEHLTRETTER & C. HEIPKE:
Illumination Invariant Dense Image Matching Based On Sparse Features. .. 584

F. POLITZ, B. KAZIMI & M. SESTER:
Classification of Laser Scanning Data Using Deep Learning. 597

D. X. U. NGUYEN, F. ROTTENSTEINER & C. HEIPKE:
Object Proposals for Pedestrian Detection in Stereo Images. 611

M. GRAF & M. CHRISTEN:
Augmented Reality an Fenstern: Herausforderungen und Lösungskonzepte. 624

A. DUTTA, J. M. GITAHI, P. GHIMIRE, R. MINK, G. PETEINATOS, J. ENGELS, M. HAHN & R. GERHARDS:
Weed Detection in Close-range Imagery of Agricultural Fields using Neural Networks. 633

R. SIHOMBING & V. COORS:
Object-Based Mobile Augmented Reality for a 3D Model. 646

M. SCHOTTEN & J. SCHIEWE:
Automatisierte Auswahl von Methoden zur Überlagerung von Geodaten mit unterschiedlichen Bezugsgeometrien. 656

T. WENDT, R. HECHT & M. BEHNISCH:
Crowdsourcing zur Informationsgewinnung von Potentialen zur Nutzung von Gebäudeintegrierter Photovoltaik. 658

A. KOUKOFIKIS & V. COORS:
Employing OGC's Portrayal Service to Interoperate Hierarchical Data Structures. A Case Study on Visualizing I3S In Cesium. 670

C. EISFELDER, H. O. KIM & C. KUENZER:
Multi-layer Land Cover Data for Remote-Sensing based Vegetation Modelling for South Korea. 679

P. MWANGI:
Analyzing the Spatial Relationship between Building Volumes and Land Surface Temperature in Upper-Hill, Nairobi, Kenya. 691

BEITRÄGE KARL-KRAUS-NACHWUCHSPREIS

D. LAUPHEIMER & NORBERT HAALA:
Deep Learning for the Classification of Building Facades. 701

L.-M. KRALISCH, H. LILIENTHAL & M. BREUER:
Untersuchung zur Nutzung der Radarfernerkundung zur Klassifizierung von Feldfrüchten. 710

C. BEIL & T. H. KOLBE:
Detaillierte Repräsentation des Straßenraums in 3D-Stadtmodellen. 717

O.R. MOSLEHIRAD, G. PIPELIDIS, D. IWASZCZUK, C. PREHOFER & U. HUGENTOBLER:
Crowdsourced Vertical Indoor Mapping. 729

H. GÜMGÜMCÜ, D. LAUMER, J. D. WEGNER, P. BEARDSLEY & K. SCHINDLER:
Hand Tracking using Kalman Filter for Safe Human-Robot Interaction. 743

A. M. PITER:
Untersuchungen zur dreidimensionalen Analyse von photogrammetrisch erfassten Betonproben. 749

D. HOLDENER:
Konzeption und Realisierung eines neuen portablen 360°-Stereokamerasystems. 755

M. SCHÄR:
Einsatz von Deep Learning zur Aktualisierung der Arealstatistik der Schweiz - Erste Untersuchungen. 761

R. NIESSNER, H. SCHILLING, S. HINZ & B. JUTZI:
Klassifikation von Fahrzeugen aus RGB- und LiDAR-Daten mit Convolutional Neural Networks. 767

F. BREITWIESER, J. SCHOLZ & P. MANDL:
Simulation raumzeitlicher Effekte des Angebotes und der Nachfrage von Hackschnitzeln in Kärnten: Ein Agentenbasierter Ansatz. 778

BEITRÄGE FÖRDERPREIS GEOINFORMATIK

T. KELM, U. KLEIN, F. NETZEL, C. MUDERSBACH, M. KRÜGER & A. PFISTER:
Entwicklung und Implementierung eines „Weather Extremity Index (WEI)“ zur Bewertung von Starkregenereignissen unter Berücksichtigung räumlicher und zeitlicher Verteilungen. 787

L. LIEBEL:
Deep Convolutional Neural Networks for Semantic Segmentation of Multispectral Sentinel-2 Satellite Imagery: An Open Data Approach to Large-Scale Land Use and Land Cover Classification. 790

C. LUDWIG, S. KASS & M. RIFFLER:
A Semi-Automatic Algorithm For Wetland Detection Using Multi-Temporal Optical Satellite Data. 794

N. PIVETEAU, J. SCHITO, M. RAUBAL & R. WEIBEL:
A Novel Approach to the Routing Problem of Overhead Transmission Lines. 798

R. RUHDORFER:
Kopplung von Verkehrssimulation und semantischen 3D-Stadtmodellen in CityGML. 802

C. BERGER:
Fusion of VHR Multispectral and Object Height Data for Urban Environmental Monitoring. 806

P. V. KUPER:
Spatio-Temporal Data Handling for Generic Mobile Geoinformation Systems. 823

T. KUTZNER:
Geospatial Data Modelling and Model-driven Transformation of Geospatial Data based on UML Profiles. 827

RT GIS: TOPMODELS FÜR DIE GEOINFORMATIK

C. ATKINSON:
Reconciling Explanatory and Constructive Modes of Modeling through Deep Modeling. 829

Z. YAO & T. H. KOLBE:
A new Approach to Model Transformation using Graph Transformation System. 831

T. KUTZNER & T. H. KOLBE:
CityGML 3.0: Sneak Preview. 835

RT GIS: Neues zur Digitalisierung in der Landwirtschaft

P. Noack:
Einsatz von Multi- und Hyperspektralsensordaten in der Landwirtschaft. 840

G. Grenzdörffer:
Flugroboter für Precision Farming und das Feldversuchswesen. 851

T. Machl & W. Ewald:
Neue Wege für die Zukunft? Erfahrungen aus dem Einsatz neuer Planungsgrundlagen für die Konzeption ländlicher Kernwegesysteme. 857

RT GIS: Geodateninfrastrukturen - offen, sicher und nachhaltig

R. van Rossem:
BGT, The Realisation of a Key Register of Large Scale Topography in The Netherlands. 862

A. Matheus:
Sicherheitsaspekte und Best Practices in Geodateninfrastrukturen für den Einsatz in Smart Cities. 864

D. Carstensen:
Langzeitspeicherung von Geodaten - eine Lösung in Sicht!. 877

RT GIS: 3D-Geoinformation öffnet neue Anwendungsfelder

A. Donaubauer & T. H. Kolbe:
Leitfaden Geodäsie und BIM. 886

G. Fiutak, C. Marx, P. Willkomm, A. Donaubauer & T. H. Kolbe:
Automatisierte Generierung eines digitalen Landschaftsmodells in 3D. 888

R. Ott:
Geoinformationen in der Immobilienbewertung. 903

RT GIS: Geoinformation zum Mitmachen - Crowdsourcing und Citizen Science

J. Schiewe:
Kollaborative Analyse und Entscheidungsfindung an Tangible User Interfaces – Herausforderungen für die Modellierung und Visualisierung von Geodaten. 907

RT GIS: GEODATENEINSATZ FÜR EIN NACHHALTIGES STRAßENMANAGEMENT

G. HERRIG:
Community Mapping im kommerziellen Umfeld am Beispiel der HERE Basiskarten. 913

F. STEINBACHER & M. VEH:
Wie die Elektromobilität unsere Infrastrukturplanung verändert. 917

R. BEHRENS:
Softwaregestützte Fortführung der Straßenbestandsdaten im GIS. 921

M. KLEEMANN & M. STECKER:
Einsatz von Real-Time Daten zur Optimierung des Winterdienstes in Bayern. 923

AUTORENVERZEICHNIS 926

PROGRAMMKOMITEE 931

Kartographische Herausforderungen bei der Herstellung der TopPlus-Web-Open

PETER KUNZ [1]

Zusammenfassung: Mit dem Produkt TopPlus-Web-Open stellt das Bundesamt für Kartographie und Geodäsie (BKG) erstmals eine frei zugängliche Webkarte bereit, welche komplett auf offenen bzw. freien Datenquellen basiert. Hierzu wurden verschiedene Ausgangsdaten aufbereitet und miteinander kombiniert, um eine bestmögliche Kartendarstellung zu erzielen. Die Dienste der TopPlus-Web-Open wurden am 26. September 2017 auf der INTERGEO freigeschaltet. Das Produkt konnte sehr schnell einen großen Zuspruch aus dem öffentlichen, privaten und wissenschaftlichen Sektor verzeichnen und wird bereits in vielen Portalen (Geoportal RLP, LANIS, openrouteservice, ArcGIS Online) als Hintergrundkarte eingesetzt. Besonders positiv werden vor allem der freie Zugang und die grenzübergreifende Kartendarstellung hervorgehoben.

1 Verfahren TopPlus

Das Verfahren TopPlus wurde beim Bundesamt für Kartographie und Geodäsie (BKG) [URL 5] entwickelt, um aus Geodaten Karten verschiedener Maßstäbe für Web und Druck ableiten zu können (KUTTERER 2015). Alle Kartenbilder des Verfahrens TopPlus zeigen eine grafisch einheitliche Gestaltung. Sie stellen auch weite Bereiche des benachbarten Auslands dar und sind häufig aktueller als die vorhandenen Topographischen Karten. Das Verfahren TopPlus ermöglicht neben der Ableitung von Karten in Monitorauflösung auch die Erstellung von hochaufgelösten Rasterdaten. Somit können aktuelle Präsentationsgrafiken in Druckqualität und als georeferenzierte Rasterdaten für Geoinformationssysteme bereitgestellt werden (KUNZ 2014). Inzwischen sind mit dem Verfahren verschiedene Webkarten und eine Serie von hochaufgelösten Präsentationsgrafiken entstanden.

1.1 Produktportfolio TopPlus

Bereits 2012 hat das BKG damit begonnen, verschiedene TopPlus-Produkte für die Verwendung im internen Bereich von Bundeseinrichtungen bereitzustellen. Neben den Darstellungsdiensten in Monitorauflösung entstanden auch Kartendarstellungen in festen Maßstäben und in einer hohen Auflösung (200 Linien/cm, 508 dpi), welche sich sehr gut für qualitativ hochwertige Ausdrucke eignen. Die TopPlus-Produktfamilie besteht inzwischen aus nachstehenden Produkten:

- TopPlus-Web (Webkarte Deutschland und Europa, UTM32, EPSG:25832)
- TopPlus-World (Webkarte weltweit, Sphärische Mercatorabbildung, EPSG:3857)

[1] Bundesamt für Kartographie und Geodäsie, Abt. Geoinformation, Geodateninfrastrukturleistungen, Richard-Strauss-Allee 11, D-60698 Frankfurt am Main, E-Mail: peter.kunz@bkg.bund.de

Präsentationsgrafiken in hoher Auflösung in UTM32, EPSG:25832, für Deutschland und das benachbarte Ausland:

- TopPlus-P5 1:5.000
- TopPlus-P10 1:10.000
- TopPlus-P17.5 1:17.500
- TopPlus-P25 1:25.000
- TopPlus-P50 1:50.000
- TopPlus-P100 1:100.000
- TopPlus-P250 1:250.000

Präsentationsgrafiken in hoher Auflösung in Lambertabbildung (LAEA), EPSG:3035, für Europa

- TopPlus-P250-Europa 1:250.000

Diese Karten basieren im Bereich der Bundesrepublik Deutschland ausschließlich auf amtlichen Daten wie dem Digitalen Basis-Landschaftsmodell (Basis-DLM), den amtlichen Hausumringen Deutschland (HU-DE) und amtlichen Hauskoordinaten Deutschland (HK-DE) und für die Höhenliniendarstellung auf dem Digitalen Geländemodell 1:10.000 (DGM10) [URL 24]. Immer wieder gab es aufgrund der Nutzungsbestimmungen dieser amtlichen Geobasisdaten Einschränkungen bei der Nutzung der Daten und Dienste der TopPlus für bestimmte Anwendungsfälle. Dies betraf insbesondere die Unterlizensierung und die Weitergabe der Daten an Dritte oder die Verwendung außerhalb des internen Bereichs der Behörden. So wurde auch die ebenenübergreifende Nutzung des Kartenmaterials in Bund, Ländern und Kommunen durch lizenzrechtliche Beschränkungen erschwert oder gar verhindert. Nicht zuletzt deshalb wurde vom BKG 2017 das offene Produkt TopPlus-Web-Open [URL 1] bereitgestellt.

Abb. 1: TopPlus-Web-Open (Zoomstufe 14), Datenquelle: OpenStreetMap

1.2 TopPlus-Web-Open – Die offene Karte

Mit der TopPlus-Web-Open wird eine Webkarte auf Basis offener amtlicher Geobasisdaten und vieler weiterer offener Datenquellen bereitgestellt, die frei und ohne Einschränkungen genutzt werden kann. Eingesetzt werden ausschließlich offene bzw. freie Datenquellen, wie z. B. die offenen Daten (Open Data) des BKG [URL 15] und der Bundesländer Hamburg, Berlin, Nordrhein-Westfalen und Thüringen, OpenStreetMap-Daten, Open Data der Deutschen Bahn AG und weitere Daten. Die TopPlus-Web-Open verfügt über 18 vordefinierte Zoomstufen – von der weltweiten Übersichtskarte bis hin zur detaillierten Stadtkarte in Deutschland – und ist in drei unterschiedliche Darstellungsbereiche unterteilt:

- Weltweite Darstellung für kleine Maßstäbe
- Europaweite Darstellung bis hin zu den mittleren Maßstäben
- Detaildarstellung für Deutschland und angrenzendes Ausland

Außerhalb der genannten Darstellungsbereiche ist im Dienst derzeit kein weiteres Kartenmaterial verfügbar. Das Kartenbild der TopPlus-Web-Open kann in der Webanwendung „Karten des BKG“ [URL 2] ausführlich begutachtet werden.

Mit diesem Vorhaben veranschaulicht das BKG, wie durch die Kombination verschiedenster Open-Data-Angebote und freier Quellen neue Produkte mit Mehrwerten geschaffen werden können und wie unerlässlich hierfür eine durchgängige Open-Data-Strategie zur Öffnung von amtlichen Datenquellen ist. Die Webkarte ist als Internetdienst über die standardisierten Open Geospatial Consortium (OGC)-konformen Schnittstellen Web Map Service (WMS) und Web Map Tile Service (WMTS) verfügbar, kann einfach in gängige GIS-Systeme und Webkartenanwendungen wie ArcGIS, QGIS, OpenLayers oder Leaflet integriert und auf den verschiedensten Endgeräten – vom Smartphone bis Desktop-PC – eingesetzt werden. Das Produkt liegt standardmäßig in der weitverbreiteten Web Mercator Projektion (EPSG:3857) vor. Über die WMS-Schnittstelle kann die Webkarte allerdings auch in weiteren gängigen Projektionen abgerufen werden.

1.2.1 Freie Nutzung der TopPlus-Web-Open

Die TopPlus-Web-Open wird unter der „Datenlizenz Deutschland – Namensnennung – Version 2.0“ [URL 4] bereitgestellt. Daten und Dokumente, die unter dieser Datenlizenz stehen, dürfen in jeder denkbaren Art und Weise genutzt werden. Ob dies zu kommerziellen oder nichtkommerziellen Zwecken geschieht, ist unerheblich. Die Nutzung ist zeitlich und räumlich unbeschränkt möglich. Die einzige Bedingung für die Nutzung ist, dass der Name der bereitstellenden Behörde genannt wird. Auf welche Weise diese „Namensnennung“ im Falle der TopPlus-Web-Open erfolgen muss, wird auf den Webseiten des Dienstleistungszentrums des BKG [URL 1] beschrieben.

2 Ausgangsdaten

Grundsätzlich können verschiedenste Geoinformationen im Verfahren TopPlus verwendet werden: amtliche Geobasisdaten [URL 7], Georeferenzdaten [URL 6] sowie nicht-amtliche Geoinformationen. In der TopPlus-Web-Open werden immer dann amtliche Datenquellen verwendet, wenn diese als offene Geodaten bereitstehen. In allen anderen Fällen wurde auf freie Daten und hier insbesondere auf OpenStreetMap-Daten zurückgegriffen. Nachstehende Daten kommen bei der TopPlus-Web-Open zum Einsatz [URL 3]:

Digitale Landschaftsmodelle

- Digitales Basis-Landschaftsmodell (Basis-DLM) von Berlin, Hamburg, Nordrhein-Westfalen und Thüringen [URL 11]
- OpenStreetMap-Daten in allen weiteren Bundesländern und im Ausland [URL 21]
- Digitales Landschaftsmodell 1:250.000 (DLM250) [URL 16]
- Digitales Landschaftsmodell 1:1.000.000 (DLM1000) [URL 17]
- EuroGlobalMap (EGM) [URL 10]
- Natural Earth [URL 23]

Weitere topographische Vektordaten

Von den Bundesländern Berlin, Hamburg, Nordrhein-Westfalen und Thüringen:

- Gebäude: Amtliche Hausumringe Deutschland (HU-DE) [URL 13]
- Industriegebäude, Öffentliche Gebäude, POIs: Amtliche 3D-Gebäudemodelle LoD1 [URL 14]
- Hausnummern: Amtliche Hauskoordinaten Deutschland (HK-DE) [URL 12]

- Geographische Namen 1:250.000 (GN250) [URL 19]
- Verwaltungsgebiete 1:250.000 (VG250) [URL 18]
- Verzeichnis der Staatennamen für den amtlichen Gebrauch in der Bundesrepublik Deutschland (Auswärtiges Amt)
- Landbedeckung: CORINE Land Cover; Europäische Umweltagentur (EUA) [URL 20]
- Flusskilometrierung: Verkehrsnetz der Bundeswasserstraßen, Wasser- und Schifffahrtsverwaltung des Bundes (WSV) [URL 9]
- Bahnhöfe und Haltestellen: Deutsche Bahn AG [URL 8]
- Siedlungen: GeoNames [URL 22]

Digitale Geländemodelle

Europa:

- Digital Elevation Model over Europe from the GMES RDA project (EU-DEM) [URL 25]

Weltweit:

- GMTED2010,U.S. Geological Survey [URL 26]
- SRTM15_PLUS Global Bathymetry, Data: SIO, NOAA, U.S. Navy, NGA, GEBCO [URL 27]

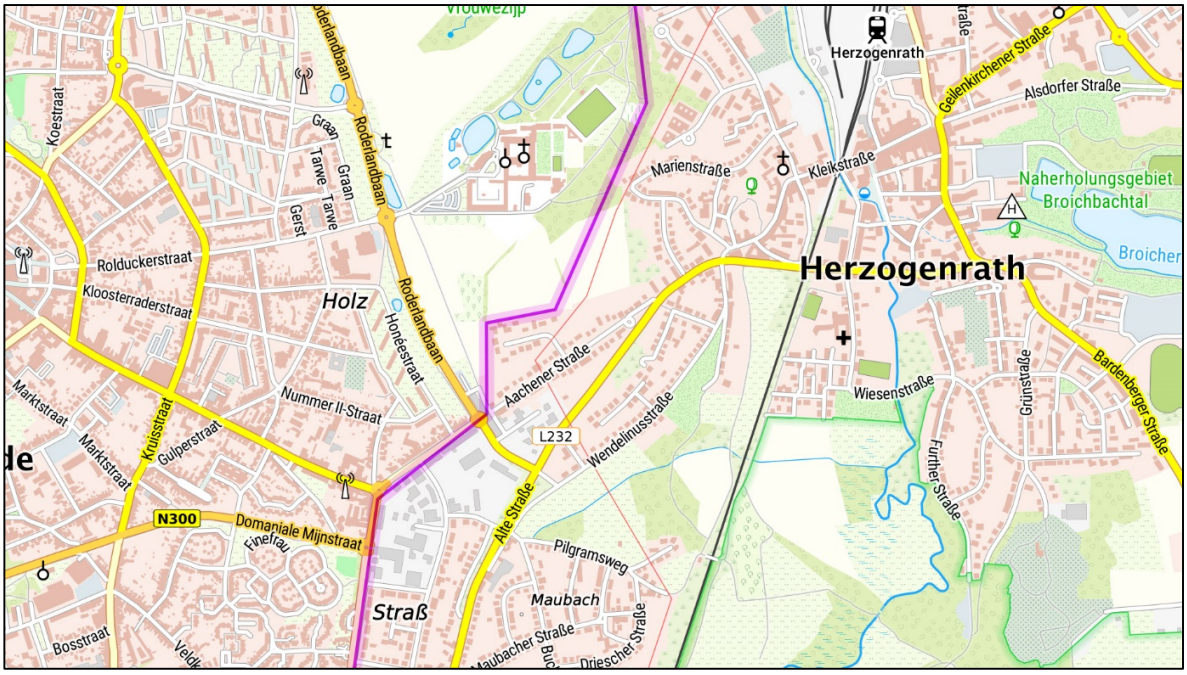


Abb. 2: TopPlus-Web-Open (Zoomstufe 15): grenzüberschreitende Darstellung: Basis-DLM in Nordrhein-Westfalen (rechts), OpenStreetMap in den Niederlanden (links)

3 Aufbereitung der Ausgangsdaten und Datenhaltung

Für das Aufbereiten vieler Ausgangsdaten und insbesondere für die Ableitung der Kartengrafiken – das Rendern der Karten – wird weitgehend auf freie Software (Open-Source-Software) zurückgegriffen. Zusätzlich kommt aber auch die kommerzielle Software FME [URL 31] für die Prozessierung der GIS-Daten zum Einsatz. Die Datenhaltung erfolgt in einer PostGIS-Datenbank [URL 28]. Mit der Programmbibliothek GDAL (Geospatial Data Abstraction Library) [URL 33] werden aus Digitalen Geländemodellen Höhenlinien und Schummerungen abgeleitet.

3.1 Datenimport

Für den Datenimport und die Konvertierung der Daten in PostGIS-Datenbanken werden im Falle der OSM-Daten die Software Imposm [URL 29] und osm2pgsql [URL 30] eingesetzt. Natural Earth Daten werden mit ogr2ogr importiert. Bei weiteren Daten wird jedoch FME für den Datenimport eingesetzt und hier erfolgt zusammen mit dem Importvorgang gleichzeitig eine teilweise umfangreiche Aufbereitung der Daten. Nachstehend werden beispielhaft Probleme beschrieben, die beim Einsatz ganz unterschiedlicher Ausgangsdaten auftreten.

3.2 Probleme bei der Kombination verschiedener Ausgangsdaten

Verschiedene Ausgangsdaten müssen auch innerhalb einer Zoomstufe kombiniert werden. So grenzen in großen Maßstäben Daten des Basis-DLM an die OpenStreetMap-Daten. In mittleren Maßstäben werden DLM250-Daten für Deutschland, EuroGlobalMap-Daten für Europa und für

alle weiteren Staaten Daten aus OpenStreetMap oder Natural Earth eingesetzt. Bei der Kombination verschiedener Ausgangsdaten sind nachstehende Herausforderungen zu beachten.

3.2.1 Kennzeichnung von Staaten und Bundesländern in OSM-Daten

Leider verfügen die OSM-Daten nicht über ein eindeutiges Attribut, welches die Lage eines Objektes in einem bestimmten Staat oder Bundesland kennzeichnet. Da die OSM-Daten aber mit anderen Daten wie dem Basis-DLM einiger Bundesländer oder EGM-Daten kombiniert werden, müssen die OSM-Daten zunächst mit den entsprechenden Verwaltungsgebieten verschnitten und mit der entsprechenden Länderkennung versehen werden. Dieser Vorgang ist aufgrund der großen Datenmenge nicht trivial und relativ aufwändig. Hinweis: Auch bei den OSM-Daten der Geofabrik [URL 21], welche für einen Staat oder ein Bundesland vorliegen, sind die Daten nicht an den jeweiligen Grenzen genau beschnitten. So fehlt vielleicht ein Teil des Rheins auf deutscher Seite im Datensatz von Deutschland und befindet sich dafür vollständig im Datensatz von Frankreich. Ebenso ist der Grenzverlauf in den OSM-Daten nicht identisch mit der entsprechenden Geometrie im Basis-DLM.

3.2.2 Unterschiedlicher Umfang der bereitgestellten Objektarten

Im Bereich der Open-Data-Bundesländer Berlin, Hamburg, Nordrhein-Westfalen und Thüringen werden nur die Objekte dargestellt, welche im Basis-DLM, den Hausumringen oder den 3D-Gebäudedaten vorkommen. Auf die Darstellung von OSM-Daten wurde hier verzichtet. So ist es zu erklären, das in diesen Open-Data-Ländern einige Objektarten wie z. B. Bushaltestellen, Einbahnstraßenpfeile oder POIs fehlen, welche in den benachbarten Bundesländern aber dargestellt werden. Dieses unterschiedliche Vorkommen bestimmter Objektarten gilt auch für die Kombination der anderen Ausgangsdaten in weiteren Zoomstufen.

3.2.3 Unterschiedliche Geometrien vergleichbarer Objektarten

Das Basis-DLM unterscheidet Straßenachsen von Fahrbahnachsen. Eine Autobahn hat normalerweise eine Straßenachse und zwei Fahrbahnachsen. In der TopPlus werden in mittleren Zoomstufen vorwiegend die Straßenachsen gezeichnet, in höheren Zoomstufen die Fahrbahnachsen. Die OSM-Daten zeigen aber im Falle von Autobahnen oder vergleichbaren Schnellstraßen nur Fahrbahnachsen. Für die Darstellung der OSM-Autobahnen wurden daher eigene Ableitungsregeln und andere Signaturen verwendet als für die Basis-DLM-Autobahnen.
Häufig lassen sich vergleichbare Objektbereiche verschiedener Datensätze aufgrund ihrer unterschiedlichen geometrischen Modellierung nicht gut kombinieren. So liegen zwar weltweite Meeresflächen in guter Qualität aus den OSM-Daten vor, jedoch lassen sich diese OSM-Meeresdaten nicht gut mit Flüssen oder Küstenlinien aus anderen Datensätzen verwenden. So reicht ein Fluss der EGM-Daten vielleicht nicht an die Küste der OSM-Daten heran oder eine Insel des Basis-DLM passt nicht in die entsprechende Freifläche im OSM-Meer. Es mussten daher topologisch richtige bzw. zueinander passende Kombinationen der Meeresflächen von Basis-DLM und OSM, DLM250 und OSM und von EGM und OSM erzeugt werden.

3.2.4 Unerwartete Attributwerte bei OSM-Daten

Immer wieder bereitet das Fehlen eindeutiger Datentypen und die Vielzahl möglicher Attributwerte der OSM-Attribute Probleme. Hier einige Beispiele für mitunter kaum interpretierbare Höhenangaben von Bergen.

Problem	**Name**	**Höhenangabe**
Angabe in Meter oder in Feet?	Mount Mackenzie	2456m;8058ft
	Mill Hill	905 feet
	Burbush Hill	209 M
	The Mount	89 Metres
Handelt es sich bei einem Punkt oder Komma um ein Tausender- oder ein Dezimaltrennzeichen?	Cerro Yuichuchu	P. 3.640 m
	Pico del Ave	952,40 m
	Volcán Chicabal	9,514 feet
Unbekannte oder ungefähre Höhenangaben	<NULL>	~1220
	Signal Mountain	0
	Fuchsberg	Fixme
	Kohleberg	around 610m.a.s.l.

Die gleichen Ungenauigkeiten treten bei der Angabe von Gewässerbreiten oder der Einwohnerzahl von Siedlungen auf und führen beim Sortieren der Daten oder bei Klassenbildungen zu Fehlern. Die Bildung von Klassen und das Gruppieren von Attributwerten werden auch durch das Vorkommen von zu vielen verschiedenen Attributwerten erschwert. So kommen in dem Attribut „amenity“ der Punktobjekte insgesamt 6451 verschiedene Attributwerte vor, darunter 196 Werte die häufiger als 100-mal weltweit vorkommen.

3.2.5 Fehlende Generalisierung der OSM-Daten

In mittleren Maßstäben werden in Europa EuroGlobalMap-Daten verwendet. Außerhalb Europas müssen diese je nach Zoomstufe durch OSM-Daten oder Natural Earth-Daten ergänzt werden. Im Falle der OSM-Daten wirkt sich in diesen Maßstäben die fehlende Generalisierung der Daten negativ aus. Immer wieder werden Flüsse oder andere linienförmige Objekte durch Straßen teilweise oder vollständig verdeckt. Bei der Verwendung von Natural Earth-Daten macht sich wiederum die zu starke Generalisierung der Natural Earth-Daten im Vergleich zu EGM bemerkbar.

3.3 Aufbereitung der Ausgangsdaten mit FME

Hier einige Beispiele der Anwendung von FME für die umfangreiche Aufbereitung der Ausgangsdaten:

- Schematransformation und Umprojektion für die Angleichung unterschiedlicher Ausgangsdaten an die gewünschte Datenbankstruktur.

- Zuweisen der Darstellungspriorität für Objekte bei höhenungleichen Kreuzungen (auf Brücken oder in einem Tunnel), wobei sich auch mehrere Brücken durchaus in einem Punkt kreuzen können. Hier ist die Ermittlung der genauen „Etage“ erforderlich.
- Zuweisen der Darstellungspriorität für die Straßenfüllungen bei höhengleichen Kreuzungen. Bei Einmündungen einer höherrangigen Straße in eine niederrangige Straße hat die niederrangige Straße die höhere Priorität (z. B. Einmündung der als Autobahn gewidmeten Rampe in eine Bundes- oder Landesstraße).
- Ermittlung der Tunnelportale bei Eisenbahntunneln. Gesucht werden Beginn und Ende einer Tunnelstrecke und in diesen Punkten wird der Azimutalwinkel eines kurzen Stückes der Tunnelachse bestimmt. In der Karte wird der Tunneleingang als Querstrich im rechten Winkel zur Tunnelachse gezeichnet. Die ermittelte Winkelangabe kann dann für den Parameter „transform =rotate“ (Drehwinkel) der SVG-Datei des Tunnelportals verwendet werden.
- Verlängerung von Straßenachsen, um mit dieser verlängerten Standlinie auch noch kurze Straßen oder solche mit langen Straßennamen beschriften zu können.
- Nachbarschaftsanalysen, z. B. Ersetzen von mehreren gleichartigen, benachbarten Objekten durch ein repräsentatives Objekt (z. B. Straßenbahnhaltestellen)

3.4 Anwendung von PostGIS für die Aufbereitung der Daten

Weitere Bearbeitungsschritte erfolgen nach dem Datenimport direkt innerhalb der PostGIS-Datenbank. Hier einige Beispiele für solche Datenaufbereitungen, bei denen entweder vorhandene Tabellen bearbeitet werden oder auch neue Datensätze entstehen:

- Prüfen der Geometrie von Polygonen auf Validität und ggf. Korrektur dieser Fehler.
- Erzeugen generalisierter Varianten der Daten für die Darstellung in mittleren und kleinen Maßstäben.
- Verschneiden von Industrieflächen und von Flächen besonderer funktionaler Prägung mit Hausumringen, um weitere Industriegebäude bzw. öffentliche Gebäude zu kennzeichnen.

3.4.1 Bestimmung von Points of Interest in den 3D-Gebäudemodelldaten (LoD1)

Im Gegensatz zu den amtlichen Hausumringen besitzen die 3D-Gebäudedaten wertvolle Informationen über die Gebäudefunktion und die Bezeichnung von Objekten. Diese Attribute werden dazu verwendet, Points of Interest und weitere topographische Einzelobjekte in den 3D-Gebäudedaten zu identifizieren. Diese Objekte werden in der Karte entweder nur beschriftet oder mithilfe einer Punktsignatur dargestellt. Zu diesen Objekten zählen z. B. Schulen, Kindergärten, Krankenhäuser, Kirchen/Kapellen, Polizei, Feuerwehr, Post, Theater usw. Da in den 3D-Gebäudedaten viele oder alle Gebäude einer Schule oder eines Krankenhauses diese Information mitbringen, werden z. B. die Einzelgebäude eines Krankenhauses zusammengefasst und durch einen oder mehrere repräsentative Punkte ersetzt. Da die amtlichen Hausumringe aktueller sind als die 3D-Gebäudedaten, werden in der TopPlus-Web-Open die Geometrien der Hausumringe zusammen mit den Attributen der 3D-Gebäude dargestellt. Dazu werden die 3D-Gebäude in die Klassen Wohn-, Industrie- und öffentliche Gebäude eingeteilt und danach mit den Hausumringen

verschnitten. Diese Informationen werden auch benötigt, bevor die Hausumringe unter Beibehaltung dieser Klassifizierung mit Change [URL 32] generalisiert werden.

3.4.2 Weltweiter Datensatz der Bahnhöfe

Es wurde ein weltweiter Datensatz der Bahnhöfe und Haltepunkte bis hinunter zur Bushaltestelle erzeugt. Kombiniert wurden Daten aus dem Basis-DLM mit Daten des DLM250, der Deutschen Bahn AG und von OSM. Dieser Datensatz beinhaltet alle Bahnhöfe, U-Bahn-, S-Bahn-, Straßenbahn- und Bushaltestellen. Die besondere Herausforderung war hier die Kennzeichnung von Bahnhöfen mit den verschiedenen Bahnkategorien, die an einem solchen Punkt halten.

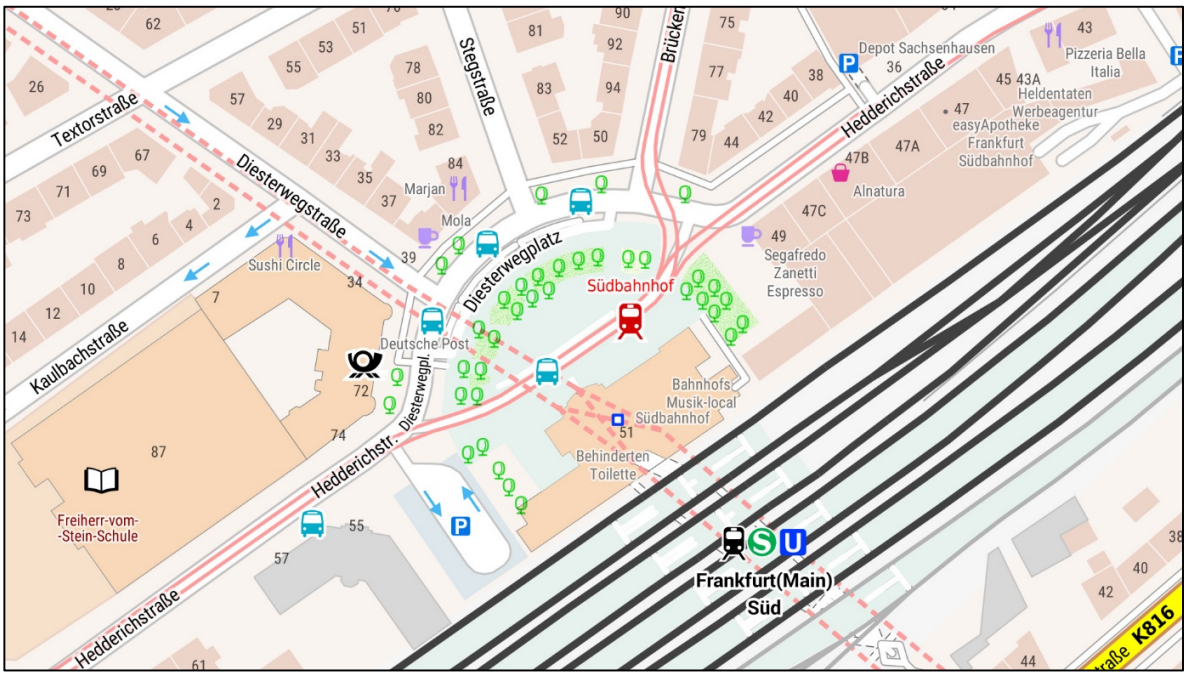


Abb. 3: TopPlus-Web-Open (Zoomstufe 18): Bahnhof Frankfurt (Main)-Süd mit Haltepunkten für Bahn (schwarz), S-Bahn (grün), U-Bahn (blau), Straßenbahn (rot) und Bus (hellblau)

3.4.3 Erstellung eines weltweiten Datensatzes der Siedlungen

Für die Darstellung der Ortsnamen wurde ein weltweiter Datensatz mit Siedlungen zusammengestellt, wobei die Namen aus amtlichen Datenquellen bevorzugt wurden. Auch diese Aufgabe ist nicht einfach lösbar, weil in den verschiedenen Datensätzen ganz unterschiedliche Schreibweisen für ein und denselben Ort verwendet werden und auch die Geometrien teilweise stark voneinander abweichen. Ebenso unterscheiden sich die Einwohnerzahlen von Datensatz zu Datensatz. Auch eine Suche im Umkreis und die Verwendung von Funktionen zum Finden ähnlicher Namen führen hier nicht immer zu einem befriedigenden Ergebnis. So ist es zu erklären, dass mitunter doch ein Ort in unterschiedlichen Schreibeweisen mehrfach in der Karte dargestellt wird. Die Kombination verschiedener Ausgangsdaten war aber erforderlich, weil ein einzelner Datensatz wie OSM alleine nicht alle Regionen und Maßstäbe abdeckt, die man durch eine Kombination verschiedener Ausgangsdaten erhalten kann. So liegt zwar mit Geonames.org ein weltweiter Datensatz vor, der sich gut für mittlere und kleine Maßstäbe eignet. In größeren

Maßstäben fehlen jedoch viele kleinere Siedlungen oder die Lage der bereitgestellten Koordinaten ist ungenau und liegt teilweise weit außerhalb der genannten Siedlung. Umgekehrt fehlen wieder viele Siedlungen aus Geonames.org in den OSM-Daten. Für die Beschriftung der Siedlungen in Deutschland wurde vorwiegend auf die Daten des DLM250 zurückgegriffen. Diese wurden im Bereich der Open-Data-Länder der BRD mit weiteren Siedlungsnamen aus dem Basis-DLM ergänzt. Eine besondere Herausforderung war in allen Fällen die Unterscheidung in einen Hauptnamen und ggf. einem Zweitnamen. Dieser Zweitname ist in Deutschland mitunter ein sorbischer (Cottbus/Chóśebuz) oder friesischer Name (Husum/Hüsem), im europäischen Ausland häufig ein deutsches Exonym (Roma/Rom). Weltweit kommen hier häufig Namen mit anderen Schriftzeichen zum Einsatz (Agadir / أكـــادير). In diesem Datensatz wurden die Daten von Basis-DLM, DLM250, EGM, Geonames.org und OSM verwendet.

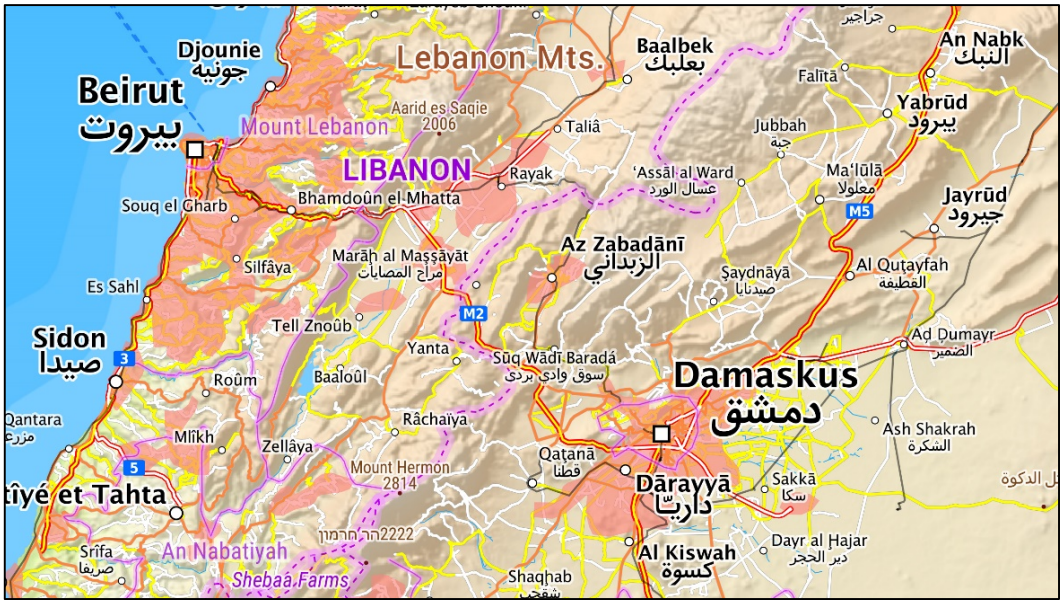


Abb. 4: TopPlus-Web-Open (Zoomstufe 9): Namen von Siedlungen im Libanon und in Syrien

4 Herstellung der Kartendaten mit Mapnik

Die zentrale Komponente zur Herstellung von Karten im Verfahren TopPlus ist die Open-Source-Software Mapnik – eine Sammlung von Bibliotheken für das Rendern von Karten [URL 34]. Die Software gehört zu den bekanntesten Tools für die Herstellung von Karten aus OSM-Daten. Beim Verfahren TopPlus wird mit Mapnik ein komplettes Kachelarchiv (ein Tile-Cache) vorprozessiert. Ein Darstellungsdienst kann anschließend diese Kacheln performant bereitstellen. Hierzu werden die vorprozessierten Tiles (Kachelarchiv, Cache) von einem Tile-Server über die standardisierten Schnittstellen Web Map Tile Service (WMTS) und Web Map Service (WMS) zugänglich gemacht. Das Kartenbild für die verschiedenen Zoomstufen/Maßstäbe wird mithilfe von umfangreichen Formatierungsangaben in XML-basierten Konfigurationsdateien beschrieben. Unter anderem werden hier die Datenquellen festgelegt und die Signaturen beschrieben, mit denen die Objekte in den verschiedenen Zoomstufen gezeichnet werden.

5 Kartengestaltung

Es wurde angestrebt, eine möglichst gut lesbare Kartengrafik für die verschiedenen Zoomstufen zu erzeugen, wobei die Kartenbilder ausschließlich aufgrund der Beschreibung der Kartengestaltung im Programmcode entstehen. Eine manuelle, interaktive kartographische Bearbeitung des Kartenbildes ist aufgrund der sehr großen Datenmenge nicht möglich. Nachstehend soll an einigen Beispielen gezeigt werden, wie die Erzeugung der gewünschten Kartengrafik insbesondere durch das Zusammenspiel von Mapnik und PostGIS unterstützt wird.

5.1 Map, Layer, Styles und Rules

Die Karte (Map) entsteht aus einer Abfolge von Layern. Jeder Layer verweist auf eine Datenquelle und auf den Style, mit welchem diese Daten in das Kartenbild gezeichnet werden. Ein Style definiert, wie die Objekte in den jeweiligen Zoomstufen dargestellt werden. Ein Style enthält mehrere Rules mit Zeichenanweisungen. Eine Rule kann mit einem Filter auf eine bestimmte Objektauswahl beschränkt werden. Mithilfe von Scaledenominatoren kann die Anwendung der Rule auf bestimmte Zoomstufen/Maßstäbe eingeschränkt werden. Jede Rule enthält einen oder mehrere Symbolizer für das Zeichnen von Texten, Linien, Polygonen, Symbolen usw. Mit umfangreichen Parametern wird genau beschrieben, wie die Signatur einer Linie, Fläche, eines Punktes oder einer Objektbeschriftung aussehen sollen. In neueren Mapnik-Versionen können die Werte dieser Parameter direkt aus Feldern der Datenbank übernommen werden.

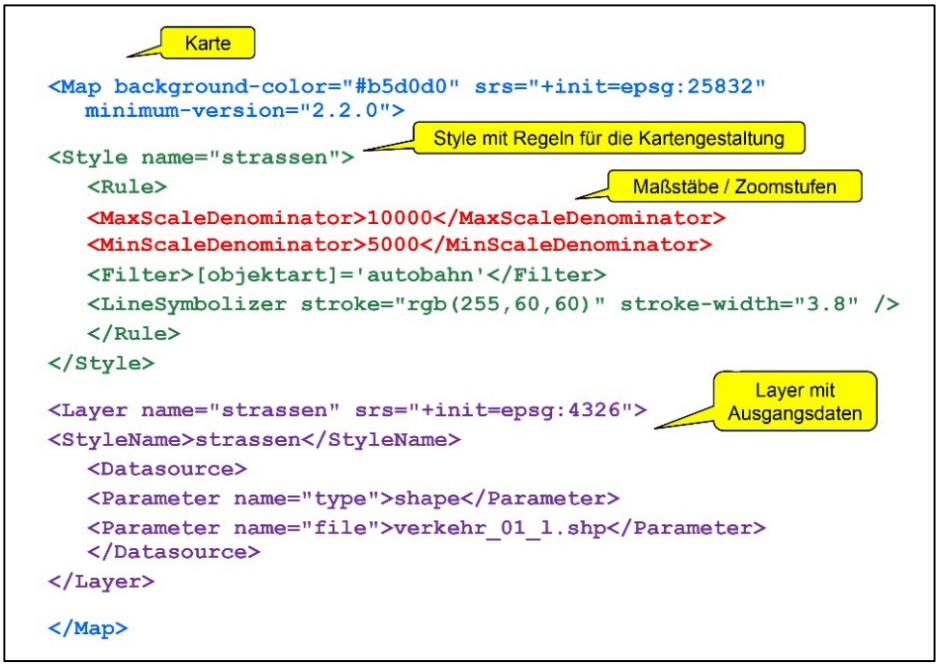


Abb. 5: Mapfile von Mapnik: Definition für eine Karte in UTM32. Ein Layer mit Quelldaten in EPSG:4326 verweist auf den Style „Strassen“. In einer Rule wird die Objektart „Autobahn“ gewählt und als einfache Linie gezeichnet. Die Rule wird nur für Maßstäbe von 1:5.000 bis 1:10.000 angewandt.

5.2 Zeichenreihenfolge und Kartenbild

5.2.1 Punkt, Linie, Fläche

Die darzustellenden Inhalte werden in Form von Layern aufgerufen und in das Kartenbild gezeichnet. Die Reihenfolge, in der die Objekte gezeichnet werden, bestimmt ganz westlich das entstehende Kartenbild. Für Flächen, Linien und Punktsignaturen gilt: Wer zuletzt ma(h)lt liegt ganz oben und ist sichtbar. Ein Element überzeichnet das vorhergehende. Beispiel: Zuerst wird die Landbedeckung und Landnutzung gezeichnet, danach erst die Seen und Flüsse, dann die Straßen usw. Das Kartenbild wird durch die Abfolge der Layer vorgegeben und durch die Reihenfolge, in welcher die Zeichenanweisungen (Rules) in einem Style aufgerufen werden. Innerhalb eines Layers ist die Reihenfolge entscheidend, mit der die Objekte eines Layers gezeichnet werden. Diese Sortierkriterien werden bei den SQL-Abfragen an die Datenbank mit dem Ausdruck „order by Attribut1, Attribut2“ bestimmt. In der SQL-Abfrage können auch weitere Hilfsattribute für die Sortierung berechnet werden. Die Objekte können dann nach Länge oder Flächengröße der Geometrie oder nach der Zeichenanzahl in einem Feld sortiert werden. Werden z. B. in einem Layer der Straßenbeschriftung mithilfe einer SQL-Abfrage die Straßenlängen berechnet und die Straßen nach der Widmung aufsteigend und nach der Straßenlänge absteigend sortiert, erfolgt anschließend zuerst die Beschriftung der Straßen höherer Widmung und innerhalb einer Klasse zuerst die der langen Straßen und erst danach die der kürzeren Straßen, sofern für deren Beschriftung noch Raum zur Verfügung steht.

5.2.2 Schriftplatzierung bei Mapnik

Im Falle der Textplatzierung kann Mapnik auf bereits platzierte Schrift Rücksicht nehmen und Schriftüberlagerungen verhindern. Dabei werden die Objekte in der Reihenfolge beschriftet, die durch eine SQL-Abfrage vorgegeben wurde. Der Layer und der Text, der zuerst platziert wird, blockiert diese Beschriftungsposition für alle weiteren Texte oder Symbole. Hier gilt im Gegensatz zu Grafikobjekten: Wer zuerst kommt ma(h)lt zuerst. Daher muss genau bestimmt werden, welche Objektbeschriftungen zu einem frühen Zeitpunkt unverrückbar platziert werden sollen und welche Schriften zu späteren Zeitpunkten sich dann noch den zur Verfügung stehenden Platz teilen müssen. Eine Beschriftung von Siedlungen nach der Einwohnerzahl würde z. B. mit dem Sortierkriterium „order by ewz DESC“ aufgerufen werden. Mapnik erlaubt innerhalb einer Zeichenanweisung/einer Rule auch die Definition von alternativen Beschriftungspositionen, Schriftgrößen oder Textvarianten. Diese verschiedenen Varianten werden dann während des Renderns in der vorgegebenen Reihenfolge durchprobiert, bis sich evtl. eine Beschriftungsmöglichkeit ergibt. Eine solche Aufzählung von Beschriftungsvarianten wäre z. B.:

1. Versuch: Schriftgröße 20pt, Name in Langform („Naturschutzgebiet Elbwiesen“)
2. Versuch: Schriftgröße 18pt, Name in Langform („Naturschutzgebiet Elbwiesen“)
3. Versuch: Schriftgröße 18pt, Name in Kurzform („NSG Elbwiesen“)
4. Versuch: Schriftgröße 16pt, Kurzform der Objektart („NSG“)

5.2.3 Kartenaufbau

Für den Kartenaufbau muss daher eine sinnvolle Reihenfolge festgelegt werden, in welcher die Kartenelemente gezeichnet werden. Nachstehende Überlegungen spielen dabei eine Rolle:

- Was stelle ich dar? Wie stelle ich es in den verschiedenen Zoomstufen dar?
- Wie erfolgt der Kartenaufbau? In welcher Reihenfolge zeichne ich die Objekte?
- Welche wichtigen Objekte haben Vorrang und sollen unbedingt dargestellt werden?
- Welche Objekte dürfen dagegen auch teilweise von anderen Objekten verdeckt oder überzeichnet werden?
- Welche Grafik muss zu einem späten Zeitpunkt gezeichnet werden, damit sie ganz oben liegt?
- Welche Texte müssen zu einem frühen Zeitpunkt platziert werden, damit sie noch einen freien Beschriftungsort finden? Welche Objekte dürfen ggf. entfallen, da sich kein geeigneter Platzierungsort findet bzw. weil der Originalort bereits belegt ist?
- Was soll an bestimmten Orten nicht dargestellt werden? Wie kann ich verhindern, dass Objekte an bestimmten Stellen platziert werden?
- Gibt es einen einzuhaltenden Mindestabstand zu nachfolgenden Objekten?
- Was soll an einer ganz bestimmten Position dargestellt werden (z. B. Einzelsignatur Kirche oder Ortspunkt)? Was kann aber auch in der Nähe des eigentlichen Objektes präsentiert werden (z. B. Siedlungsbeschriftung)?
- Soll der Name oder ein Symbol mehrfach im Verlauf einer Linie wiederholt werden (z. B. Straßennummern, Gewässernamen)? In welchen Abständen soll der Name/das Symbol wiederholt werden?
- Wie ist der maximal erlaubte Winkel zwischen zwei Zeichen, damit im Verlauf der Straße platzierte Text eines Straßennamens noch gut lesbar ist?

Aus diesen und vielen weiteren Überlegungen ergibt sich dann die Festlegung der Reihenfolge, in der die Objekte dargestellt werden.

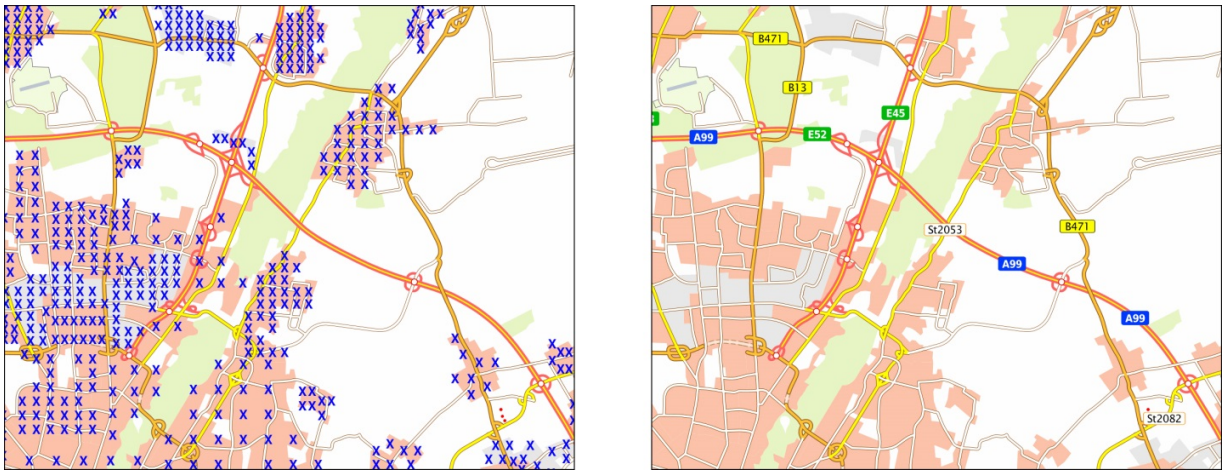


Abb. 6: Transparente Zeichen (hier in blau, links) verhindern das Platzieren von Straßennummern innerhalb der Siedlungsflächen (rechts)

5.2.4 Verhindern der Darstellung an bestimmten Orten

Häufig ist es erforderlich, das Platzieren von Symbolen oder die Beschriftung von Objekten an bestimmten Orten zu verhindern. So wird z. B. keine Darstellung erwünscht im Falle von:

- Ortsnamen auf Autobahnkreuzen, Anschlussstellen
- Straßennummern über kleinen Siedlungsflächen, Ortspunkten oder Straßenkreuzungen
- Gewässernamen auf Straßen/Bahnen
- Einbahnstraßenpfeile auf kreuzenden Brücken einer andern Straße oder Bahn
- Höhenlinienzahlen in Gewässern, auf Gebäuden, über Straßen, Eisenbahnen, Grenzen

Mapnik erlaubt hier die Platzierung von unsichtbaren Textzeichen (z. B. „X“) mit der Textfarbe „transparent“ über diesen unerwünschten Orten. Diese nicht sichtbare Schrift verhindert anschließend die Platzierung weiterer Texte oder von Symbolen an diesen Stellen.

5.3 Verwendung von PostGIS-Funktionen

Bei der Verwendung von PostGIS-Datenbanken als Datenquelle werden in Mapnik die darzustellenden Objekte eines jeden Layers mithilfe von SQL-Skripten ausgewählt. Mapnik erweitert diese SQL-Skripte während der Laufzeit derart, dass unmittelbar vor dem Rendern ein Ausschnitt in der Größe der zu rendernden Kachel aus der PostGIS-Datenbank selektiert und für das Rendern bereitgestellt wird. Diese SQL-Skripte können auch alle Datenbank- und GIS-Operationen ausführen, welche eine PostGIS-Datenbank bietet. Dazu zählen z. B.

- Umbenennen von Attributen oder das Erzeugen von neuen Attributen
- Bearbeitung von Strings (z. B. durch das Ersetzen von Textteilen)
- Formatierung von Texten (z. B. Zeilenumbrüche)
- Generalisierung von Geometrien
- Klassenbildung
- Anwendung von Aggregat-Funktionen (Zusammenfassung von Objekten)
- Sortieren von Objekten nach Geometriemerkmalen oder Attributen

Neben der Verwendung von Standard-Funktionen können aber auch selbst geschriebene PostGIS-Funktionen für die Lösung bestimmter Aufgaben verwendet werden. Hier einige Beispiele für den Einsatz solcher Funktionen in der Mapnik-Kartenbeschreibung:

Umbruch von Siedlungsnamen je nach Länge und Schreibweise
Abhängig von der Gesamtlänge des Namens, der Länge der Teilnamen und der Art der Trennung zwischen den Namensbestandteilen erfolgen ein oder mehrere Umbrüche.

Ohne Umbruch	**Mit Zeilenumbruch**
Bruch unter dem Deich	Bruch unter dem Deich
Heinz-Dieter-Keller-Platz	Heinz-Dieter- -Keller-Platz
Kreba-Neudorf - Chrjebja-Nowa Wjes	Neu-Drauschkowitz - - Nowe Družkecy

Funktion zum Ermitteln von abgekürzten Straßennamen

Um auch die Beschriftung kurzer Straßen zu ermöglichen, werden abgekürzte Bezeichnungen der Straßennamen ermittelt. Zuerst erfolgt ein Platzierungsversuch in der Langform, dann erst in der Kurzform.

Langform	Kurzform
Bahnhofstraße	Bahnhofstr.
Zur alten Mühle	Z.a. Mühle
Pfarrer-Heinrich-von-Solre-Straße	Pfr.-Heinrich-v.-Solre-Str.

Beschriften von Hausnummern mit Abstand zur Straße

Die Punkte der amtlichen Hauskoordinaten befinden sich häufig so dicht an einer Straße, dass nach dem Platzieren des Straßennamens kein Platz für die Beschriftung einer Hausnummer bleibt. Für das Beschriften der Hausnummern wird daher der Winkel des Punktes zum nächstgelegenen Punkt einer Straße ermittelt. Aufgrund dieses Winkels wird die Schriftposition festgelegt, mit welcher die Hausnummer relativ zum Bezugspunkt/zur Hauskoordinate platziert wird. Zum Beispiel führt ein ermittelter Azimutalwinkel von 60° zu einer Schriftposition im NW des Bezugspunktes. Die Hausnummer wird so mit einem gewissen Abstand von der Straßenachse platziert. Da zuvor die Straßennamen im Straßenraum platziert wurden, können jetzt aufgrund der neuen Schriftposition wesentlich mehr Hausnummern platziert werden. Platzierungskonflikte zwischen Straßennamen und Hausnummer werden so weitgehend vermieden.

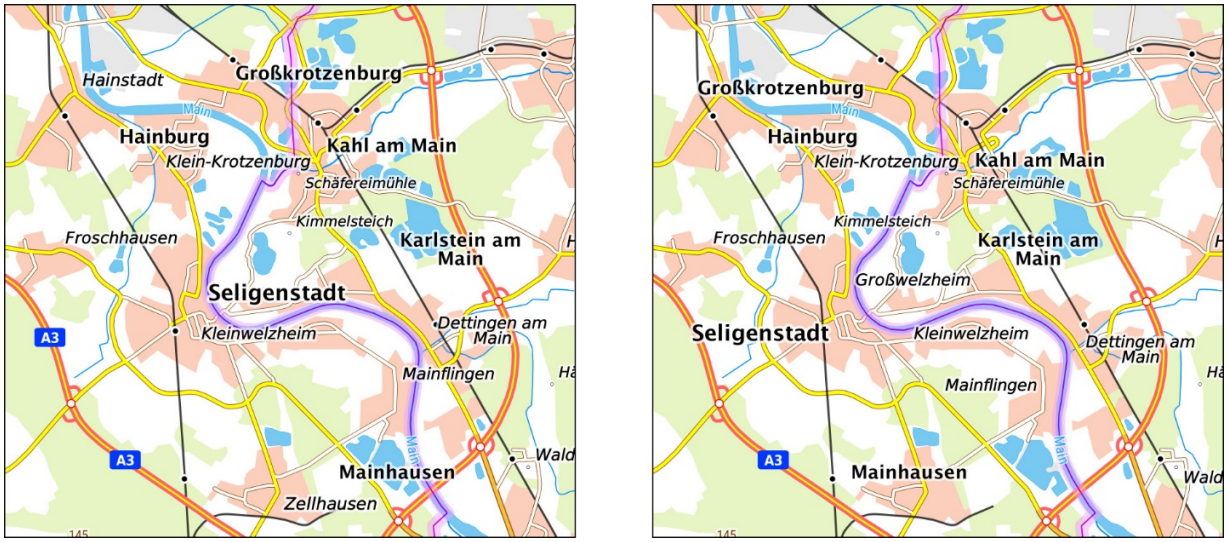


Abb. 7: Platzierung der Ortsnamen ohne Berücksichtigung der Lage des Ortes relativ zur Landesgrenze und zum Fluss (links). Berechnung des Azimutalwinkels der Ortspunkte zum nächstgelegenen Punkt der Landesgrenze oder des Flusses und Platzierung der Namen unter Anwendung des berechneten Winkels (rechts).

Abb. 8: TopPlus-Web-Open (Zoomstufe 8): Bevorzugte Platzierung der Namen von an der Küste gelegenen Siedlungen im Meer

Das gleiche Prinzip wurde auch angewandt für die:

- Beschriftung von Ortsnamen auf der „richtigen Seite“ des Flusses, der Verwaltungsgrenzen oder von Autobahnen (siehe Abb. 7)
- Platzierung der Namen von Küstenorten im Meer (siehe Abb. 8)
- Beschriftung von POIs und die Beschriftung öffentlicher Gebäude möglichst auf der „richtigen“ Straßenseite

6 Massenhaftes Rendern aller Kacheln

Für die Verfahrensentwicklung und die hochperformante Durchführung des Renderns wurden Tools auf der Basis des Microsoft.NET Framework implementiert. Diese erlauben die benutzerfreundliche Auswahl des zu erzeugenden Gebietes, der gewünschten Produktart, Zoomstufe und weiterer Parameter des Herstellungsprozesses. Mit der Skriptsprache Python wird Mapnik gesteuert und der Prozess des Renderns zur Erstellung eines kompletten Kachelarchives gestartet. Dieses Render-Programm erlaubt auch die Überwachung des Rendervorgangs und protokolliert evtl. auftretende Fehler. Es erfolgt eine Qualitätskontrolle, Prüfung auf Vollständigkeit und auf Fehler in den PNG-Dateien. Auf diese Weise werden auch Dateifehler einzelner Kacheln (Datei ist vorhanden, lässt sich aber nicht öffnen) gefunden. So kann nach einem ersten Renderdurchgang der Prozess für noch fehlende Kacheln erneut ausgeführt werden.

6.1 Bereitstellung der Renderergebnisse

Das Rendern wird mit Großkacheln in der Größe 2048×2048 Pixel durchgeführt. Diese Großkacheln werden anschließend in kleine Kacheln 256×256 Pixel zerschnitten und in die gewünschte Verzeichnisstruktur verteilt. In einem getrennten Vorgang erfolgt dann pro Zoomstufe die Speicherung der endgültigen Kacheln in SQLite-Datenbanken nach der MBTiles-Spezifikation. Im Falle der Präsentationsgrafiken in hoher Auflösung erfolgt das Rendern in Kacheln mit einer Größe von 8000×8000 Pixel. Danach erfolgt mithilfe von GDAL das Erstellen von GeoTiffs aus den gerenderten PNG-Dateien.

7 Ergebnisse und Ausblick

Die Bundesländer Brandenburg und Rheinland-Pfalz werden ab 2018 ihre Geobasisdaten ebenfalls für die Verwendung in der TopPlus-Web-Open bereitstellen. Es wäre zu wünschen, dass sich auch weitere Bundesländer diesem Schritt anschließen würden, so dass Deutschland in einer offenen Webkarte vollständig mit amtlichen Daten repräsentiert wird.
Zukünftig soll auf Basis der im TopPlus-Verfahren unterstützten Druckausgabe ein Map-on-Demand-Dienst zum Abruf von Präsentationsgrafiken (Rasterdaten, PDF-Dokumente) in Druckqualität via Internet realisiert werden. Hierzu wird serverseitig auf Basis der vektoriellen Originaldaten die gewünschte Präsentation erstellt und an den Client zum Druck übersandt. Auch die Daten des Erdbeobachtungsprogramms Copernicus der Europäischen Union können zukünftig in den Kartendarstellungen des TopPlus-Verfahrens verwendet werden. Erste Untersuchungen hierzu liegen bereits vor (z. B. Erstellung von Satellitenbildkarten).
Die Produkte aus dem TopPlus-Verfahren werden heute von vielen Bundeseinrichtungen erfolgreich verwendet. Vom BKG werden die Präsentationsgrafiken in hoher Auflösung umfassend für die Herstellung von Sonderkarten eingesetzt. Ob für den G7-Gipfel auf Schloss Elmau 2015 oder den G20-Gipfel in Hamburg 2017 – eine Vielzahl von Sonderkarten wurde vom BKG auf der Grundlage der TopPlus-Produkte für die Einsatzplanung dieser Großereignisse bereitgestellt. Im BKG wird das Verfahren heute intensiv zur Erfüllung des gesetzlichen Auftrags zur Aufbereitung, Aktualisierung und Bereitstellung von orts- und raumbezogenen Daten eingesetzt.

8 Literaturverzeichnis

KUTTERER, H. & KUNZ, P., 2015: Karten nach Maß mit dem TopPlus-Verfahren. gis.Business, 1/2015, 28-30.

KUNZ, P., 2014: TopPlus – von der detaillierten Stadtkarte bis zur europaweiten Übersichtskarte. Kartographische Nachrichten, 2/2014, 59-67.

KUNZ, P., 2012: Produktion topographischer Webkarten aus amtlichen Geobasisdaten. Tagungsband FOSSGIS 2012, 43-54.

8.1 Web-Adressen (Zugriff 01/2017)

[URL 1] TopPlus-Web-Open (Beschreibung): http://www.bkg.bund.de/TopPlusOpen

[URL 2] TopPlus-Web-Open (Webanwendung):
http://sg.geodatenzentrum.de/web_bkg_webmap/applications/bkgmaps/minimal.html?CRS=EPSG:3857

[URL 3] TopPlus-Web-Open (Datenquellen):
http://sg.geodatenzentrum.de/web_public/Datenquellen_TopPlus_Open.pdf

[URL 4] Datenlizenz Deutschland – Namensnennung – Version 2.0:
https://www.govdata.de/dl-de/by-2-0

[URL 5] Bundesamt für Kartographie und Geodäsie (BKG): http://www.bkg.bund.de

[URL 6] Dienstleistungszentrum des BKG (DLZ): http://www.geodatenzentrum.de

[URL 7] Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV): http://www.adv-online.de

[URL 8] Open-Data-Portal Deutsche Bahn AG: http://data.deutschebahn.com/

[URL 9] Wasserstraßen- und Schifffahrtsverwaltung des Bundes: https://www.wsv.de/

[URL 10] EuroGeographics, Produkte: http://www.eurogeographics.org/products-and-services

Vektordaten

[URL 11] Digitales Basis-Landschaftsmodell (Basis-DLM): http://www.adv-online.de/AdV-Produkte/Geotopographie/Digitale-Landschaftsmodelle/Basis-DLM/

[URL 12] Amtl. Hauskoordinaten (HK-DE): http://www.adv-online.de/AdV-Produkte/Liegenschaftskataster/Amtliche-Hauskoordinaten/

[URL 13] Amtl. Hausumringe (HU-DE): http://www.adv-online.de/AdV-Produkte/Liegenschaftskataster/Amtliche-Hausumringe/

[URL 14] Amtl. 3D-Gebäudemodelle LoD1 (LoD1-DE): http://www.adv-online.de/AdV-Produkte/Weitere-Produkte/3D-Gebaeudemodelle-LoD/

[URL 15] BKG, Open Data:
http://www.geodatenzentrum.de/geodaten/gdz?l=down_opendata_top

[URL 16] Digitales Landschaftsmodell 1:250.000 (DLM250):
http://www.geodatenzentrum.de/geodaten/gdz?l=down_dlm250

[URL 17] Digitales Landschaftsmodell 1:1.000.000 (DLM1000):
http://www.geodatenzentrum.de/geodaten/gdz?l=down_dlm1000

[URL 18] Verwaltungsgebiete 1:250.000 (VG250):
http://www.geodatenzentrum.de/geodaten/gdz?l=down_vg250_0101

[URL 19] Geographische Namen 1:250.000 (GN250):
http://www.geodatenzentrum.de/geodaten/gdz?l=down_gn250

[URL 20] Corine Land Cover: https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-corine

[URL 21] OpenStreetMap-Daten: http://download.geofabrik.de/

[URL 22] GeoNames: http://www.geonames.org/

[URL 23] Natural-Earth-Daten: http://www.naturalearthdata.com

Geländemodelle
[URL 24] Digitales Geländemodell Gitterweite 10 m (DGM10):
http://www.geodatenzentrum.de/docpdf/dgm10.pdf
[URL 25] European Digital Elevation Model (EU-DEM):
https://land.copernicus.eu/pan-european/satellite-derived-products/eu-dem
[URL 26] Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010):
http://topotools.cr.usgs.gov/gmted_viewer/
[URL 27] SRTM15_PLUS V1 global bathymetry:
http://topex.ucsd.edu/WWW_html/srtm30_plus.html

Software
[URL 28] PostgreSQL: http://www.postgresql.org, PostGIS: http://postgis.refractions.net
[URL 29] https://imposm.org/
[URL 30] osm2pgsql: https://github.com/openstreetmap/osm2pgsql
und http://wiki.openstreetmap.org/wiki/Osm2pgsql
[URL 31] Safe FME: http://www.safe.com
[URL 32] Change: https://www.ikg.uni-hannover.de/index.php?id=software
[URL 33] GDAL und OGR: http://www.gdal.org
[URL 34] Mapnik: http://mapnik.org

ETRS89/UTM-Umstellungsstrategie in Bayern für das Liegenschaftskataster – Umformungsverfahren und Möglichkeiten der Umsetzung

CLEMENS GLOCK[1] & HENNING BIRKENBEUL[2]

Zusammenfassung: In Bayern steht nach der flächendeckenden Einführung von ALKIS für das Liegenschaftskataster die Einführung des einheitlichen europäischen Bezugssystems ETRS89/UTM bevor. Mit der Einführung von ETRS89/UTM im Liegenschaftskataster wird zeitlich der amtliche Bezugssystemwechsel vollzogen. Das Projekt KanU (Kataster nach UTM) schafft die Grundlage, dass die Produkte des Liegenschaftskatasters in dem neuen Bezugssystem ETRS89/UTM abgegeben werden können. Es wurden verschiedene Umformungsmethoden für die unterschiedlichen Objektbereiche entwickelt. Als wichtigstes Überführungsverfahren wurde in dem Projekt KanU das neuartige Verfahren „Umformung durch Ausgleichung nach dem Ortra-Ansatz“ entwickelt. Mit der Umstellung stehen somit Fachanwender vor der Herausformung, große Datenmengen zu transformieren. Der Kundentransformationsansatz NTv2-KanU zur Überführung von Geofachdaten mit Katasterbezug wurde dafür herstellerneutral entwickelt. Transformationssoftwareprodukte oder GIS-Produkte können diesen Kundentransformationsansatz nutzen. Mit einem Transformationssoftwareprodukt und einem GIS-Produkt wurden Erfahrungen gesammelt.

1 Einleitung

In Bayern steht dem Nutzer von Geobasisdaten, dem Fachanwender, dem freien Ingenieurberuf und dem amtlichen Vermessungswesen Ende des Jahres 2018 ein Wechsel des amtlichen Bezugssystems bevor. Zum Jahreswechsel 2018/2019 wird in Bayern der Übergang vom bisherigen nationalen Bezugssystem Deutsches Hauptdreiecksnetz (Abkürzung DHDN) mit dem Gauß-Krüger-Abbildungssystem zum Europäischen Terrestrischen Referenzsystem 1989 (ETRS89) mit der Universalen Transversalen Mercator Projektion in der Zone 32 (Abkürzung UTM32) als neues Abbildungssystem erfolgen.
Der Bezugssystemwechsel wird bei der Bayer. Vermessungsverwaltung (BVV) sowohl durch die Überführung des amtlichen Liegenschaftskatasters nach ETRS89/UTM als auch durch den formalen Abschluss der Umstellung der Produktion und des Vertriebs der amtlichen Geobasisdaten der Landesvermessung nach ETRS89/UTM markiert.
Damit endet im gesamten Gebiet der Bundesrepublik Deutschland nach mehreren Jahrzehnten die Zeit des Gauß-Krüger-Systems als amtliches Koordinatensystem.
Die Einführung des neuen amtlichen europäischen Bezugssystems ETRS89/UTM zur Realisierung eines einheitlichen europäischen Referenzrahmens stellt die Vermessungsverwaltungen der Bundesländer vor große Herausforderungen. Besonders bei der Umstellung des Liegenschaftskatasters nach ETRS89/UTM sind hohe Genauigkeitsanforderungen zu erfüllen.

[1] Landesamt für Digitalisierung, Breitband und Vermessung, Projekt KanU, Alexandrastraße 4, D-80538 München, E-Mail: clemens.glock@ldbv.bayern.de
[2] con terra GmbH, Martin-Luther-King-Weg 24, D-48155 Münster, E-Mail: h.birkenbeul@conterra.de

Der Raumbezugswechsel bringt folgende Vorteile, er

- ermöglicht eine einheitliche Basis für Geodaten und Karten in Europa,
- unterstützt eine großräumige und länderübergreifende Nutzung der Geodaten,
- vereinfacht die Zusammenführung von Geodaten aus verschiedenen Quellen,
- bildet die Grundlage für eine europaweite Geodateninfrastruktur (INSPIRE) und
- lässt GNSS-Messungen im amtlichen System zu.

Zum Zweck der Überführung der Geobasisdaten des Liegenschaftskatasters ins Bezugssystem ETRS89/UTM wurde 2012 am damaligen Landesamt für Vermessung und Geoinformation (jetzt: Landesamt für Digitalisierung, Breitband und Vermessung, LDBV) in Bayern das Projekt KanU „Kataster nach UTM“ eingerichtet (HAMPP & GLOCK 2017).
Für den zweiten Bereich der amtlichen Geodaten, die Geobasisdaten der Landesvermessung (Geotopographie), wird in diesem Jahr die mehrjährige Umstellung beim LDBV abgeschlossen. Für die Bezugssystemumstellung wurden bereits durchgängige Produktionsketten in ETRS89/UTM aufgebaut (z. B. in der Luftbildproduktion). Bei der Umstellung von Vektor- und Rasterdaten der Landesvermessung sind Genauigkeitsanforderungen ausreichend, die z. B. in Form eines NTv2-Verfahrens die Bundeseinheitliche Transformation für ATKIS (BeTA2007) zur Verfügung stellt.
Nach jetziger Planung wird die Überführung des Liegenschaftskatasters während eines ca. zweiwöchigen Umstellungszeitraums zum Jahreswechsel 2018/2019 erfolgen. Anschließend steht das Amtliche Liegenschaftskatasterinformationssystem (ALKIS) für die Produktion und die Abgabe von Produkten im neuen Bezugssystem ETRS89/UTM zur Verfügung.
Fachanwender müssen sich auf den Bezugssystemwechsel nach ETRS89/UTM einstellen. Geofachdaten, die einen Bezug zu den amtlichen Geobasisdaten haben, müssen für eine Kombination mit den Geobasisdaten in das neue System überführt werden. Die BVV stellt hierzu Transformationsmodelle bereit. Die katastergenaue Transformationsdatei steht nach Abschluss der BVV-internen Transformation voraussichtlich im Januar 2019 zur Verfügung.
Der Artikel konzentriert sich auf die Umformungsansätze zur Überführung der verschiedenen Objektbereiche von ALKIS.
Das LDBV hat mit dem Ortra-Verfahren (Ortra steht für „Orthogonale Transformation“) ein numerisches Homogenisierungsverfahren, basierend auf einer Ausgleichung entwickelt. In der Ausgleichung für das Ortra-Verfahren sind elastomechanische Analogien vorhanden.
Der Beitrag beschreibt die Ausgleichungen detailliert, die innerhalb des Projekts KanU („Kataster nach UTM“) entwickelt wurden. Diese werden zum Bezugssystemwechsel für die Überführung der ALKIS-Objektbereiche Flurstücke, Gebäude und Bauwerke eingesetzt. Zusätzlich wird ein weiteres Ortra-Ausgleichungsverfahren zur Erstellung der Gitterpunkte des Kundentransformationsansatzes NTv2 verwendet.
Nach Darstellung der Ausgleichungsansätze für Flurstücke und Gebäude (Bauwerke werden weitgehend wie Gebäude behandelt und deshalb nicht näher vorgestellt) wird anschließend der Kundentransformationsansatz NTv2 mit der engmaschigen NTv2-Gitterdatei BY-KanU dargestellt. Er kann zur Überführung von Geofachdaten mit Katasterbezug genutzt werden. Der not-

wendigen Kürze des Beitrags geschuldet, wird auf die Darstellung der Umformung durch Nachbarschaft verzichtet, die die übrigen Objektbereiche von ALKIS umformt.

2 Transformation und Homogenisierung des Liegenschaftskatasters durch das Ortra-Verfahren

Das Ortra-Verfahren wurde als Ausgleichungsverfahren konzipiert und ab 2013 am LDBV für die einzelnen Objektbereiche entwickelt. Es führt während der Umformung eine Homogenisierung der Daten durch. Im alten Bezugssystem noch vorhandene Netzspannungen werden durch eine Minimierung nach dem Prinzip der kleinsten Quadrate reduziert.
Das Verfahren ist mit den Kartenhomogenisierungsverfahren der 1990er Jahre verwandt, die für eine effiziente Datenakquisition manuell digitalisierte Beobachtungen mit Hilfe der Ausgleichung über Verknüpfungspunkte konsistent in ein übergeordnetes Koordinatensystem überführen konnten (Hampp, 1992).
Das Prinzip des Ortra-Verfahrens ist Verbesserungsgleichungen mit Hilfe einer Vielzahl von künstlichen Beobachtungen, sogenannten Pseudobeobachtungen, für die Grundrissgeometrie des Objektbereichs zu definieren. Die folgende Abbildung zeigt die Umsetzung des Prinzips am Beispiel der Flurstücke:

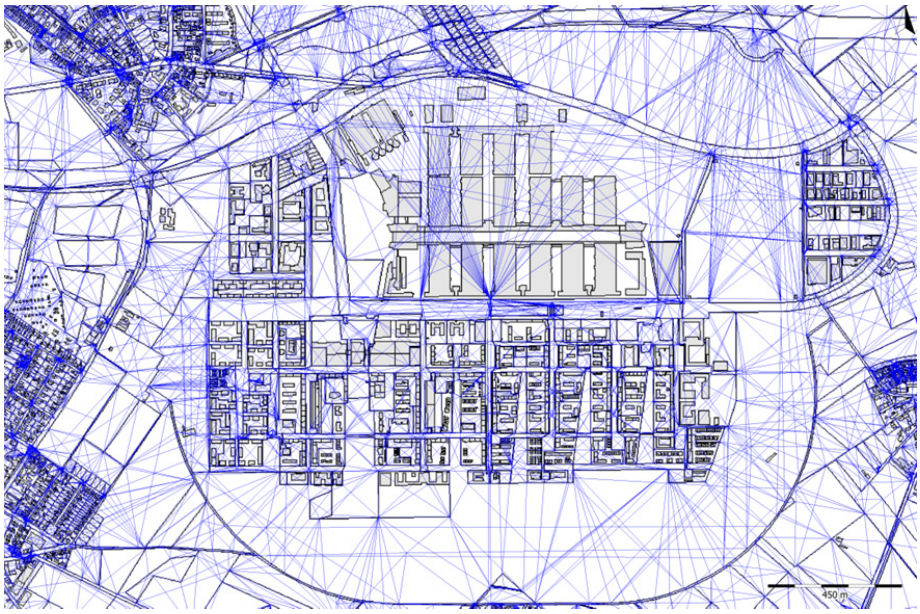


Abb. 1: Vernetzung des Flurstücks-Grundrisses durch künstliche Beobachtungen (Pseudobeobachtungen), München, Messegelände Riem

Ausgangspunkt für die Vernetzung ist der in ALKIS vorliegende Grundriss im Koordinatenausgangssystem Gauß-Krüger mit seinen punkt-, linien- und flächenförmigen Objekten.
Nach dem Aufbau eines geodätischen Netzes ausschließlich aus Grundrisselementen im alten Bezugssystem Gauß-Krüger wird für den Zweck der Umformung ins ETRS89/UTM-System das

gebildete geodätische Netz auf die identischen Punkte, die sowohl GK- als auch UTM-Raumbezug haben, gezogen.

2.1 Flurstücksausgleichung

Zuerst erfolgt die Beschreibung der Ausgleichung für den ALKIS-Objektbereich „Angaben zum Flurstück".
Der ALKIS-Objektbereich „Angaben zum Flurstück" wird, da die Flurstücksdaten rechtlich die höchste Relevanz besitzen, als erster Objektbereich nach UTM umgeformt. Nach erfolgter Überführung bilden die Flurstücke das bereits homogenisierte „Rahmenwerk" in ETRS89/UTM für weitere Ausgleichungen. Nachfolgende Ausgleichungen können dann die in UTM vorliegenden Flurstücksgrenzpunkte als zusätzliche identische Punkte nutzen.
Zwei Arten von Beobachtungen werden für die Flurstücke definiert:

2.1.1 1. Art: Flurstücksgrenzbeobachtungen und Beobachtungen entlang der Winkelhalbierenden

Für jeden Flurstücksgrenzpunkt wird ein Ortra-Kleinsystem aufgebaut, in dem Beziehungen zu den Flurstücksgrenzpunkten der unmittelbaren Nachbarschaft definiert werden. Im lokalen Koordinatensystem, das bei Überführung nach UTM gedehnt, gestaucht und i. d. R. gedreht werden kann, und damit winkelerhaltende Eigenschaften besitzt, wird durch Definition von Koordinatendifferenzen zu den Nachbarpunkten eine lokale Metrik aufgebaut, die nach UTM übertragen wird. Die definierten Nachbarschaftsbeziehungen (Strecken- und Winkelbeziehungen) werden bei der Ausgleichung dann so gut wie möglich erhalten.
Die Grundformel von Ortra mit Beziehungen zwischen GK- und UTM-Koordinaten im lokalen Koordinatensystem lautet:

$$E_2 = E_1 + o_1(x_2 - x_1) + a_1(y_2 - y_1)$$
$$N_2 = N_1 + a_1(x_2 - x_1) - o_1(y_2 - y_1)$$

Die von der Ausgleichung verwendeten, dazugehörigen Verbesserungsgleichungen lauten:

$$0 + \hat{v}_{E_2} = \hat{E}_1 + \hat{o}_1(x_2 - x_1) + \hat{a}_1(y_2 - y_1) - \hat{E}_2$$
$$0 + \hat{v}_{N_2} = \hat{N}_1 + \hat{a}_1(x_2 - x_1) - \hat{o}_1(y_2 - y_1) - \hat{N}_2$$

Eine Pseudobeobachtung wird dann zu einem benachbarten Flurstücksgrenzpunkt aufgebaut, wenn eine gemeinsame Flurstücksgrenze beide Punkte verbindet (siehe Abbildung 2). Auf diese Weise definieren sich ein Vorblick vom Ausgangspunkt zu dem entlang des Flurstückpolygons im Uhrzeigersinn am nächsten gelegenen Punkt und ein Rückblick zu dem zurückliegenden Grenzpunkt.
Die innerhalb eines Gebiets für alle Flurstücksgrenzpunkte gebildeten Beobachtungen entlang der Flurstücksgrenzen führen zu einer exakten Nachbildung des Grundrisses. Es ist jedoch erforderlich, dass der Grundriss noch zusätzlich durch weitere Beobachtungen versteift wird. Zu diesem Zweck werden Flurstücksquerverstrebungen zu dem gegenüberliegenden Teil des Flurstücks gebildet.

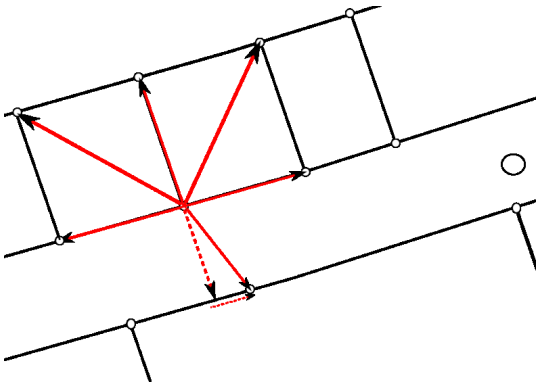


Abb. 2: Aufbau der Querverstrebung, Definition der Beobachtung in der Winkelhalbierenden zwischen zwei Flurstücksgrenzen

Für die Flurstücks-Querverstrebung wird eine Winkelhalbierende in die von diesem Punkt abgehenden Grenzen definiert. Diese schneidet, i. d. R. auf der gegenüberliegenden Seite des Flurstücks, eine Flurstücksgrenze. Auf der geschnittenen Flurstücksgrenze wird dann derjenige Grenzpunkt als Endpunkt der Beobachtung ausgewählt, der näher zum Schnittpunkt liegt. Durch diese Querverstrebungen wird einer geometrischen Entartung bei langgestreckten Flurstücken (z. B. Flussläufe, Straßen) entgegengewirkt.

2.1.2 2. Art: Beobachtungen zu den identischen Punkten

In Bayern existieren ca. 880.000 mit GNSS gemessene identische Punkte. Der überwiegende Teil davon sind Katasterfestpunkte. Als identische Punkte finden auch Flurstücksgrenzpunkte, die mit GNSS gemessen wurden, Verwendung. Diese identischen Punkte sind für den Ortra-Ansatz der übergeordnete, in UTM vorgegebene, homogene Koordinatenrahmen, in dem das in GK aufgebaute künstliche Netz eingespannt wird. Die Vernetzung wird bei der Umformung vom gegebenen lokal-bestanschließenden Bessel-Ellipsoid in das geozentrisch gelagerte, die GRS80-Dimensionen realisierende ETRS89-Ellipsoid gezogen und schließlich mit den identischen Punkten verknüpft. Auf diese Weise erfährt das Netz über jeden Knotennetzpunkt mechanische Spannungen, da eine verzerrungsfreie Übertragung zwischen verschiedenen Ellipsoiden geometrisch nicht möglich ist. Diese Spannungen führen in Verbindung mit dem Gesetz der minimalen Formänderungsenergie dazu, dass die einzelnen Ortra-Kleinsysteme einen UTM-Maßstab bzw. eine gute Näherung dieses Maßstabs annehmen, wenn lokale Netzspannungen gegeben sind. Zuletzt wird das auf diese Weise bereits auf UTM skalierte künstliche geodätische Netz auf dem neuen Bezugsellipsoid ETRS89 in Ellipsoidhöhe 0 m verankert.
Durch Minimierung der Netzspannungen bei dem Vorgang ergibt sich gleichzeitig eine Verbesserung der globalen Lage der Netzpunkte.
Folgende Pseudobeobachtungen werden für identische Punkte definiert:
Von einem identischen Punkt werden Beobachtungen zu allen Flurstücksgrenzpunkten definiert, die sich im Umkreis von 30 m befinden. Der Wert von 30 m für den Umkreis wurde im Projekt empirisch festgelegt. Falls sich innerhalb des Umkreises keine Flurstücksgrenzpunkte befinden, werden im umliegenden Flurstück, in dem sich der identische Punkt befindet, über alle Richtungen verteilt Beobachtungen zu den Punkten des Flurstücks generiert.

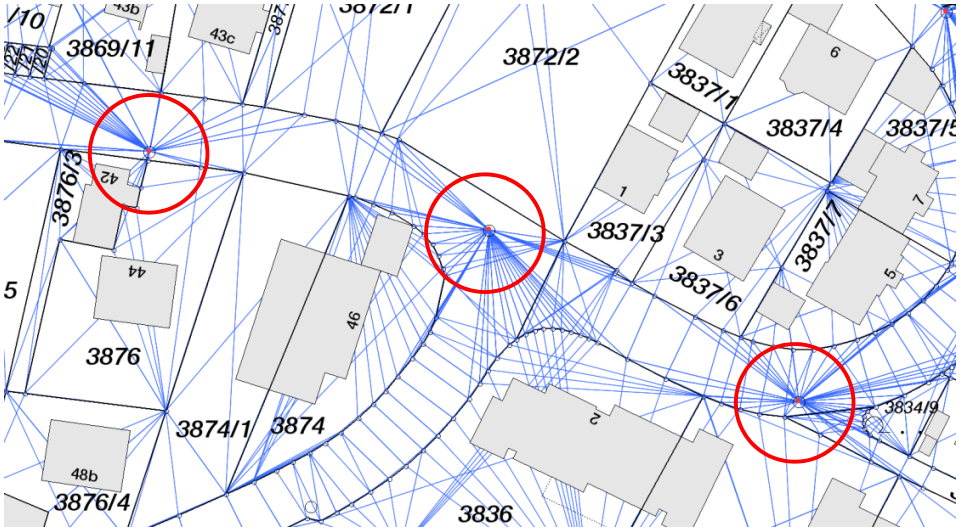


Abb. 3: Beobachtungsdefinition für die identischen Punkte

Die Verbesserungsgleichungen für die Beobachtung von identischen Punkten zu einem benachbarten Flurstücksgrenzpunkt sind bereits oben angegeben. Zusätzlich gehen für den identischen Punkt die Koordinaten als direkte Beobachtungen in die Beobachtungsgleichungen ein:

$$E_i + \hat{v}_{E_i} = \hat{E}_i$$

$$N_i + \hat{v}_{N_i} = \hat{N}_i$$

Mit den vorgestellten Beobachtungsgleichungen können die nach UTM zu transformierenden Flurstücksgrenzpunkte mit den identischen Punkten in Beziehung gebracht werden. Ein identischer Punkt behält bei der Ausgleichung seine gemessene Koordinate, wenn in den oben angegebenen direkten Beobachtungen für Koordinaten jeweils ein hohes Gewicht angesetzt wird. Das ganze Netz kann bei „unbeweglichen“ identischen Punkten in UTM durch das mathematische Hilfsmittel der Ausgleichung in den vorgegebenen ETRS89/UTM-Koordinatenrahmen fixiert werden.
Die einzelnen Pseudobeobachtungen sind zu elastisch wirkenden Einheiten zusammengefügt worden (siehe Abbildung 2), die sich nach physikalischen Gesetzen verformen. Eine wichtige Funktion spielt das bereits erwähnte Gesetz der minimalen Formänderungsenergie, das bei statischen Netzen, die sich im Gleichgewicht befinden, angewendet werden kann. Dieses Gesetz bildet die Grundlage, um den geforderten guten Geometrieerhalt bei gegebenen Netzspannungen beim Ortra-Ansatz zu erlangen.

Abb. 4: Aus Pseudobeobachtungen definiertes, mechanisch wirkendes „KanU“-Fachwerk

Über die dargestellten linearen Beziehungen zwischen Koordinaten im Ausgangssystem und Knotenpunkt-Koordinaten im Zielsystem entsteht ein kinematisch definiertes mechanisches Fachwerk, das z. B. auch über das Weggrößenverfahren der Statik berechnet werden kann, um elastische Eigenschaften des Ortra-Ansatzes genauer zu untersuchen.

2.2 Gebäudeausgleichung

Die Gebäudeausgleichung wird nach der Flurstücksausgleichung durchgeführt. Damit erfolgt eine Einpassung der Gebäude durch Ausgleichung in den durch das Ortra-System homogenisierten Flurstücksgrundrissrahmen. Die identischen Punkte für die Gebäude sind die bereits bei den Flurstücken genutzten Festpunkte des Objektbereichs „Angaben zum Netzpunkt“ und die in vorheriger Ausgleichung nach UTM32 umgeformten Flurstücksgrenzpunkte. Bei der Gebäudeausgleichung kommt die bereits vorgestellte Grundformel von Ortra zur Anwendung. Im Vergleich zur Flurstücksvernetzung wird die Vernetzung der Gebäude differenzierter vorgenommen. Zwischen drei verschiedenen Vernetzungsarten wird bei der Gebäude-Ausgleichung unterschieden:

- Gebäudeinnenvernetzung
- Gebäudeaußenvernetzung
- Gebäudezwischenvernetzung

Folgende Gründe waren maßgeblich für die Definition von verschiedenen Gebäudevernetzungen:
Durch das Katasterpunktfeld verursachte Netzspannungen in der Umgebung des Gebäudes sollen homogen durch eine vermittelnde Ausgleichung der Gebäude innerhalb des Flurstücks weiterverteilt werden, dabei soll jedoch innerhalb des Flurstücks die Form der Gebäude möglichst gut erhalten bleiben. Die Grenzabstände der Gebäude zu den Flurstücksgrenzen sind zu erhalten.

2.2.1 Gebäudeinnenvernetzung

Die Gebäudeinnenvernetzung ist identisch zur Flurstücksvernetzung. Die Beobachtungen in Richtung der Winkelhalbierenden führen zu einer Versteifung des Gebäudegrundrisses. Für jeden Gebäudepunkt definierte Vor- und Rückblicke modellieren den Gebäudegrundriss. Die Gebäudeinnenvernetzung hat zur Aufgabe, eine möglichst verzerrungsfreie Überführung der Gebäudeform durchzuführen.

Abb. 5: Gebäudeinnenvernetzung, Ausschnitt München, Messehallen Riem

2.2.2 Gebäudeaußenvernetzung

Durch die Gebäudeaußenvernetzung wird der Bezug des Gebäudes zum umgebenden Flurstück definiert. Die Pseudobeobachtung wird vom Gebäudepunkt in Richtung der nächsten Flurstücksgrenze gebildet. Dazu werden vom Gebäudepunkt die beiden anstoßenden Gebäudelinien nach außen verlängert, wenn kein Gebäude im Weg steht, bis die Verlängerung eine Flurstücksgrenze des umgebenden Flurstücks trifft.

Für eine gefundene nächste Flurstücksgrenze werden zwei Pseudobeobachtungen definiert. Eine Beobachtung geht vom Gebäudepunkt zum Anfangspunkt der gefundenen Flurstücksgrenze, die andere zum Endpunkt.

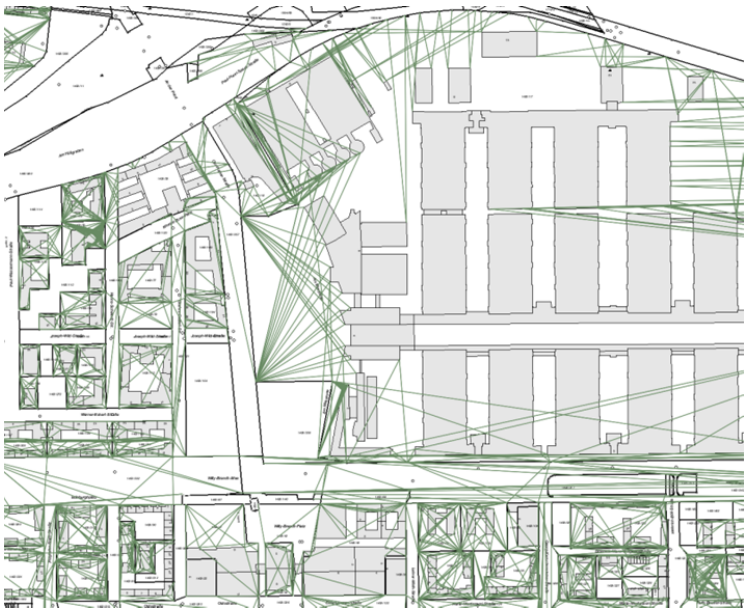


Abb. 6: Gebäudeaußenvernetzung, Ausschnitt München, Messehallen Riem

2.2.3 Gebäudezwischenvernetzung

Die Aufgabe der Gebäudezwischenvernetzung ist es, die Gebäudeabstände beim Übergang nach ETRS89/UTM so gut wie möglich zu übertragen. Die Gebäudezwischenvernetzung wird wie folgt definiert: Für jeden Gebäudepunkt werden im Umkreis von 30 m die nächstgelegenen Gebäudepunkte und die mit ihnen verbundenen, noch im Fangkreis liegenden Gebäudelinien ermittelt. Eine gefundene Gebäudelinie wird für die Definition von zwei Pseudobeobachtungen verwendet, wenn dabei das eigene Gebäude nicht geschnitten wird.

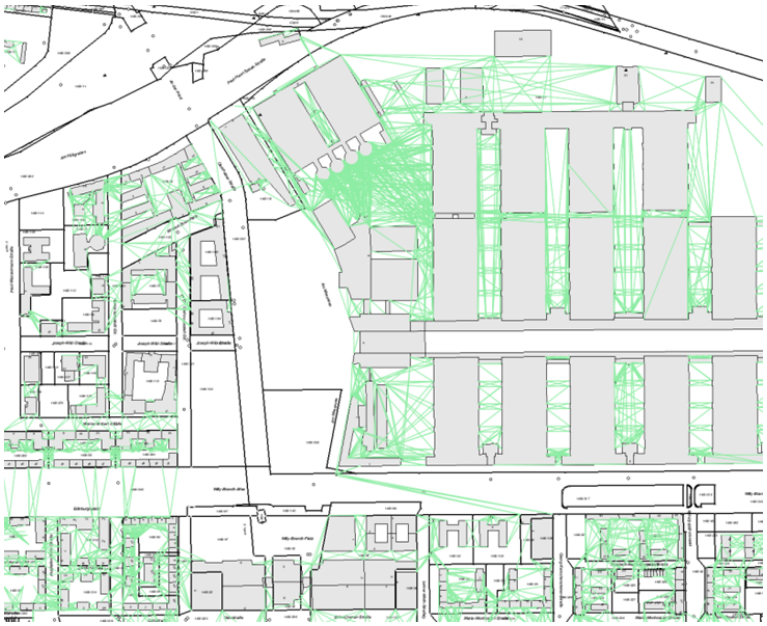


Abb. 7: Gebäudezwischenvernetzung, Ausschnitt München, Messehallen Riem

3 Transformation von Geofachdaten

Mit der Einführung des neuen amtlichen Systems ETRS89/UTM stellt sich für die Nutzer die Aufgabe, ihre **Geofachdaten in das neue System zu überführen**. Für die Umstellung der Geofachdaten vom alten Bezugssystem DHDN90/GK4 ins neue Bezugssystem ETRS89/UTM (und in umgekehrter Richtung) steht der NTv2-Ansatz „BY-KanU“ als amtliches Kunden-Transformationsmodell zur Verfügung. NTv2 ist eine gitterbasierte 2D-Transformationsmethode, die von der kanadischen Landesvermessung 1996 für Kanada definiert wurde. Mit ihr wird ein Bezugssystemwechsel von einem Quellsystem in ein Zielsystem bzw. umgekehrt vollzogen (WITTMANN 2015).

BY-KanU bezeichnet neben dem Kundentransformationsansatz auch die engmaschige NTv2-Datei, deren Gitterpunkte durch einen Ausgleichungsansatz aus den Ergebnissen der Ortra-Transformation der Flurstücke, Gebäude und Bauwerke abgeleitet werden.

Es ist geplant, die amtliche NTv2-Datei „BY-KanU“ zeitnah nach Abschluss der Arbeiten der ALKIS-Umstellung auf UTM den Fachkunden auf der Webseite der BVV zur UTM-Umstellung zur Verfügung zu stellen.

Bereits jetzt stehen vorläufige NTv2-Dateien, die vom Aufbau den späteren amtlichen NTv2-Dateien entsprechen, am selben Ort bereit. Damit können sich GIS-Dienstleister und Fachkunden mit dem Transformationsmodell vertraut machen und beispielsweise in ihrer GIS-Umgebung oder mit Open-Source-Transformationsansätzen (z. B. GDAL/OGR, proj4) Probetransformationen durchführen. Die derzeit bereitgestellten Daten dienen ausschließlich zu Testzwecken und entsprechen nicht dem späteren amtlichen Transformationsmodell.

Die Bayer. Vermessungsverwaltung stellt aktuelle Informationen zur ETRS89/UTM-Umstellung in Bayern auf der Internetseite www.ldbv.bayern.de/vermessung/utm_umstellung.html bereit.

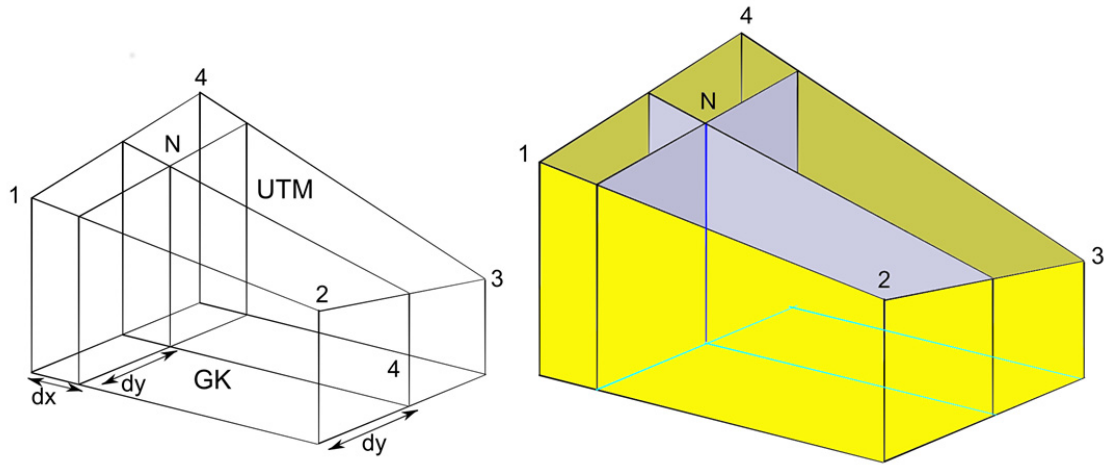


Abb. 8: Transformationsansatz NTv2-Gitter, UTM-Koordinaten des Punkts N über bilineare Interpolation aus den Ausgangswerten der Gitterpunkte berechnet

3.1 Zum Aufbau der NTv2-Datei BY-KanU

Die landesweite Transformationsvorschrift für die Transformation von Fachdaten mit Katasterbezug ist in der NTv2-Datei „BY-KanU“ vollständig abgelegt. Die Gebietsgrenzen werden mit außerbayerischen Gebietseckpunkten definiert, so dass sämtliche Landesteile von Bayern mit Shiftwerten abgedeckt sind. Die Gitterweite der NTv2-Datei beträgt eine geographische Bogensekunde (1“ = 1/3600 Grad), das sind in Bayern ca. 30 m x 30 m große Gittermaschen.
Bei der späteren amtlichen NTv2-Datei „BY-KanU“ werden für Gitterpunkte außerhalb Bayerns Shiftwerte aus der bundeseinheitlichen Transformation für ATKIS (BeTA2007) abgeleitet. Für innerbayerische Gitterpunkte wird folgende Verfahrensweise gewählt:
Für die Bestimmung der NTv2-Gitterpunkte wird ein Ausgleichungsansatz mit Ortra-Pseudobeobachtungen angewendet. Als Ausgangsdaten dienen alle nach UTM transformierten Punkte der Objektbereiche Flurstücke, Gebäude und Bauwerke. Zusätzlich kommen als Ausgangsdaten ca. 880.000 identische Punkte des Objektbereichs Angaben zum Netzpunkt hinzu.
Die NTv2-Gitterpunkte sind die Unbekannten der Ausgleichung. Die Ortra-Pseudobeobachtungen für einen NTv2-Gitterpunkt werden so definiert, dass in den inzidierenden vier NTv2-Gittermaschen Objektpunkte der genannten Objektbereiche gesucht werden, um mit ihnen die Beobachtungen zu bilden. Es werden wenigstens Beobachtungen zu fünf identischen Punkten von jedem Gitterpunkt aus erzeugt. Alle benachbarten NTv2-Gitterpunkte werden mit Beobachtungen verknüpft (Abbildung 9 Mitte).

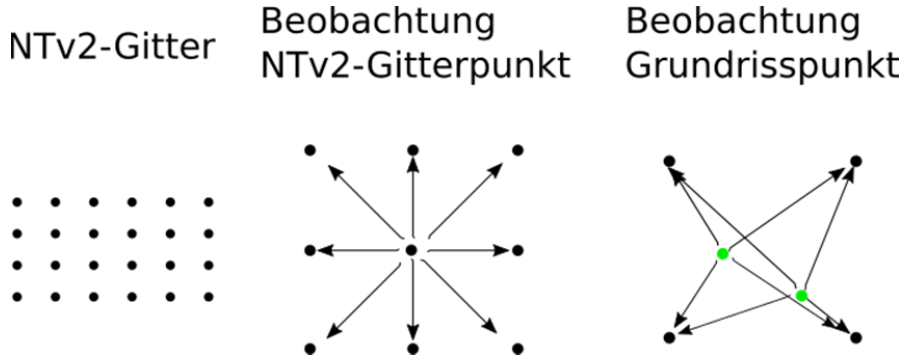


Abb. 9: Bildung der Pseudo-Beobachtungen für NTv2-Gitterpunkte

In die bekannten Ortra-Verbesserungsgleichungen fließen GK-Abszissen- und GK-Ordinatendifferenzen ein, die aus den GK-Koordinaten der Objektpunkte und den ableitbaren GK-Koordinaten der NTv2-Gitterpunkte gebildet werden.

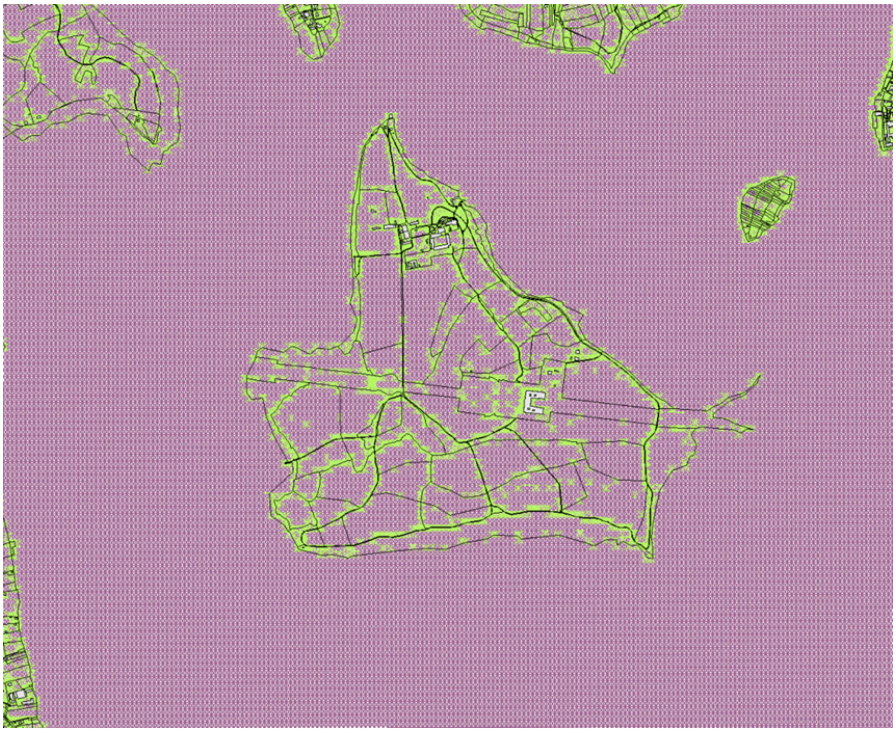


Abb. 10: Ausgleichung für die Bestimmung der NTv2-Gitterpunkte, Ausschnitt Insel Herrenchiemsee, Beobachtungsdefinition, Verknüpfung der Flurstücksgrenzpunkte mit dem NTv2-Gitter

Aus der mechanischen Betrachtungsweise entsteht ein in den Passpunktbereichen ausreichend verknüpftes und in passpunktfreien Bereichen steifes NTv2-Gitter, das Gebiete ohne Passpunkte optimal überbrücken kann. Das Ergebnis der Ausgleichung sind gesuchte UTM-Koordinaten der NTv2-Gitterpunkte.
Mit den vollständig vorliegenden GK- und UTM-Koordinatenpaaren können die geographischen Shiftwerte der NTv2-Gitterpunkte erzeugt werden, die anschließend binär in der NTv2-Datei gespeichert werden. Die Shiftwerte sind Differenzen in Bogensekunden in den geographischen Koordinaten zwischen Ausgangs-Koordinatensystem DHDN (Bessel-Ellipsoid) und Ziel-Koordinatensystem ETRS89.

3.2 Berechnung einer UTM-Koordinate aus der GK-Ausgangskoordinate mittels NTv2-Ansatz

Als Beispiel soll eine Koordinate eines Neupunkts N von GK nach UTM transformiert werden (siehe Abbildung 8). Dazu ist die Gittermasche des NTv2-Gitters zu ermitteln, in die die Neupunktkoordinate fällt. Dazu wird die GK-Koordinate zunächst nach geographischen Koordinaten (Bezugssystem DHDN/Bessel-Ellipsoid) umgerechnet. Über die Gebietsdefinition und die vordefinierte Gitterweite lässt sich die Gittermasche und die Position des Punktes dx, dy innerhalb des Gitters bestimmen. Aus den Shiftwerten der vier Gitterpunkte wird über bilineare Interpolation ein repräsentativer Shiftwert für den Neupunkt N interpoliert. Die im Shiftwert enthaltenen Punkt-Verschiebeinformationen (geographische Koordinatendifferenzen) werden zu den geographischen Koordinaten des Punktes addiert. Diese sind dann interpolierte geographische Koordinaten für den Neupunkt N auf dem ETRS89/GRS80-Ellipsoid. Abschließend können diese dann

für eine vorgegebene Zone durch eine konforme Abbildung in UTM-Koordinaten verebnet werden.

4 Nutzung der NTv2-Datei „BY-KanU“ in Softwareprodukten zur UTM-Transformation

Anhand von ersten ausgewählten Testumsetzungen von Anwendern in ausgewählten Transformationssoftwareprodukten wie beispielsweise FME oder GIS-Produkten (z.B. ArcGIS) konnte aufgezeigt werden, dass eine reibungslose und einfache Integration des NTv2-Transformationsansatzes möglich ist und die anwenderseitig gewünschten Genauigkeitsanforderungen auch softwareseitig erreicht wurden. Durch den Einsatz von Transformationssoftwareprodukten besteht für Fachanwender von Geodaten außerdem die Möglichkeit, sowohl große Mengen von Daten, als auch Daten unterschiedlichster Formate (z.B. CAD-, GIS-, Raster-Formate oder nicht räumliche Daten wie CSV, XLS etc.) in das neue Raumbezugssystem zu überführen. Dies wird anhand von ausgewählten Testszenarien demonstriert.

5 Literaturverzeichnis

BENNING, W. & SCHOLZ, T., 1990: Modell und Realisierung der Kartenhomogenisierung mit Hilfe strenger Ausgleichungstechniken. Zeitschrift für Vermessungswesen (ZfV), DVW Gesellschaft für Geodäsie, Geoinformation und Landmanagement e.V., 1990/2. 45-55.

GLOCK, C. & VETTER, C., 2014: Strategie zur Einführung des neuen Bezugssystems in Bayern. Vortrag bei „UTM in Bayern, was geht uns das an?“, Geodätisches Kolloquium der Hochschule Würzburg-Schweinfurt (FHWS), Würzburg.

GLOCK, C., 2016: Projekt KanU – Strategie zur UTM-Umstellung des Liegenschaftskatasters in Bayern. Vortrag bei GIS@Aktuell - Neue Koordinaten für alle – von Gauß-Krüger nach UTM, 27.04.2016, München.

HAMPP, D. & GLOCK, C., 2017: Ein ausgeglichenes Bayern, DVW-Bayern, Heft 2/2017, 105ff, München, 2017. http://www.dvw.de/sites/default/files/landesverband/bayern/anhang/beitragskontext/ 2017/hampp.pdf.

HAMPP, D., 1992: Digitalisierung, Homogenisierung und numerische Fortführung – ein schneller Weg zur aktuellen Digitalen Flurkarte. Mitteilungsblatt DVW Bayern, Heft 3/1992: 239-252. München.

JOOS, G. & GLOCK C., 2017: Bezugssystemwechsel auf ETRS89/UTM-Grundlagen, Kap. 2, S. 16-45, im Leitfaden Runder Tisch GIS e. V., Hrsg.: Donaubauer A., Kolbe T., München. http://rundertischgis.de/images/2_publikationen/leitfaeden/Bezugssystemwechsel/Leitfaden_Bezugssystemwechsel.pdf.

WITTMANN, J., 2015: Die gitterbasierte Transformation NTv2 zur Umrechnung zwischen DHDN90/GK und ETRS89/UTM in Bayern. Bachelorarbeit, Hochschule München, 2015. https://www.ldbv.bayern.de/file/pdf/10515/Bachelorarbeit_Wittmann_Johanna.pdf.

Modeling Subsurface Soil Moisture Based on Hyperspectral Data - First Results of a Multilateral Field Campaign

SINA KELLER[1], FELIX M. RIESE[1], NIKLAS ALLROGGEN[2], CONRAD JACKISCH[3] & STEFAN HINZ[1]

Abstract: Soil moisture dynamics and its spatial distribution represent important data of a landscape. Until today, a precise and spatially continuous measurement of these is difficult. Current remote sensing approaches focus on an upscaling of measured hyperspectral and hydrological data from laboratory setups by means of hydrological models. In this paper, we follow an alternative approach on a pedon-scale. We present a multilateral dataset from an irrigation campaign in August 2017 employing a comprehensive hydrological, geophysical, and hyperspectral sensor setup. In a first regression modeling approach, we apply machine learning methods to estimate subsurface soil moisture based on hyperspectral data and hydrological reference data. The derived results reveal the potential of estimating field-measured soil moisture. They provide the basis for further studies of modeling approaches based on this unique dataset.

1 Introduction

Soil moisture is a key variable of a landscape. Knowledge about its dynamics and spatial distributions is crucial for many application areas from agricultural planning to disaster mitigation. Characterizing soil moisture of field sites usually is based on sparse or incomplete experimental data. Common approaches to measure the spatial-temporal dynamics of soil moisture rely either on point-wise measurements e.g. with probes relying on estimations of the relative electrical permittivity or soil water tension, or on procedures such as satellite-based remote sensing techniques. The point-wise measurements provide a high temporal but sparse spatial resolution. Satellite-based techniques are characterized by a high spatial, heterogeneous, and poor temporal resolution. This gap between spatial and temporal resolution is well-know (RUBIN & HUBBARD 2005; ROBINSON et al. 2008).

To close this gap, researchers increasingly apply terrestrial remote sensing techniques in particular with passive sensors since the 1970s (HUETE & USTIN 2004). One possible solution opens up with recent developments in close-range hyperspectral remote sensing. It provides a high spatial and spectral resolution for analyzing the dynamics of surface signatures. To retrieve spectral signatures from the recorded hyperspectral data presents a major task in this context (cf. SALISBURY & DARIA 1992).

[1] Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Englerstraße 7, D-76131 Karlsruhe, e-mail: [Sina.Keller, Felix.Riese, Stefan.Hinz]@kit.edu
[2] University of Potsdam, Institute of Earth and Environmental Science, Karl-Liebknecht-Straße 24-25, D-14476 Potsdam-Golm, e-mail: Niklas.Allroggen@geo.uni-potsdam.de
[3] Karlsruhe Institute of Technology (KIT), Institute of Water and River Basin Management, Otto-Amman-Platz 1, D-76131 Karlsruhe, e-mail: Conrad.Jackisch@kit.edu

Hyperspectral sensors measure only the reflectance of the top soil layer. However, the soil surface consists of mixed features. They overlay in the reflectance spectra which results in interference. Mixed reflections often impede inference on the state and the process of the investigated object (cf. PLAZA et al. 2003; CLARK et al. 2003). Hence, such effects constitute the modeling of subsurface states and processes based on hyperspectral data as a remaining and challenging task. Many studies which deal with the estimation of soil moisture from spectral signatures reduce the overlaying mixture aspects by examining only bare soil. These studies are mentioned in the subsequent section. They conduct these experiments primarily in a laboratory setting or combine these settings with a field experiment.
From a model-based perspective, two approaches are commonly applied to estimate soil-moisture contents based on hyperspectral data: First, the widely used spectral indices which essentially engineer features by combining spectral bands (FABRE et al. 2015; OLTRA-CARRIÓ et al. 2015; VEREECKEN et al. 2014); second, the emerging data-driven machine learning approaches which increasingly are applied in remote sensing (STAMENKOVIC et al. 2013; ALI et al. 2015). The latter are justified by the potentials of machine learning to process large data, in case of satellite-based data, and handle non-linear regression problems such as modeling soil-moisture dynamics. The majority of these machine learning approaches integrate supervised learning, e.g. Random Forest (RF), Support Vector Regression (SVR), or artificial neural networks.
Despite of all progress in the combination of hyperspectral data and machine learning approaches, the uncertainty to derive estimations from the spectral surface states on subsurface conditions remains. This uncertainty of state measurements even exists in the domain of point-based pedo-hydrological monitoring (JACKISCH et al. 2018). Techniques which can quantify water distribution in the soil continuum gradually emerge (KOESTEL AND LARSBO 2014; SCHLÜTER et al. 2016, JACKISCH et al. 2017). From a process-perspective, the soil surface is known as central area related to the partitioning and redistributing of precipitation before the latter infiltrates into the soil matrix and flows in subsurface paths (among others BROOKS et al. 2015, JARVIS 2007, NIMMO 2016).
Referring to this characterization, we seek to address the complex modeling of soil moisture dynamics based on spectral data and by conducting a multi-sensor field experiment:

- by creating defined surface conditions during irrigation and subsequent drying of the study area and
- by explicitly monitoring the development of the thermal conditions in the subsurface and ground at the atmospheric boundary in one dimension with diverse sensor systems.

In addition to the multi-sensor field experiment, we present a framework which involves standard supervised learning approaches for estimating subsurface soil moisture dynamics solely with hyperspectral data in order to analyze the potential of remote sensing as a stand-alone tool.
After presenting the field experiment in Section 2, we explain in detail the measured dataset in Section 3. Subsequently in Section 4, we describe the applied machine learning approaches to estimate soil moisture values and dynamics. Thereupon, we present the derived results and discuss these results in Section 5. The paper concludes in Section 6 and we give a summary of prospective studies.

2 Field Campaign

To correspond as much as possible to real-world conditions, we performed our field experiment on a grassland site and differ thereby from the bare-soil laboratory experiments. The field campaign was conducted between the 15th and 21st August 2017 in Linkenheim-Hochstetten, Germany. The grassland site on loamy sand was chosen for its behavioral hydrological properties and its accessibility.

The experiment was designed in eight plots covering an area of one square meter each (Fig. 1). The plots in row 1 are intensively monitored and the plots in row 2 functioned as replica with fewer observations. The pedo-hydrological states and dynamics of the plots were monitored using several time domain reflectometry (TDR, IMKO Pico32) probes and tensiometers (METER T5). In addition, a Bromide tracer (3 g/L KBr) was used in the irrigation water, which was resampled two hours after the end of irrigation in percussion drilled core samples with 8 cm diameter. These cores were sampled in 2 cm increments to about 40 cm depth for Bromide analysis; for methodological details see JACKISCH 2017.

The irrigation was facilitated by a drip irrigation platform (ALLROGGEN et al. 2017), fed from a common reservoir. Outflow was monitored by a mechanical meter. To maximize the number of functional replica, the irrigation scheme is based on pulses of five liter, except for the last pulse in column D with 15 L. Moreover, the irrigation intensities increased (see height of irrigation bar in Fig. 1) in steps of approximately 0.3 L/min, 0.8 L/min, 1.2 L/min, and 2 L/min. To provide data on the dynamics of subsurface conditions, we employed ground penetrating radar (GPR) measurements to record repeated shot gather and common offset profiles across the experimental plots (Fig. 1).

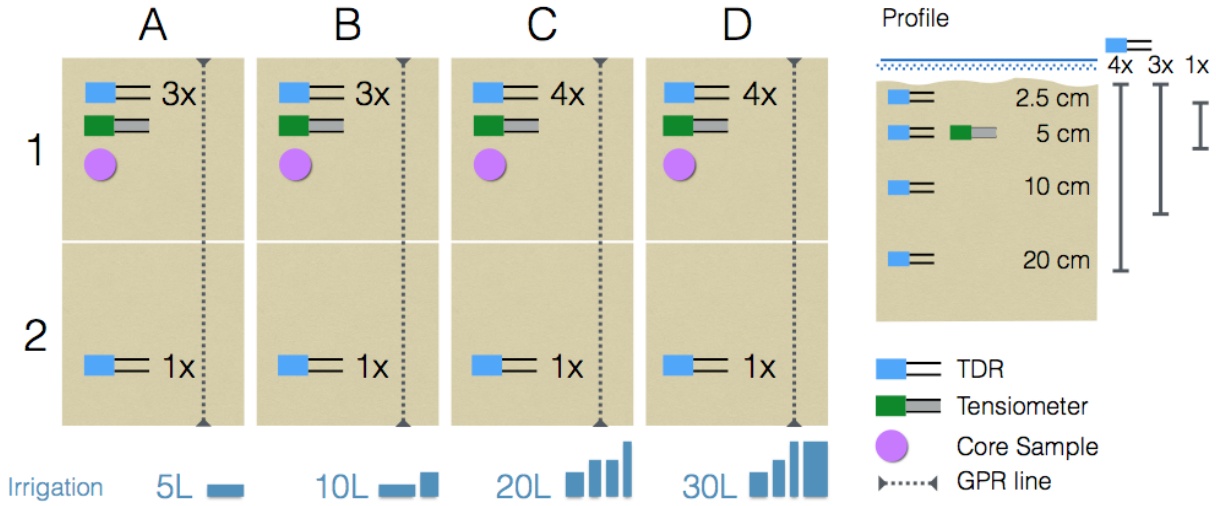


Fig. 1: Sketch of the experimental plots and the instrumentation. Each plot has a size of 1m^2. The irrigation pulses have been 5L each with increasing intensity. The last pulse at D1 and D2 were 15L.

Due to the high dependency of the electromagnetic wave propagation velocity to soil moisture variations, the GPR is frequently used for mapping spatial variations of soil moisture (e.g. HUISMAN et al. 2003). In principle, GPR transmits an electromagnetic wave and records the

signal at one or several receivers. The recorded signals are used to calculate the GPR propagation velocity and estimate the soil moisture content.
As remote sensing techniques, we installed a hyperspectral snapshot camera, a hyperspectral spectroradiometer, and an infrared (IR) camera mounted on a stage at 10 m distance to avoid interference with the GPR. A hyperspectral snapshot camera (Cubert) presents in this paper the primary remote sensing sensor system. It recorded frames of the plot areas. Additionally, we measured hyperspectral, point-based spectra (ASD). Thus, we are able to compare the hyperspectral snapshot data and the hyperspectral point-measured data (cf. Fig. 2). With the long-wave infrared (LWIR) measurements (FLIR), we extend the spectral range. Table 1 outlines the descriptions of the applied remote sensing sensors.

Table 1: Description of the installed remote sensing sensor systems

Sensor System	Spectrum	Channels	Measurement Frequency	Spatial Resolution
Cubert UHD 285 (hyperspectral snapshot)	450 - 950 nm	125	20 / hour	50 × 50
ASD FieldSpec 4 (spectroradiometer)	350 - 2500 nm	2151	1 / hour	1 × 1
FLIR Tau2 640 (LWIR)	7.5 - 13.5 µm	1	10 / hour	640 × 512

Fig. 2 provides a summary of the used sensors system and measured data during the field campaign. The linkage between soil-moisture values as reference data and hyperspectral images as input data is used in this paper to address the estimation of subsurface soil moisture with supervised learning approaches.

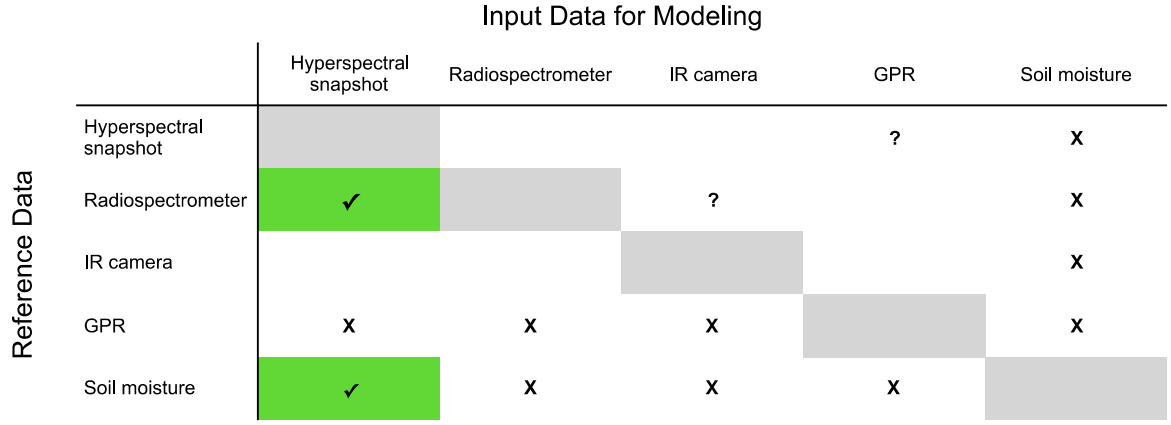


Fig. 2: Summary of the measured sensor data split into reference data (rows) and input data for modeling (columns). Green cells with "✓" characterize the data used in this paper. Cells labeled with "X" characterize the linkages of reference and input data which we will approach in prospective studies. Cells labeled with "?" symbolize linkages which we intent to study, if they can be pursued.

3 Data Processing & Analysis

3.1 Hydrological Processes

The different irrigation schemes and respective soil water redistribution result in stepwise increases of soil moisture as shown in Fig. 3. A conversion of the changes in soil moisture to the recovery of absolute water input outlines the boundaries of certainty of the experiment and sensor-based soil water monitoring. Fig. 3 cumulates the total water recovery at the sensors of the four plots (A1, B1, C1, and D1). With a perfect recovery, each step would be exactly 5 mm and the recession curves would be very gentle as in A1 and C1 of Fig. 3. The deviation can be explained an occurrence of free water, a lateral diffusive redistribution from the plot, and a non-uniform flow field. Peaking maxima and steep recession curves suggest partially quick lateral redistribution of free water.

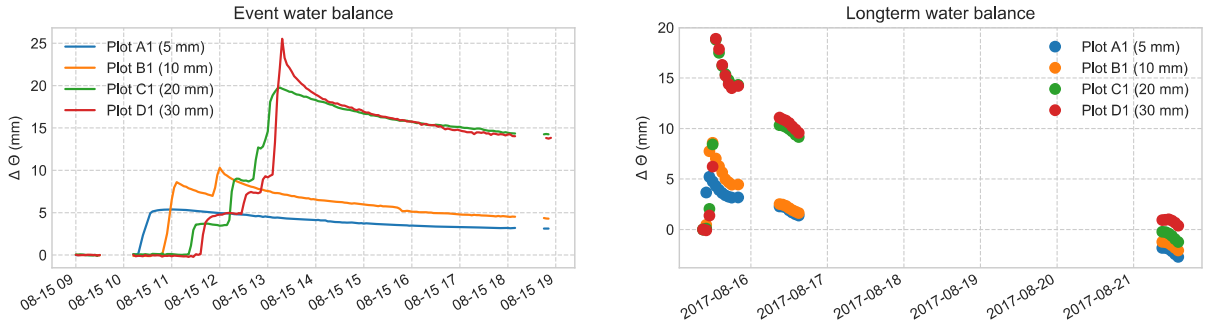


Fig. 3: Left: Water balance in the plots A1 to D1 on the day of the irrigation. Right: Water balance over the course of the experiment.

As complementary reference, we derived soil water retention properties for the test site based on undisturbed ring samples (Fig. 4). When plotting the observed pairs of soil moisture and matric potential, deviations from the laboratory-derived dying curve can be attributed to the existence of free water. As such, these periods may be least suitable to assume locally homogeneous and well-mixed states.

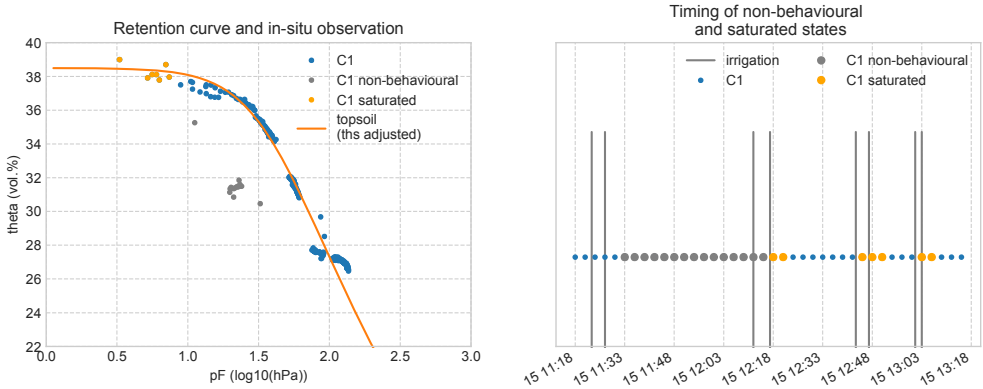


Fig. 4: Exemplary laboratory and in-situ soil water retention to identify non-behavioral states of plot C1. These mostly occur immediately after the irrigations.

As a reference for the overall event water distribution into the subsurface, the concentration profiles of Bromide tracer are considered. For the light irrigation with 5 L in 17 min (A1), infiltrating water reaches almost 0.2 m depth with a peak around 0.15 m (Fig. 5, left). Interestingly, this local peak persists over the different plots. Moreover, relative and absolute surface concentrations do not deviate much from the different irrigations. According to the cumulative recovery plots in Fig. 5 (right), it can be concluded that high irrigation intensities do not necessarily result in stronger surface signals. Furthermore, approximately 45% of the irrigated tracer mass represents the maximum of the relative recovery in the top 4 cm. At this level, we expect to deduce the surface signals.

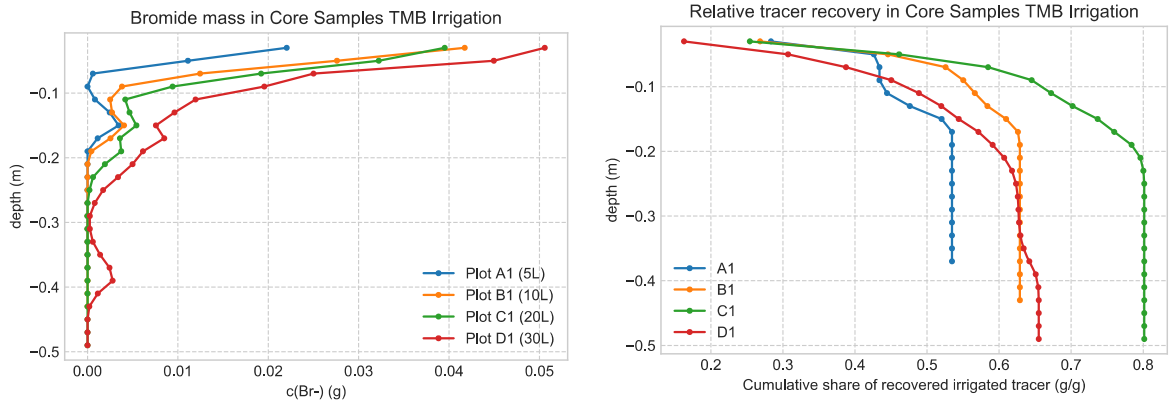


Fig. 5: Bromide tracer recovery.

During the drying process of the experiment, the surface states of the different plots return to antecedent state, with continued drying after five days. The established state variance between the plots converges only slowly (Fig. 3, right).

Based on the hydrological findings, we assume that the irrigated water established a non-uniform flow field into the subsurface and reached a depth range between 0.2 m to 0.4 m. The post-irrigation recession is not uniquely defined by the irrigation amount alone. Different mobilizations in the subsurface may result in similar surface soil moisture dynamics. This effect can be seen in Fig. 3 (left) at 20 mm and 30 mm irrigation. With respect to functional coherence, we have shown that datapoints after initial irrigation and briefly after the successive irrigation pulses should be excluded from references for the spectral signals (see Section 4).

3.2 Ground Penetrating Radar Data Processing

GPR data are collected during the first two days of the experiment to provide information on the subsurface heterogeneity and the spatial soil moisture patterns. We used a pulseEKKO PRO GPR system equipped with a pair of 1 GHz shielded antennas. To ensure accurate and repeatable positioning, the GPR antennas were attached to wooden guides and their position was recorded by a self-tracking total station while we recorded a series of common offset profiles and shot gathers (not shown here). In total, we recorded 87 GPR profiles at irregular time intervals.

The GPR data processing follows a standard processing scheme including a bandpass filter, time-based amplitude scaling, first-arrival based zero-time correction, gridding to a regular 0.01 m spacing and a Kirchhoff migration approach using a constant velocity of 0.1 m/ns.
In Fig. 6, we present a processed common offset GPR profile originating from plot B. The data show direct signals until a traveltime of 8 ns. Later signals are reflected at subsurface structures. For interpreting these data in terms of soil moisture changes, we identify traveltime differences to a reflector over the course of the experiment by applying a correlation based automatic picking algorithm starting with a seed traveltime of 15 ns (corresponding to a depth of 0.6 m). The results of this approach (exemplary shown as a red line in Fig. 6) show traveltime variations in the order of 0.5 ns. For a hydrological interpretation, we apply a petrophysical translation of traveltime differences to soil moisture variations based on the complex refractive index model (CRIM) as presented by ALLROGGEN et al. (2015) using a realistic estimate for the initial soil velocity of 0.1 m/ns.

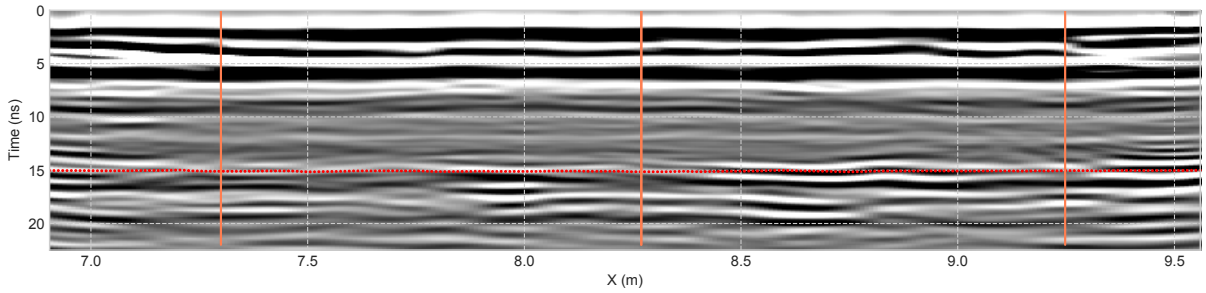


Fig. 6: Exemplary common offset GPR profile from plot B recorded at 10:57 in the morning of the first day of the experiment. Red dots mark the traveltimes selected by our correlation analysis. Orange lines mark the boundaries of the irrigation plots

In Fig. 7, we present exemplary results for the retrieved spatial soil moisture variation for plots A and B. We notice a good agreement of the total detected changes for the irrigated water (5 mm for plot A and 10 mm for plot B) in the time shortly after the end of the irrigation and a gradual drying afterwards. It should be noted that several structures in the signal persist over the course of the experiment. With regard to remote sensing application this method provides spatial details on the interpretation of soil moisture dynamics. However, one has to notice the integration depth of 0.6 m. Thus, surface moisture cannot be resolved. For making these signals possible, we recorded shot gathers during the field experiment, which by analyzing the direct arrivals can provide information on the shallower soil moisture variations. With regard to point-based soil moisture measurements, the deviations in event water recovery (Fig. 4) can also be attributed to limited representatively of single TDR probes in the diverse state field.

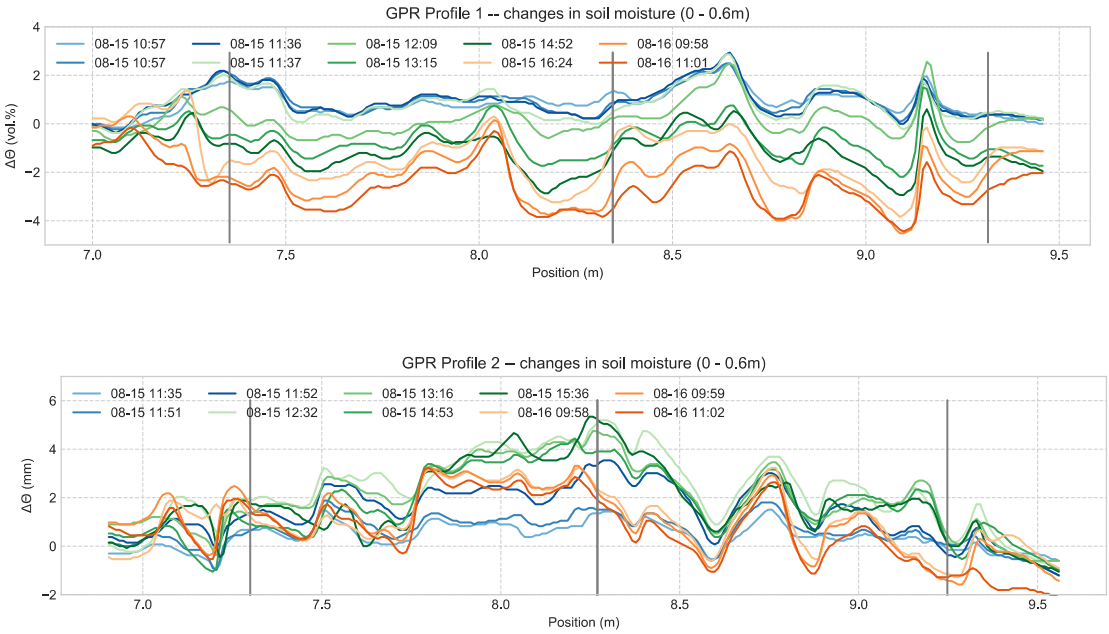


Fig. 7: Changes in soil moisture along common offset profile inferred from changes in apparent two-way traveltime to reflector.

3.3 Remote Sensing Data Processing

Since the hyperspectral snapshot sensor provides the highest temporal resolution, the major analysis in the remote sensing part of this paper is performed on the snapshot data. To apply machine learning (ML) approaches, the feature selection presents a prerequisite for the latter.

In this paper, we conduct two different feature selection procedures. First, the average spectra of each plot and measurement are calculated with respect to the approximated position of the TDR probes (Fig. 8, left). The wooden guides (cf. Fig. 8, left) are masked and ignored in the calculation of the average spectrum. The spectralon (cf. Fig. 8, left) functions as white reference of the spectrum in every image. Second, we dismiss five bands at the beginning and five bands at the end of the spectrum to remove occurring sensor artefacts. On the basis of the comparison in Fig. 8 (right) between a spectroradiometer and a snapshot, these artefacts are identifiable. The spectroradiometer data functions as a cross check for the snapshot data. The shapes of both spectra are similar, but the amplitudes of both distributions differ clearly. The observed offset indicates differences in the calibration and measurement angle.

Two aspects have to be considered: First, the estimation in the following sections relies on a specific sensor, in this case the hyperspectral snapshot sensor. Second, with respect to the similar spectra shapes, we assume that both hyperspectral sensors measured the same reflectance of the surface soil moisture states.

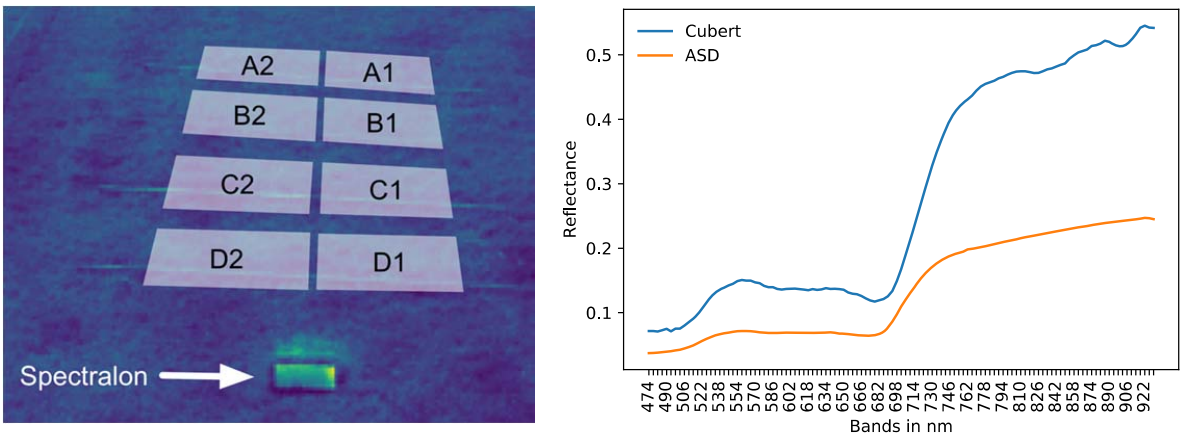


Fig. 8: Left: An example of a hyperspectral snapshot. The 50×50 image is pan-sharped to 1000×1000 due to visualization purposes. In context of the modeling approach, the raw images with the 50×50 resolution are used. The dark areas are masked and, based on the white grid areas, the average spectra are calculated. Right: Exemplary spectra of the spectroradiometer and the snapshot sensors recorded at the same plot and at the same time.

In addition to the hyperspectral data acquisition, the infrared (IR) data are averaged to calculate an average temperature value. No further pre-processing is applied on the IR data. After a pre-analysis of the data, we decided to exclusively focus on the estimation quality of the hyperspectral data importance in the regression. Within the entire multilateral dataset, the IR as well as the hyperspectral spectroradiometer measurement are included and can be analyzed in future studies.

In the following, we refer to the measured soil moisture value (5 cm) as reference data combined with the simultaneous recorded, pre-processed as well as averaged hyperspectral data as a single datapoint. In total, the dataset consists of 1940 datapoints.

4 Regression Methods

The supervised learning approach is performed with the hyperspectral image data as input vector and the soil moisture data as desired output value. For the soil-moisture estimation, we apply common regression methods i.e. Random Forest (RF) (cf. BREIMAN 2001), extremely randomized tress (ET) (cf. GEURTS et al. 2006) as an extension of the RF algorithm, as well as Support Vector Regression (SVR) (cf. VAPNIK 1995). These regression methods represent model-independent approaches.

The RF and the ET regressors are implemented in the Python package *scikit-learn* (PEDREGOSA et al. 2011). We perform the regression with the default settings of hyperparameters, defining the number of estimators in both methods. Hyperparameters are parameters set before the estimator's training process. The regression result converges with 1000 estimators. The RF and ET regressors ranks the importance of the input variables after the training (cf. Section 5). The

support vector regressor (SVR) is based on the package *libsvm* (CHANG 2011) and is also implemented in Python's *scikit-learn*. The SVR hyperparameters are tuned with a grid-search approach.

The dataset is split randomly into a train subset of 382 datapoints and a test subset of 1558 datapoints. For a reasonable modeling of continuous soil moisture variables, the distributions of both subsets have to be representative of the soil moisture distribution in terms of distribution range and, if possible, in terms of distribution shape. As outlined in Section 3.1, the soil moisture distribution differs between the different plots. Thus, we have to consider these effects in context of the soil moisture modeling based on hyperspectral data. Pre-emptively, we exclude combinations between two plots (e.g. A1 and B1) for the training and validation.

Providing the regressor with data from all plots enables the machine learning approaches to identify spectral bands which are correlated to the soil-moisture values. Fig. 9 provides histograms of the soil moisture distribution in the test and training subsets. All datapoints have been cross-checked with the hydrological findings introduced in Section 3.1. In coherence with the hydrological interpretation of the observations, we exclude datapoints which are measured after initial irrigation and briefly after the successive irrigation pulses.

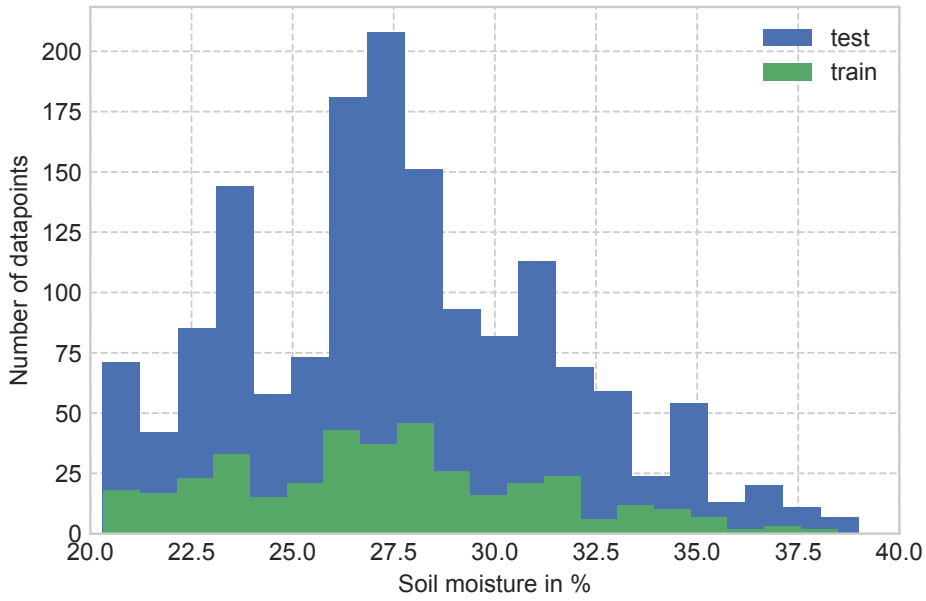


Fig. 9: Soil-moisture distribution of the test (blue) and train (green) subset.

The regression parameters are learned from conjugate single input and reference datapoints. Thus, time is not used in these approaches, neither for training nor for estimation. By this, it is possible to estimate soil moisture also from one single hyperspectral image, like in the satellite-case. The modeling process is outlined in Fig. 10.

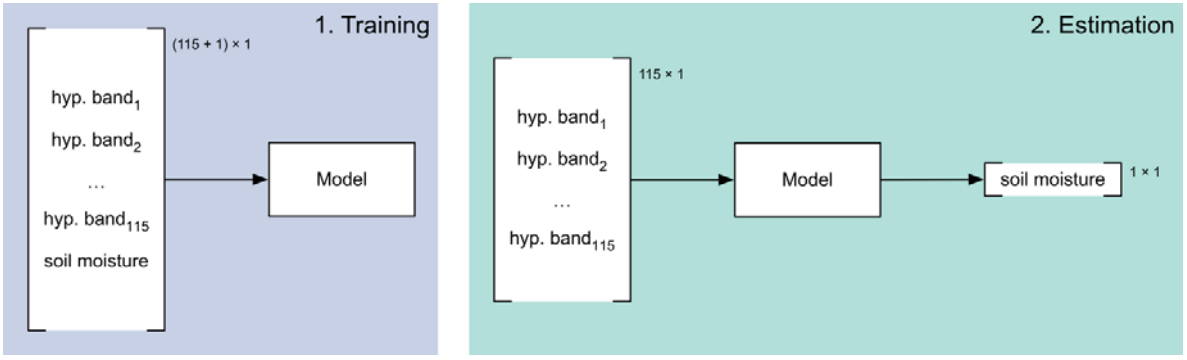


Fig. 10: Regression schema for estimating soil moisture

5 Results & Discussion

The regression results are summarized in Fig. 11 and Table 2. The SVR performs badly with a coefficient of determination R^2 of 56%. RF and ET outperform the SVR to solve the presented regression problem. RF reaches a R^2 of 73% and a root mean squared error (RMSE) of 2.00 (%-soil moisture). The best regression results are achieved by the ET with a R^2 score of 78% and a RMSE of 1.81 (%-soil moisture).

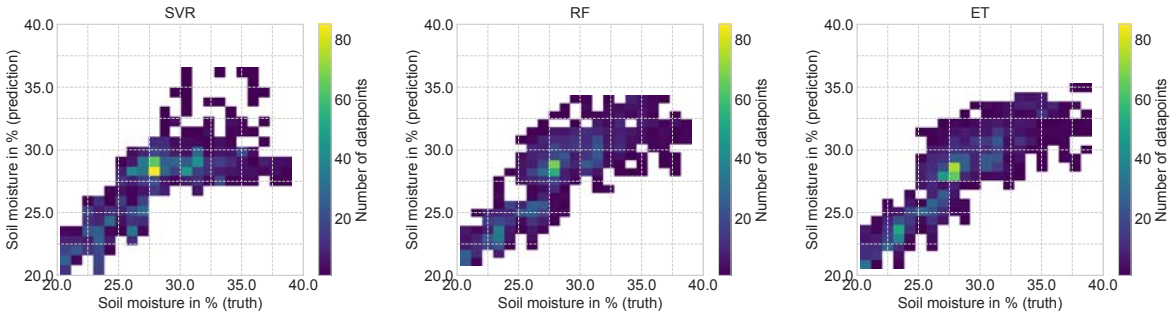


Fig. 11: Regression results for the SVR, RF, and ET.

Table 2: Regression results with the coefficient of determination R^2 and the RMSE.

	R^2 in %	RMSE in %-soil moisture
SVR	56.38	2.58
RF	73.69	2.00
ET	78.41	1.81

Fig. 12 shows the feature importance of the hyperspectral bands based on the ET regressor as well as the mean spectrum of the dataset with its respective standard deviation. In contrast to widely used spectral indices, the supervised learning approaches RF and ET weight the input vector (hyperspectral data) by means of their importance for the estimation quality. Considering the feature importance, we recognize that the spectral bands between 720 nm and 850 nm are significantly more important than the remaining ones. While the variance of the mean spectrum is relatively large in the upper third of the band spectrum, the feature importance distribution shows a peak between 720 nm and 850 nm. The bands between 850 nm and 930 nm possess a minor feature importance and a large variance. This finding indicates that these band exhibit noise, e.g. occurring due to weather and sensor conditions. Furthermore, the RF and ET separate the reflectance of vegetation and soil moisture without further information on the underlying processes is inserted.

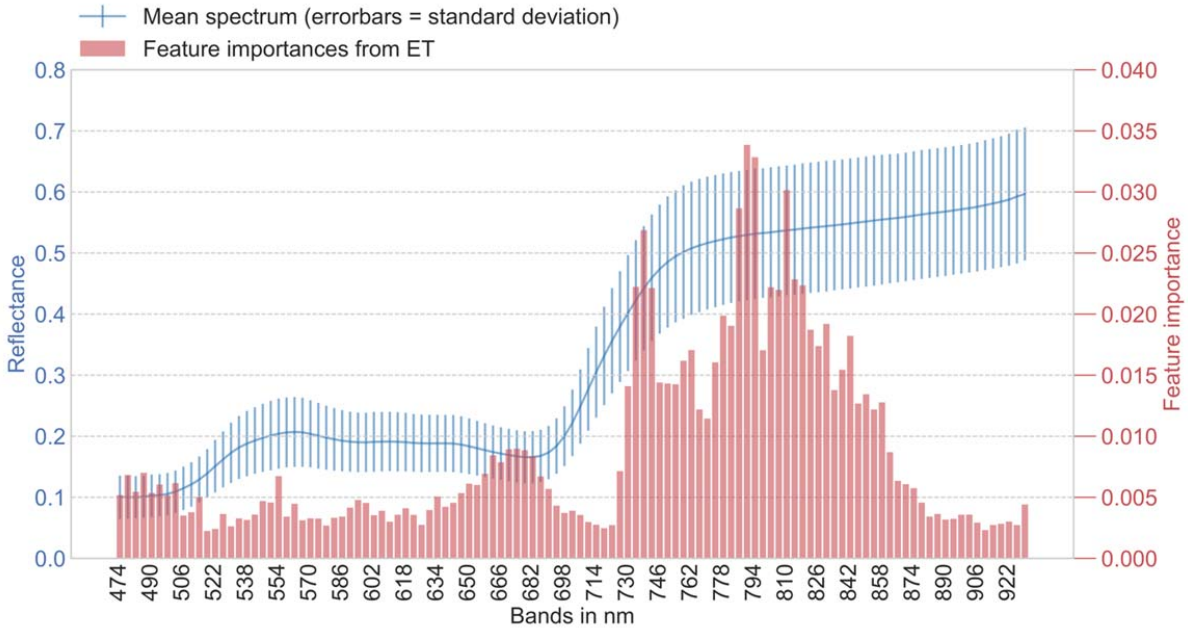


Fig. 12: Mean spectrum of the dataset with its standard deviation as error bars (blue line), feature importance of the ET regressor (red bar plot).

6 Conclusion & Outlook

In this paper, we investigated the potential of hyperspectral data for surface soil moisture modeling in grass-land like areas. We acquired a multi-sensor dataset measured in August 2017 in Linkenheim-Hochstetten, Germany. The resulting dataset consists of data from a suite of different hydrological, geophysical, and remote sensing systems. We presented the respective sensor data as an overview. In extension to previous remote sensing studies focusing on bare soil

experiments in the context of modeling surface soil moisture, we transferred the data acquisition to a grass side.
As input data hyperspectral and hydrologically reasonable data are used. The applied standard machine learning approaches only rely on the described input data: only artefacts were excluded, but no pre-processing was conducted. The derived first results reveal that an estimation of surface soil moisture can be achieved based on hyperspectral data in combination with machine learning approaches. An additional implementation including e.g. tuning of regressors will be conducted in future studies. Furthermore, we demonstrated that an experiment at a complex environment, e.g. grass soil, is feasible. We showed that the proposed supervised methods are able to handle this complexity, if the proposed setting of hyperspectral input and desired soil moisture data is given.
Coming in short to the hydrological perspective, further research will be conducted in modeling the excluded datapoints (measurements immediately at the irrigation point). One promising aspect to tackle this challenge might be to consider the infrared data additionally in the context of machine learning. The GPR data show a level of spatial heterogeneity, which was not modeled in this study. Nevertheless, when studying spatial variations this method provides the possibility for validating and training spatial classification approaches.
For prospective studies, we intend to publish the aggregated dataset.

7 Acknowledgements

We thank our colleagues Torsten Schuler, Jürgen Hauth, and Armin Hofmann from the Institute of Technology and Management in Construction at the Karlsruhe Institute of Technology for their technical support. We also thank Wolfgang Groß from the Fraunhofer Institute of Optronics, System Technologies and Exploitation for lending us a spectroradiometer and the research group “Catchments as Organized Systems” for collaborative discussions.

8 References

ALI, I., GREIFENEDER, F., STAMENKOVIC, J., NEUMANN, M. & NOTARNICOLA, C., 2015: Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. Remote Sensing, **7**, 16398-16421.

ALLROGGEN, N., VAN SCHAIK, N.L.M.B. & TRONICKE, J., 2015: 4D ground-penetrating radar during a plot scale dye tracer experiment. Journal of Applied Geophysics, **118**, 139-144.

ALLROGGEN, N., JACKISCH, C. & TRONICKE, J., 2017: Four-dimensional gridding of time-lapse GPR data, presented at the 2017 9th International Workshop on Advanced Ground Penetrating Radar (IWAGPR), 1-4.

BREIMAN, L., 2001: Random forests. Machine Learning, **45**(1), 5-32.

BROOKS, P. D., CHOROVER, J., FAN, Y., GODSEY, S. E., MAXWELL, R. M., MCNAMARA, J. P. & TAGUE, C. (2015). Hydrological partitioning in the critical zone: Recent advances and opportunities for developing transferable understanding of water cycle dynamics. Water Resources Research, **51**(9), 6973-6987.

CHANG, C.C., 2011: LIBSVM: A library for support vector machine. ACM Transactions on Intelligent Systems and Technology, **2**(3), 1-27.

CLARK, R.N., SWAYZE, G.A., LIVO, K.E., KOKALY, R.F., SUTLEY, S.J., DALTON, J.B. & GENT, C.A., 2003: Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. Journal of Geophysical Research: Planets, **108**(E12).

FABRE, S., BRIOTTET, X. & LESAIGNOUX, A., 2015: Estimation of Soil Moisture Content from the Spectral Reflectance of Bare Soils in the 0.4–2.5 µm Domain. Sensors, **15**, 3262-3281.

GEURT, P., ERNST, D. & WEHENKEL, L., 2006: Extremely randomized trees. Machine Learning, **63**(1), 3-42.

HUETE, X. & USTIN, Y., 2004: Manual of Remote Sensing, Remote Sensing for Natural Resources Management and Environmental Monitoring, Manual of Remote Sensing, **4**.

HUISMAN, J. A., HUBBARD, S. S., REDMAN, J. D. & ANNAN, A. P., 2003: Measuring Soil Water Content with Ground Penetrating Radar: A Review. Vadose Zone Journal, **2**(4), 476-491.

JACKISCH, C., ANGERMANN, L., ALLROGGEN, N., SPRENGER, M., BLUME, T., TRONICKE, J. & ZEHE, E., 2017: Form and function in hillslope hydrology: in situ imaging and characterization of flow-relevant structures. Hydrol. Earth Syst. Sci., **21**(7), 3749-3775.

JACKISCH, C., GRAEFF, T., GERMER, K. & DURNER, W., 2018: Soil moisture and matric potential – what do we measure? Vadose Zone Journal, in submission.

JARVIS, N. J., 2007: A review of non-equilibrium water flow and solute transport in soil macropores: principles, controlling factors and consequences for water quality. *European* Journal of Soil Science, **58**(3), 523-546. http://doi.org/10.1111/j.1365-2389.2007.00915.x

KOESTEL, J. & LARSBO, M, 2014: Imaging and quantification of preferential solute transport in soil macropores. Water Resources Research, 50(5), 4357-4378.

NIMMO, J. R., 2016: Quantitative Framework for Preferential Flow Initiation and Partitioning. Vadose Zone Journal, **15**(2).

OLTRA-CARRIÓ, R., BAUP, F., FABRE, S., FIEUZAL, R. & BRIOTTET, X., 2015: Improvement of Soil Moisture Retrieval from Hyperspectral VNIR-SWIR Data Using Clay Content Information: From Laboratory to Field Experiments. Remote Sensing **7**, 3184-3205.

PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M. & DUCHESNAY, E., 2011: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, **12**, 2825-2830.

PLAZA, A., MARTÍNEZ, P., PLAZA, J., & PÉREZ, R., 2003: Spatial/spectral analysis of hyperspectral image data. IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, 298-307.

ROBINSON, D. A., CAMPBELL, C. S., HOPMANS, J. W., HORNBUCKLE, B. K., JONES, S. B., KNIGHT, R. & WENDROTH, O., 2008: Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review. Vadose Zone Journal, **7**(1), 358-389.

SALISBURY, J.W. & D'ARIA, D. M., 1992: Emissivity of terrestrial materials in the 8–14 µm atmospheric window. Remote sensing of Environment, **42**(2), 83-106.

SCHLÜTER, S., BERG, S., RÜCKER, M., ARMSTRONG, R.T., VOGEL, H.-J., HILFER, R. & WILDENSCHILD, D., 2016: Pore-scale displacement mechanisms as a source of hysteresis

for two-phase flow in porous media. Water Resources Research, **52**(3), 2194-2205. http://doi.org/10.1002/2015WR018254

VAPNIK, V., 1995: The nature of statistical learning theory. Springer science & business media.

VEREECKEN, H., HUISMAN, J.A., PACHEPSKY, Y., MONTZKA, C., VAN DER KRUK, J., BOGENA, H. & VANDERBORGHT, J., 2014: On the spatio-temporal dynamics of soil moisture at the field scale. Journal of Hydrology, **516**, 76-96.

Estimation of Chlorophyll a, Diatoms and Green Algae Based on Hyperspectral Data with Machine Learning Approaches

PHILIPP M. MAIER[1], STEFAN HINZ[1] & SINA KELLER[1]

Abstract: Monitoring of inland waters is a major topic in terms of water quality and environmental issues. We suggest that hyperspectral remote sensing can provide a valuable data source to monitor important water parameters. In this study, we sampled hyperspectral data on The River Elbe from a research ship. The data, which we collected, contains 1383 datapoints and demonstrates similar spectral behavior to the data in the literature. A Random Forest regression shows that concentration of chlorophyll a, green algae, and diatoms is predictable based on hyperspectral data with a R^2 of around 80%. In general, the results reveal the potential of estimating chlorophyll a concentration and concentration of different algae taxa with machine learning approaches. For prospective studies, we intend to develop a generic model approach which is able to process various types of input data, measured at different inland waters.

1 Introduction

During the last decades, monitoring of inland waters has become a major research topic in terms of water quality and environmental issues. The monitoring of area-wide water bodies, is highly data intensive. Most of the currently available datasets to evaluate the quality of inland waters consist of sampled point data. To derive information on the entire water body, this data can be unreliable. Recent attempts towards such an area-wide coverage of water quality monitoring have included the application of hyperspectral sensors to gather image-based, remote sensing data. Chlorophyll a (Chl a) and turbidity function as indicators of algae existence which in turn characterize water quality and nutrition supply.

The first approaches to monitor Chl a concentrations with hyperspectral measurements were undertaken by NEVILLE & GOWER (1977). They discovered the correlation between the absorption peak as a minimum of the reflectance at a wavelength of 685 nm and the Chl a concentration in freshwater. GITELSON & KEYDON (1990) and GITELSON (1992) described hyperspectral charts of different types of inland waters. They sampled hyperspectral spectrometer-based data in different climate regions during various seasons and at several trophic states. According to their research, three significant extrema in the charts exist.

One maximum corresponds to the Chl a peak (minimum of the reflectance) discovered by NEVILLE & GOWER (1977), although it lays in a range between 670 nm to 680 nm. Another one covers the range between 550 nm to 570 nm as a maximum of the reflectance. It is identified as a backscattering feature from suspended particles in water (GITELSON 1992). The last maximum varies in the range between 685 nm to 700 nm. The latter increases in the course of higher Chl a

[1] Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Englerstraße 7, D-76131 Karlsruhe, E-Mail: [Philipp.Maier, Stefan.Hinz, Sina.Keller]@kit.edu

concentration (GITELSON & KEYDON 1990). Thereupon, GITELSON (1992) figured out that the maximum of the reflectance between 675 nm and 730 nm shifts towards longer wavelengths with increasing Chl a concentration.
GITELSON (1992) proposed two different types of approaches to predict Chl a concentration: the first one is based on the peak height in the range between 675 nm to 730 nm. The second one relies on the ratio of specific bands. The ratio approach has been pursued and enhanced afterwards by several researchers (see e.g. SCHALLES et al. 1998). MANNHEIM et al. (2004) presented an area-based approach, where the area under the peak without the baseline correlates with the Chl a concentration. RUNDQUIST (1996), FRASER (1998), CRAIG et al. (2006) attempted to include derivatives up to the fourth order, to enhance the predictability of Chl a.
Besides the Chl a concentration, the differentiation of algae taxa is an important topic of modeling environmental processes and water quality. Especially, the cyanobacteria, commonly known as blue-green algae, is a taxon, where a distinction is necessary. It contains several species that can be harmful for animals and human beings when they disperse in drinking water reservoirs. SIMIS et al. (2007) conducted lots of research about cyanobacteria and their unique pigment phycocyanin. The latter features its typical absorption peak at 620 nm. Spectral signatures of further algae taxa were researched e.g. by GITELSON et al. (1999) and HUNTER et al. (2008) in field experiments as well as laboratory experiments such as artificial tanks. In the laboratory studies, they varied the taxa and the Chl a concentration to observe the change in the spectral reflectance signatures. According to their studies, different algae taxa can be distinguished based on the hyperspectral signature. At this point, the hyperspectral sensor demonstrates enhanced performance over the multispectral sensor (HUNTER et al. 2008).
In general, it can be stated: the higher the Chl a concentration, the more the distinguishability between the algae taxa increases. HUNTER et al. (2008) also created synthetic algae cultures by mixing two distinct taxa within different concentrations and demonstrated that they still can be distinguished, while the spectral signature appears as mixture.
GITELSON (1999) remarked that the occurrence of suspended particular organic matter (SPOM), suspended particular inorganic matter (SPIM), and colored dissolved organic matter (CDOM) challenge the prediction of Chl a concentration with spectral information. The three matters occur in most inland water bodies. They influence the backscattering behavior by interacting with the backscattering of chlorophyll and other algal pigments (GITELSON et al. (1999)). Generally, SPIM and SPOM increase the reflectance of the water body (HUNTER et al. 2008). To handle high SPIM and SPOM concentration in the context of predicting Chl a, ZHOU et al. (2013) proposed a multi-band ratio. Reviews of most of the available approaches for Chl a prediction and prediction of several taxa can be found in MATTHEWS et al. (2010) and PALMER et al. (2015).
In this contribution, we present a hyperspectral dataset with reference data, sampled on The River Elbe. We propose a basic machine learning approach to investigate the Chl a concentration in a first step. In a second step we focus on the distinction of two algae taxa: the Green Algae and the Diatoms. The preliminary regression results are conducted with a standardized machine learning regressor, the Random Forest (RF) regression.

2 Data Acquisition & Data Processing

The data, presented in this contribution, was sampled with a hyperspectral sensor mounted on a research ship (Fig. 1), along The River Elbe from Bad Schandau to Geesthacht in Germany. The data acquisition was embedded in the scope of the Elbschwimmstaffel, which took place in summer 2017 in Germany and was funded by the Federal Ministry of Education and Research (BMBF). We sequentially performed measurements along an approximately 500 km long section on The River Elbe, which result in a large set of datapoints. The applied hyperspectral sensor was a Cubert UHD 285 characterized by an amount of 125 bands in the range from 450 nm to 950 nm. Additionally, the multi-sensor system PhycoSens (invention of BBE Moldaenke) sampled in-situ water parameters. Concentration of Chl a, diatoms, and green algae as well as turbidity represent the target variables.

Fig. 1: Hyperspectral sensor mounted on the bow of the research ship

The hyperspectral sensor was calibrated every 20 minutes with a white reference, i.e. spectralon, to compensate the varying sun altitude. Every minute we captured a hyperspectral snapshot within a 70° angle towards the water surface. The spectralon was placed on the railing so that a part of it is visible in every snapshot to control the reflectance. In addition, we equalized minor radiative fluctuations such as slight cloud occurrences. The reference data was sampled every five minutes by the PhycoSens. To expand the dataset, we interpolated the reference data linearly. Since the measured Chl a concentration has changed in a continuous matter, the interpolation is feasible in practice.

During the data processing, the measured mean spectra was calculated by manually selecting an area, which was undisturbed by bubble formations, shadows, or waves. The effect of bubbles or waves can be seen in Fig. 2, resulting in a high variance and higher reflection values. In a next step, we applied two distinctive types of cuts.

First, based on the spectralon reflectance values of the calibration, we removed outlying images, whose reflectance on the spectralon differed around the factor of 0.3. The outlying behavior was mainly affected by shadow occurrence caused by ship turnings, bridge crossings, or sudden cloud coverages. Second, data points, sampled by the PhycoSens, with a Chl a concentration above 200 µg/l, were dismissed. These values exceed the measurement range. In total, we removed more than 500 of our former 2000 hyperspectral images.

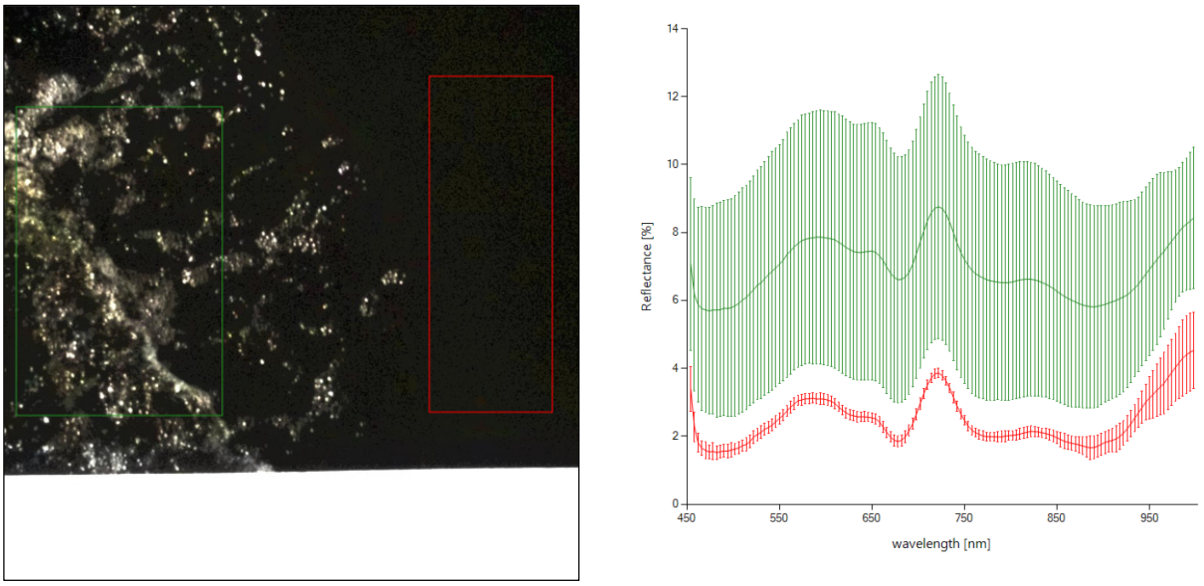


Fig. 2: Left: Hyperspectral snapshot shown as a RGB image. The white area on the bottom of the image represents the spectralon. The red and the green boxes are the selected areas to calculate the mean reflectance. Right: The charts of the selected areas (cf. left part of the figure). The green plot is characterized by “bubble” artefacts and a high variance. The red plot represents undisturbed water surface

Finally, the dataset consists of 1383 datapoints. Each of these refers to 105 hyperspectral values in the range between 486 nm to 902 nm and four measured parameters by the PhycoSens. The bands in the range between 450 nm to 485 nm and 903 nm to 950 nm were cut due to inconsistent behavior and noise.

3 Methods

For the preliminary analysis of the dataset, we perform a RF regression to predict the Chl a concentration based on hyperspectral data. The modeling of the green algae and the diatoms concentration, follows the same approach.

To perform the RF regression, we apply the *Ranger* package (WRIGHT & ZIEGLER 2017) in R. The model is tuned with a grid search employing the train function from the *caret* package (KUHN 2008) to find the best number of variables that are randomly sampled at each split. As splitting rule for the RF algorithm we process *extratrees*.

Several splits of the dataset are implemented: the training dataset is divided in 20, 50 and 80 percent of the dataset. The counterpart of the dataset represents the validation dataset respectively.
In addition to the splitting of the dataset, we evaluate the performance of the framework using a Principle Component Analysis (PCA) to reduce the dimensionality. Then, both approaches are trained with the first 20 components. The regression performance is expressed by the coefficient of determination R^2 and the mean absolute error MAE.

4 Results & Discussion

Fig. 3 shows the dataset in boxplots. Two peaks can be distinguished regarding the mean values:the first is in the range between 582 nm and 606 nm and the second one appears between 714 nm and 718 nm. The first one indicates, that the peak shifts to longer wavelengths with increasing reflectance. In GITELSON (1990) this maximum appears in the range between 550 nm and 570 nm and is affected by reflectance of Chl a as well as by backscattering on suspended particles in water.

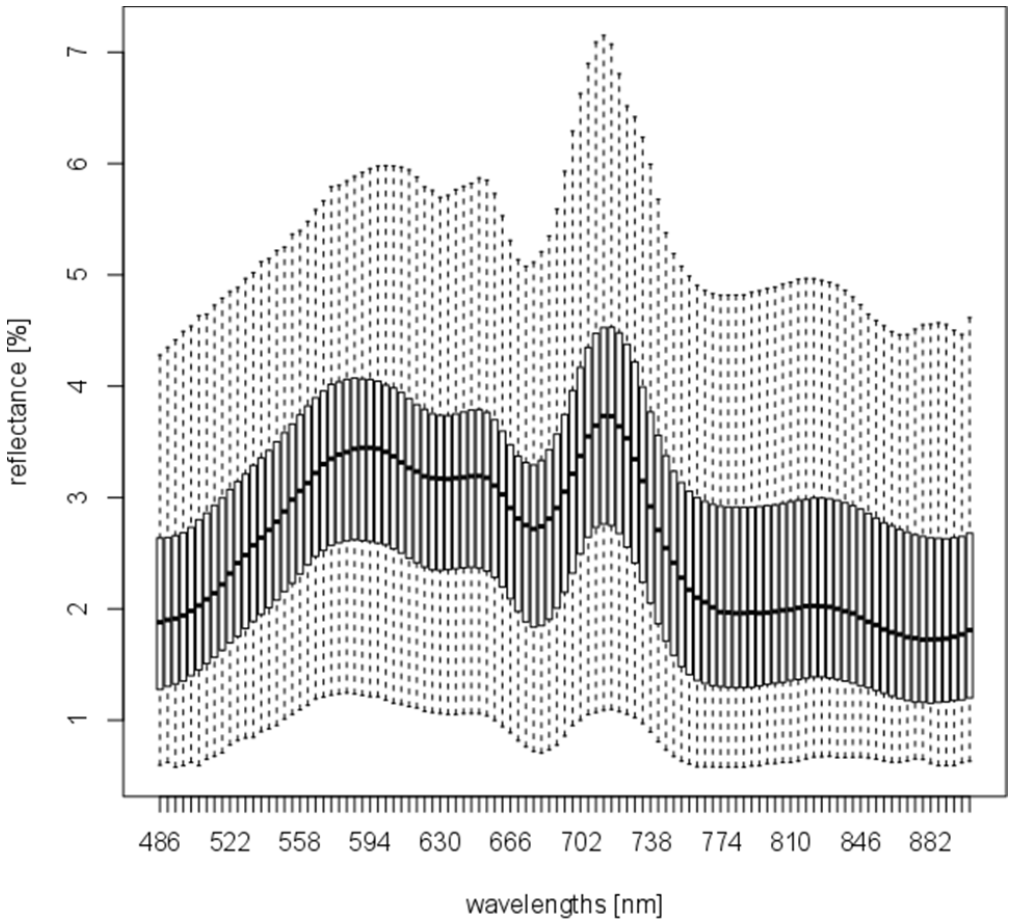


Fig. 3: The dataset is presented as boxplots between the wavelength of 486 nm and 902 nm over all 1383 datapoints. The whiskers represent the extreme values

The River Elbe is well known for its high concentration of SPIM. It can be possible that the maximum shifts even further to longer wavelengths, with a growing SPIM concentration. However, we don't have reference data to proof this behavior. The second peak is related to Chl a reflectance and is similar to values found in the literature (GITELSON 1990). The reflectance minimum at 678 nm represents a significant minimum, due to the absorption of Chl a and also suits to the literature values (GITELSON 1990).
In addition to the main extrema, slight maxima occur at 650 nm and 822 nm respectively. The general reflectance of the wavelengths varies between < 1% and 7%. The highest amplitude in the data is located in the area around the peak between 714 nm and 718 nm. All things considered, the dataset shows no inconsistent behavior.
In Fig. 4 the green line represents the mean spectrum of high Chl a concentration (>150 µg/l) and the blue line shows the mean spectrum of the datapoints with low Chl a concentration (< 50 µg/l). The reflectance of the low Chl a concentration spectrum is generally lower than the high Chl a concentration spectrum. The global peak shifts slightly from 714 nm to 718 nm with higher Chl a concentration.
Additionally, the amplitude from the minimum at 678 nm to the maximum is higher in case of the high Chl a concentration with respect to the lower concentration. In general, a trend can be observed, that an increasing concentration of Chl a, results in higher. But this may also be a consequence of higher concentration of suspended particles in the water.

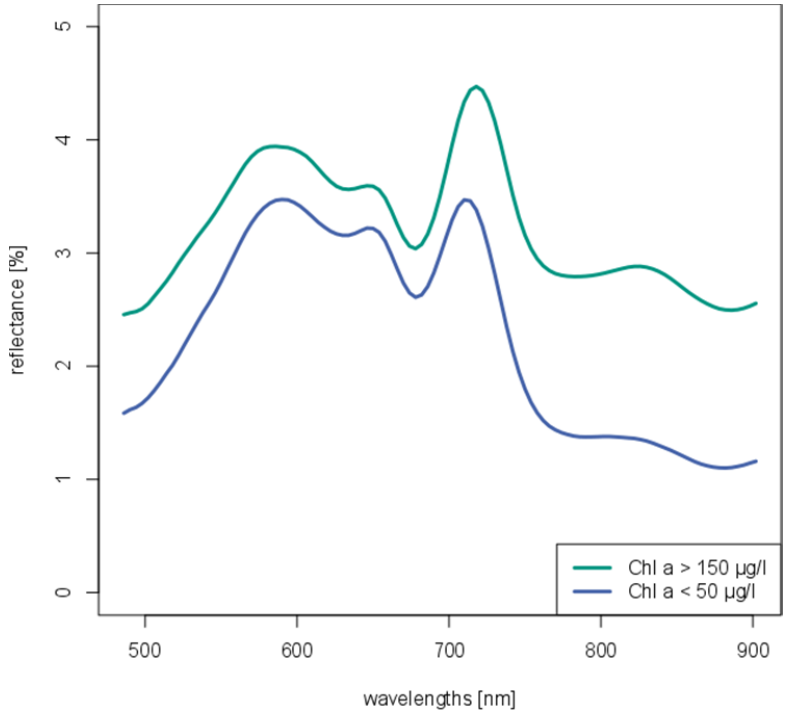


Fig. 4: Mean spectra of two different Chlo a concentration

Tab. 1 outlines the regression results for Chl a concentration with the RF model as estimator and the hyperspectral data as input vector. The regressor is trained with a different amount of data. The quality of the RF model is shown as R^2 and MAE.

In general, the performance increases with a higher amount of training data. But using more than 50% of the training data improves the results only slightly. Applying a PCA before the regression to reduce features enhances the performance. In this modeling, we pick the first 20 components of the PCA to fit the RF model.

Tab. 1: R² and MAE for LR and RF and PCA for the regression

training data	RF		RF + PCA	
[%]	R² [%]	MAE [µg/l]	R² [%]	MAE [µg/l]
20	65	15,2	72	13,0
50	80	11,1	85	9,6
70	83	10,0	87	8,3

Tab. 2 summarizes the regression results for the estimation of green algae and diatoms concentration with the RF algorithm. In general, the estimation of diatoms performs better than the estimation of the green algae, since the range of the concentration of green algae is about double the range of diatom concentration. This is an important aspect when it comes to the training of the algorithm.
PCA has a similar influence on the estimation of the Chl a concentration: performing a PCA before model fitting improves the regression results significantly because of highly correlated variables.

Tab. 2: R² and MAE for Green Algae and Diatom concentration operating the RF algorithm

training data	Green Algae		Green Algae + PCA		Diatoms		Diatoms + PCA	
[%]	R² [%]	MAE [µg/l]	R² [%]	MAE [µg/l]	R² [%]	MAE [µg/l]	R² [%]	MAE [µg/l]
20	57	10,5	63	9,8	72	6,5	79	5,8
50	69	8,6	76	7,6	79	5,3	86	4,5
70	77	7,4	82	6,4	82	4,7	88	4,0

5 Conclusion & Outlook

In general, the results reveal the potential of estimating Chl a concentration and concentration of different algae taxa with machine learning approaches. Although the river Elbe is a challenging inland water with high concentration in suspended particular substances, the RF performs well for all three target variables.
For prospective studies, we intend to develop a generic model approach, which is able to process various types of input data, measured at different kinds of inland waters. The model should later work with hyperspectral data collected by a UAV.

6 Acknowledgment

We thank the organizing committee of the Elbschwimmstaffel for the opportunity to participate in this field campaign. We also thank BBE Moldaenke for sharing with us the reference data. Finally, we express our gratitude to the crew of the research ship "Elbegrund" for their technical. support.

7 References

CRAIG, S.E., LOHRENZ, S.E., LEE, Z., MAHONEY, K.L., KIRKPATRICK, G.J., SCHOFIELD, O.M. & STEWARD, R.G., 2006: Use of hyperspectral remote sensing reflectance for detection and assessment of the harmful alga, Karenia brevis. Applied Optics, **45**(21), 5414-5425.

FRASER, R.N., 1998: Multispectral remote sensing of turbidity among Nebraska Sand Hills lakes. International Journal of Remote Sensing, **19**(15), 3011-3016.

GITELSON, A.A. & KEYDAN, G.P., 1990: Remote Sensing of Inland Surface Water Quality—Measurements in the Visible Spectrum. ratio, **2**, 7.

GITELSON, A., 1992: The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. International Journal of Remote Sensing, **13**(17), 3367-3373.

GITELSON, A.A., SCHALLES, J.F., RUNDQUIST, D.C., SCHIEBE, F.R., & YACOBI, Y.Z., 1999: Comparative reflectance properties of algal cultures with manipulated densities. Journal of Applied Phycology, **11**(4), 345-354.

HUNTER, P.D., TYLER, A.N., PRÉSING, M., KOVÁCS, A.W. & PRESTON, T., 2008: Spectral discrimination of phytoplankton colour groups: The effect of suspended particulate matter and sensor spectral resolution. Remote Sensing of Environment, **112**(4), 1527-1544.

KUHN, M., 2008: Caret package. Journal of Statistical Software, **28** (5).

MANNHEIM, S., SEGL, K., HEIM, B. & KAUFMANN, H., 2004: Monitoring of lake water quality using hyperspectral CHRIS-PROBA data. In Proc. of the 2nd CHRIS/PROBA workshop, ESA/ESRIN, Frascati, Italy.

MATTHEWS, M.W., BERNARD, S. & WINTER, K., 2010: Remote sensing of cyanobacteria-dominant algal blooms and water quality parameters in Zeekoevlei, a small hypertrophic lake, using MERIS. Remote Sensing of Environment, **114**(9), 2070-2087.

NEVILLE, R.A. & GOWER, J.F.R., 1977: Passive remote sensing of phytoplankton via chlorophyll α fluorescence. Journal of Geophysical Research, **82**(24), 3487-3493.

PALMER, S.C., KUTSER, T., & HUNTER, P.D., 2015: Remote sensing of inland waters: Challenges, progress and future directions.

RUNDQUIST, D.C., HAN, L., SCHALLES, J.F. & PEAKE, J.S., 1996: Remote measurement of algal chlorophyll in surface waters: the case for the first derivative of reflectance near 690 nm. Photogrammetric Engineering and Remote Sensing, **62**(2), 195-200.

SCHALLES, J.F., GITELSON, A.A., YACOBI, Y.Z., & KROENKE, A.E., 1998: Estimation of chlorophyll a from time series measurements of high spectral resolution reflectance in an eutrophic lake. Journal of Phycology, **34**(2), 383-390.

SIMIS, S.G., RUIZ-VERDÚ, A., DOMÍNGUEZ-GÓMEZ, J.A., PEÑA-MARTINEZ, R., PETERS, S.W. & GONS, H.J., 2007: Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass. Remote Sensing of Environment, **106**(4), 414-427.

WRIGHT, M.N. & ZIEGLER, A.,2017: ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, **77**(1), 1-17.

Erste Erfahrungen mit der Nutzung von Sentinel-2 Daten zur multispektralen und multitemporalen Analyse mit verschiedenen Klassifizierungsverfahren

MARTIN WEINMANN[1] & UWE WEIDNER[1]

Zusammenfassung: In diesem Beitrag steht die multitemporale Analyse von Sentinel-2 Daten zur Klassifizierung bezüglich diverser Landbedeckungs- und Landnutzungsklassen im Vordergrund. Basierend auf den Reflektanzwerten der einzelnen spektralen Kanäle wird mit verschiedenen Klassifizierungsverfahren eine semantische Interpretation eines betrachteten Gebiets durchgeführt, wobei Sentinel-2 Daten für verschiedene Zeitpunkte separat und in Kombination betrachtet werden. Anhand von definierten Kontrollgebieten erfolgt jeweils eine Bewertung der erzielten Ergebnisse. Dabei wird deutlich, dass durch die multitemporale Analyse von Sentinel-2 Daten eine signifikante Verbesserung der entsprechenden Klassifizierungsergebnisse erreicht werden kann.

1 Einleitung

Die Sentinel-2 Satelliten werden zur systematischen Erfassung von Beobachtungen über Land und Küstengebieten genutzt (SPOTO et al. 2012; KUKAWSKA et al. 2017). Sie bieten im Vergleich zu vorherigen Erdbeobachtungssystemen sowohl hinsichtlich der Auflösung als auch bezüglich der Aufnahmerate eine verbesserte Verfügbarkeit flächendeckender Multispektraldaten. Dadurch kann eine Szenenanalyse anhand von Satellitendaten deutlich begünstigt werden. Während die multispektrale Analyse einer Szene mit verbesserter geometrischer und spektraler Auflösung beispielsweise zur genaueren Kartierung und besseren Unterscheidung von diversen Landbedeckungs- und Landnutzungsklassen genutzt werden kann, erlaubt die multitemporale Analyse eine Beschreibung von dynamischen Prozessen. Jedoch kann die multitemporale Analyse auch zur Verbesserung der mit der multispektralen Analyse erzielten Ergebnisse beitragen. Dies ist insbesondere dadurch bedingt, dass das spektrale Verhalten verschiedener Objekte von zahlreichen Einflussfaktoren abhängt (z.B. vom Wettergeschehen, von der Jahreszeit oder von der Oberflächenbeschaffenheit von Objekten) und durch die Datenaufzeichnung über mehrere Zeitpunkte hinweg zuverlässigere Aussagen über die jeweilige Klassenzugehörigkeit möglich werden.
Im Rahmen dieses Beitrags wird die multispektrale und multitemporale Analyse von Sentinel-2 Daten genauer beleuchtet. Dabei werden Sentinel-2 Daten eines gemischt urban-ländlich geprägten Gebietes, welche für verschiedene Zeitpunkte vorliegen, sowohl pro Zeitpunkt als auch für die Kombination bezüglich verschiedener Zeitpunkte analysiert. Für jeden dieser Fälle werden anhand von definierten Trainingsgebieten zunächst die internen Parameter von verschiedenen Klassifizierungsverfahren ermittelt, wobei letztere durch einen Nearest-Neighbor-Klassifikator, einen naiven Bayes-Klassifikator, eine lineare Diskriminanz-Analyse, einen Random Forest und zwei Varianten der Support Vector Machine gegeben sind. Anschließend erfolgt pro Klassifizie-

[1] Karlsruher Institut für Technologie, Institut für Photogrammetrie und Fernerkundung, Englerstraße 7, D-76131 Karlsruhe, E-Mail: [Martin.Weinmann, Uwe.Weidner]@kit.edu

rungsverfahren eine Klassifizierung für die gesamte Szene, z.B. wie in Abb. 1 angedeutet, sowie eine Bewertung der erzielten Ergebnisse anhand von definierten Kontrollgebieten. Die erzielten Ergebnisse zeigen das Potential einer multitemporalen Analyse deutlich auf, wobei für die betrachtete Szene eine signifikante Steigerung der Gesamtgenauigkeit mit bis zu >10% erreicht wurde.
Nach einem kurzen Abriss des aktuellen Stands der Forschung (Kapitel 2) folgt eine detaillierte Erläuterung der verwendeten Methodik (Kapitel 3). Daraufhin werden die durchgeführten Experimente beschrieben (Kapitel 4) und die erzielten Ergebnisse diskutiert (Kapitel 5). Abschließend folgen eine Zusammenfassung und ein Ausblick (Kapitel 6).

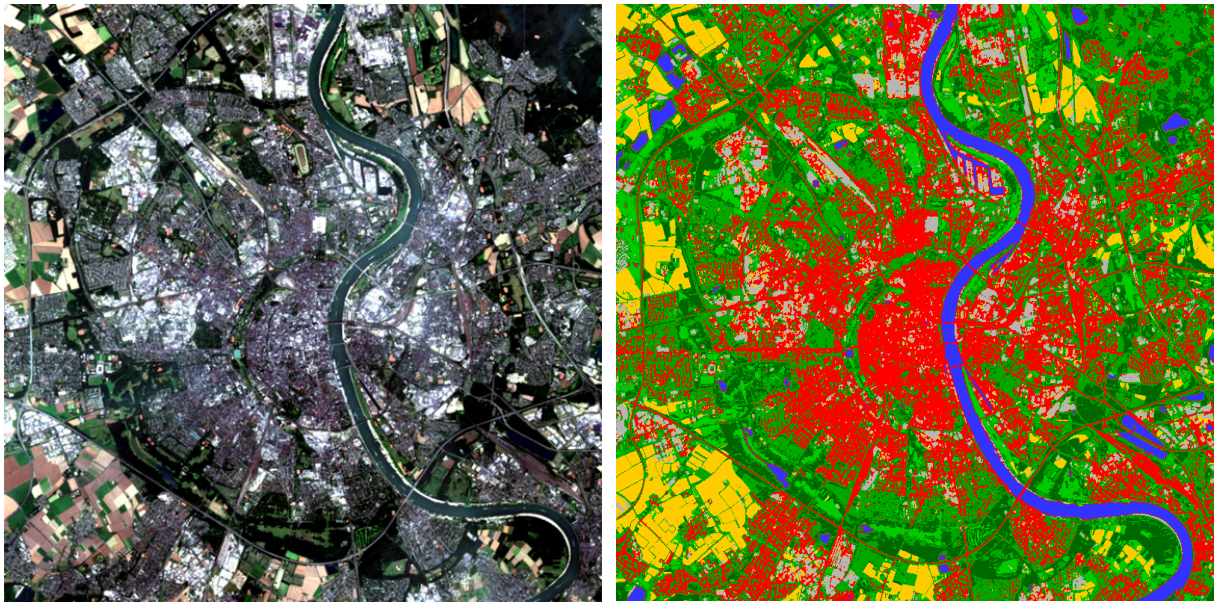


Abb. 1: Eine betrachtete Szene (links) und das entsprechende Klassifizierungsergebnis (rechts), wobei die Klassen durch *Siedlung* (rot), *Industrie* (grau), *Feld* (gelb), *Wald* (dunkelgrün), *Garten* (grün), *Wiese* (hellgrün) und *Wasser* (blau) gegeben sind

2 Stand der Forschung

Unter der Annahme von gegebenen repräsentativen Trainingsdaten kann die Klassifizierung von Sentinel-2 Daten bezüglich diverser Landbedeckungs- und Landnutzungsklassen in Analogie zur Klassifizierung von anderen Typen von Satellitenbildern durchgeführt werden. Beispielsweise können bei Sentinel-2 Daten, welche für einen Zeitpunkt erfasst wurden und typischerweise als Bildstapel vorliegen, die Reflektanzwerte der einzelnen Spektralkanäle pro Pixel als Einträge von Merkmalsvektoren dienen. Demgegenüber können auch spektrale Indizes wie z.B. der Normalized Difference Vegetation Index (NDVI) oder der Normalized Difference Water Index (NDWI) für die Definition von Merkmalsvektoren genutzt werden sowie eine bestimmte Auswahl von relevanten Kanälen (ZHANG et al. 2017). Die definierten Merkmalsvektoren können wiederum als Grundlage, d.h. als Eingangswerte, für diverse Klassifizierungsverfahren genutzt werden, wobei Standardklassifizierungsverfahren durch eine Support Vector Machine (CORTES

& VAPNIK 1995) oder einen Random Forest (BREIMAN 2001) gegeben und mittlerweile in vielen Softwarepaketen vorhanden sind.
Um bei der Analyse von Satellitenbildern zusätzlich den Einfluss verschiedener Jahreszeiten zu berücksichtigen, die sich insbesondere auf die Erscheinung von Vegetation in den entsprechenden Daten auswirken, wird in der Regel eine multitemporale Analyse durchgeführt. Hierbei werden für ein betrachtetes Gebiet die korrespondierenden Daten für verschiedene Zeitpunkte analysiert, wobei beispielsweise für jedes Pixel die Reflektanzwerte der einzelnen Spektralkanäle pro Zeitpunkt zur Definition von Merkmalsvektoren genutzt werden. Die resultierenden Merkmalsvektoren können dann wiederum mit Standardklassifizierungsverfahren wie z.B. einer Support Vector Machine oder einem Random Forest den entsprechend definierten Klassen zugeordnet werden (GÓMEZ et al. 2016). Solche Klassifizierungsstrategien lassen sich auch direkt auf die multitemporale Analyse von Sentinel-2 Daten übertragen, was beispielsweise zur Schätzung der Verteilung von Baumarten in bewaldeten Ökosystemen genutzt wurde (KARASIAK et al. 2017). Desweiteren kann mittels einer multitemporalen Analyse von Sentinel-2 Daten eine sehr genaue Analyse von landwirtschaftlichen Nutzflächen betrieben werden (BRUZZONE et al. 2017). Hierbei wurde eine Support Vector Machine zur Klassifizierung genutzt und, um zusätzlich noch limitierende Faktoren durch Störungen in den Daten (etwa durch Wolkenbedeckung) und die gegebene temporale Auflösung der vorliegenden Sentinel-2 Daten zu berücksichtigen, eine Rekonstruktion der Zeitreihe durchgeführt. Letzteres gelang durch den Einsatz eines nicht-parametrischen Regressions-Modells in Form eines Multi-Layer Perceptrons, mit welchem eine Sentinel-2 Zeitreihe in Tagesschritten rekonstruiert wurde. Neben der spektralen und temporalen Information in den Daten wurden im Rahmen einer anderen Studie (SOLANO-CORREA et al. 2017) auch räumliche Informationen in den Satellitenbildern mittels Zusammenhangskomponenten einbezogen, was die Separierung von einzelnen Feldern und deren Analyse mit einer hohen zeitlichen Auflösung ermöglicht.

3 Methodik

Die verwendete Methodik zur Analyse von multitemporalen Sentinel-2 Daten basiert auf der Verwendung von Reflektanzwerten, die zu Spektralkanälen korrespondieren und zur Definition von Merkmalsvektoren genutzt werden (Kapitel 3.1). Die Merkmalsvektoren dienen wiederum als Eingangsgrößen für ein Klassifizierungsverfahren, welches die jeweiligen Klassenzugehörigkeiten ermittelt (Kapitel 3.2).

3.1 Merkmalsextraktion

Anhand der für ein Pixel gegebenen Sentinel-2 Daten werden zur Definition von korrespondierenden Merkmalsvektoren die Reflektanzwerte der einzelnen Spektralkanäle als Merkmale betrachtet und in einem Vektor gespeichert. Bei der Betrachtung von multitemporalen Sentinel-2 Daten wird der Merkmalsvektor über die Reflektanzwerte der einzelnen Spektral-kanäle für alle betrachteten Zeitpunkte definiert.
Im Rahmen dieser Arbeit werden Reflektanzwerte betrachtet, welche zu den Spektralkanälen C2 (zentrale Wellenlänge: 490 nm), C3 (560 nm), C4 (665 nm), C5 (705 nm), C6 (740 nm), C7 (783 nm), C8a (865 nm), C11 (1610 nm) und C12 (2190 nm) korrespondieren. Reflektanzwerte, wel-

che zu den Spektralkanälen C1 (443 nm), C9 (945 nm) und C10 (1375 nm) korrespondieren, werden nicht berücksichtigt, da sie Bereiche im elektromagnetischen Spektrum abdecken, in denen die atmosphärische Durchlässigkeit nur gering ist oder der Wasserdampf das Verhalten dominiert. Der Spektralkanal C8 (842 nm) wird nicht berücksichtigt, da er mit dem Spektralkanal C8a überlappt, aber deutlich breiter und daher weniger charakteristisch ist.
Für die Reflektanzwerte aller betrachteten Spektralkanäle erfolgt eine atmosphärische Korrektur mittels der Sen2Cor Software (http://step.esa.int/main/third-party-plugins-2/sen2cor/).

3.2 Klassifizierung

Bei der Klassifizierung wird auf eine überwachte Klassifizierung fokussiert, d.h. anhand von repräsentativen Trainingsdaten wird ein Klassifizierungsverfahren trainiert. Erst danach kann eine Klassifizierung von weiteren Merkmalsvektoren durchgeführt werden. Im Rahmen dieser Arbeit werden verschiedene gängige Klassifizierungsverfahren verwendet.

3.2.1 Nearest-Neighbor-Klassifikator

Beim Nearest-Neighbor-Klassifikator (NN) erfolgt kein Training im eigentlichen Sinn. Vielmehr wird jeder zu klassifizierende Merkmalsvektor mit allen Merkmalsvektoren aus den Trainingsdaten verglichen, wozu oftmals der Euklidische Abstand als Ähnlichkeitsmetrik verwendet wird, und es wird dann die Klassenzugehörigkeit des ähnlichsten Merkmalsvektors aus den Trainingsdaten zugewiesen. Demzufolge ist es nicht erforderlich, eine Parameterschätzung durchzuführen oder vorab bestimmte Modellannahmen zu treffen. Allerdings wirkt sich gerade eine große Menge an durchzuführenden Vergleichen negativ auf den Rechenaufwand aus, da dieser dann in der Regel beachtlich ansteigt.

3.2.2 Naiver Bayes-Klassifikator

Beim naiven Bayes-Klassifikator (NB) handelt es sich um ein statistisches Klassifizierungsverfahren, welchem ein probabilistisches Modell basierend auf dem Bayes'schen Theorem zugrunde liegt und bei welchem angenommen wird, dass alle betrachteten Merkmale statistisch unabhängig voneinander sind. In der Trainingsphase werden anhand der Trainingsdaten die Wahrscheinlichkeit des Auftretens von Klassenzugehörigkeiten sowie die Wahrscheinlichkeiten von Merkmalen für diese Klassenzugehörigkeiten geschätzt. Damit kann dann für jeden zu klassifizierenden Merkmalsvektor die wahrscheinlichste Klassenzugehörigkeit ermittelt werden.

3.2.3 Lineare Diskriminanz-Analyse

Bei der linearen Diskriminanz-Analyse (LDA) handelt es sich um ein probabilistisches Klassifizierungsverfahren, bei welchem in der Trainingsphase eine multivariate Gauß-Verteilung an die gegebenen Trainingsdaten angepasst wird, d.h. es werden für jede Klasse die Parameter einer Gauss-Verteilung geschätzt. Vereinfachend wird hierbei angenommen, dass für alle Klassen identische Kovarianzmatrizen vorliegen und somit lediglich die Mittelwerte variieren. Für jeden zu klassifizierenden Merkmalsvektor werden dann die Wahrscheinlichkeiten ausgewertet, mit denen dieser Merkmalsvektor zu den verschiedenen Klassen gehört, und es wird dementsprechend die wahrscheinlichste Klassenzugehörigkeit zugewiesen.

3.2.4 Random Forest

Bei einem Random Forest (RF) (BREIMAN 2001) besteht die Grundidee in der strategischen Kombination von vielen schwachen Hypothesen zu einer starken Hypothese. In der Trainingsphase wird über ein Bootstrap Aggregating („Bagging“) ein Ensemble von randomisierten Entscheidungsbäumen erzeugt. Hierbei wird jeder Entscheidungsbaum trainiert, indem eine sukzessive Aufteilung der jeweils betrachteten Trainingsdaten in immer kleinere Untermengen durchgeführt wird, so dass jede Untermenge „so rein wie möglich“ ist. Diese Aufteilung wird in Form einer Sequenz von Entscheidungen formalisiert. Für jeden zu klassifizierenden Merkmals-vektor wird dann eine Sequenz von Entscheidungen durchlaufen, bis ein Blatt des Entscheidungsbaumes mit einer entsprechenden Hypothese bezüglich einer Klassenzugehörigkeit erreicht wird. Eine Betrachtung der schwachen Hypothesen („Votes“) über alle Entscheidungsbäume führt zu Klassenwahrscheinlichkeiten und der Mehrheitsentscheid („Majority Vote“) liefert eine starke Hypothese für eine geeignete Zuweisung einer Klassenzugehörigkeit.

3.2.5 Support Vector Machine

Bei der Support Vector Machine (SVM) (CORTES & VAPNIK 1995) müssen keine Annahmen über die Verteilung der gegebenen Daten getroffen werden. Stattdessen basiert die SVM auf der Idee, dass ein Klassifizierungsergebnis insbesondere dann gut ist, wenn es die einzelnen Klassen sehr gut separieren kann. Letzteres wird dann erreicht, wenn in der Trainingsphase zwischen den Datenpunkten, welche durch die Merkmalsvektoren gegeben sind, eine Trennfläche derart gezogen wird, dass der kürzeste Abstand zu den Datenpunkten einer Klasse maximiert wird. In der ursprünglichen Version ist eine SVM ein binäres Klassifizierungsverfahren, das darauf trainiert wird, zwei Klassen linear zu separieren. Im hochdimensionalen Merkmalsraum entspricht dies dem Einpassen einer Hyperebene, welche die Trainingsdaten entsprechend der Klassen aufteilt. Diese lineare Trennbarkeit stellt die Grundlage der linearen SVM (SVM-lin) dar. Sind die Daten nicht linear separierbar, wird der sogenannte „Kernel-Trick“ verwendet, mit dem eine nichtlineare Abbildung des Merkmalsraums in einen höherdimensionalen Merkmals-raum erfolgt, in dem die Daten linear separierbar sind. Hierzu wird oftmals eine radiale Basis-Funktion in Form einer Gauß-Funktion verwendet, was zu einer modifizierten SVM-Variante (SVM-rbf) führt. Eine Verallgemeinerung auf das Mehrklassen-Problem geschieht dadurch, dass mehrere binäre SVMs kombiniert werden. Hierbei kann z.B. für jedes Paar an Klassen eine eigene SVM trainiert werden, die entsprechende Merkmalsvektoren voneinander unterscheidet.

4 Experimente und Ergebnisse

Im Folgenden wird zunächst der verwendete Datensatz genauer beschrieben (Kapitel 4.1) und dann genauer auf die Bestimmung von Trainingsgebieten und Kontrollgebieten eingegangen (Kapitel 4.2). Anschließend werden die erzielten Ergebnisse vorgestellt (Kapitel 4.3).

4.1 Datensatz

Das betrachtete Gebiet ist in Abb. 2 dargestellt. Es hat eine Größe von 16 km × 16 km und enthält die Stadt Köln und deren Umgebung. Mit einer Ground Sampling Distance von 20 m lassen

sich die entsprechenden Sentinel-2 Daten als Bildstapel repräsentieren, wobei für jeden betrachteten Spektralkanal ein Bild mit 800 × 800 Pixeln vorliegt. Da im Rahmen dieser Arbeit lediglich die Spektralkanäle C2, C3, C4, C5, C6, C7, C8a, C11 und C12 betrachtet werden, ergibt sich somit pro Zeitpunkt ein Bildstapel bestehend aus neun Bildern. Insgesamt liegen Sentinel-2 Daten für die Zeitpunkte des 05.05.2016, 21.05.2016, 16.08.2016 und 26.08.2016 vor. Anhand dieser Sentinel-2 Daten soll eine Klassifizierung entsprechend der Klassen *Siedlung*, *Industrie*, *Feld*, *Wald*, *Garten*, *Wiese* und *Wasser* durchgeführt werden.

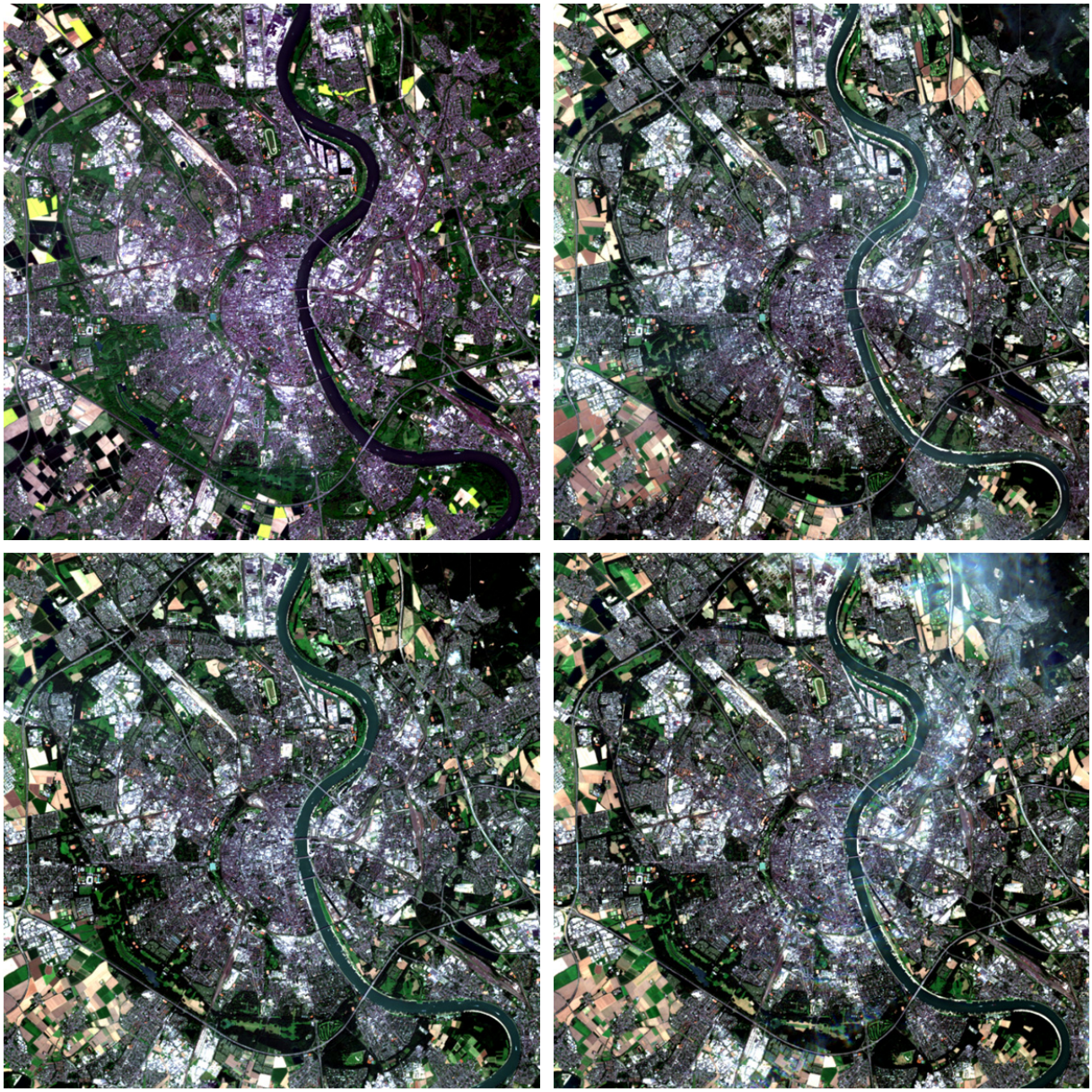


Abb. 2: Aus der Kanalkombination C4-C3-C2 abgeleitete Visualisierung für das betrachtete Gebiet um Köln zu den Zeitpunkten des 05.05.2016 (oben links), 21.05.2016 (oben rechts), 16.08.2016 (unten links) und 26.08.2016 (unten rechts)

4.2 Bestimmung von Trainingsgebieten und Kontrollgebieten

Zunächst wurden im Rahmen dieser Arbeit 34 Trainingsgebiete ausgewählt, welche den definierten Klassen entsprechen. Diese Trainingsgebiete umfassen insgesamt 4143 Pixel und entsprechen damit einer Fläche von 165,72 ha. Anhand dieser Trainingsgebiete mit bekannten Klassenzugehörigkeiten wurde pro Pixel ein Trainingsbeispiel bestehend aus dem entsprechenden Merkmalsvektor und der korrespondierenden Klassenzugehörigkeit definiert. Um eine quantitative Evaluierung der vorgestellten Methodik zu ermöglichen wurden zudem 57 Kontrollgebiete definiert, welche insgesamt 16859 Pixel und damit eine Fläche von 674,36 ha abdecken. Anhand der Kontrollgebiete mit bekannten Klassenzugehörigkeiten werden wiederum Testbeispiele definiert, die aus dem entsprechenden Merkmalsvektor und der korrespondierenden Klassenzugehörigkeit bestehen. Die entsprechenden Verteilungen über die einzelnen Klassen sind in Tab. 1 und Tab. 2 angegeben. Dabei wird deutlich, dass für die Klasse *Feld* prozentual gesehen sehr viele Trainingsbeispiele vorliegen, um die zeitliche Variabilität von Feldern möglichst gut abzudecken. Der hohe Anteil an Beispielen für die Klasse *Wasser* ist dadurch begründet, dass bei der Verwendung von statistischen Verfahren auch für diese Klasse eine gewisse Varianz vorliegt.

Tab. 1: Verteilung der Trainingsbeispiele für die definierten Klassen

Klasse	# Beispiele	Fläche [ha]	Anteil [%]
Siedlung	187	7,48	4,5
Industrie	207	8,28	5,0
Feld	2454	98,16	59,2
Wald	201	8,04	4,9
Garten	236	9,44	5,7
Wiese	233	9,32	5,6
Wasser	625	25,00	15,1
∑	4143	165,72	100,0

Tab. 2: Verteilung der Testbeispiele für die definierten Klassen

Klasse	# Beispiele	Fläche [ha]	Anteil [%]
Siedlung	5226	209,04	31,0
Industrie	2972	118,88	17,6
Feld	2168	86,72	12,9
Wald	1830	73,20	10,9
Garten	2149	85,96	12,7
Wiese	1132	45,28	6,7
Wasser	1382	55,28	8,2
∑	16859	674,36	100,0

4.3 Ergebnisse

Bei der Klassifizierung erfolgt eine Betrachtung von vier Fällen. Zunächst erfolgt jeweils eine separate Klassifizierung von Sentinel-2 Daten für die Zeitpunkte des 21.05.2016 und des 16.08.2016. Anschließend erfolgen eine Klassifizierung unter Berücksichtigung dieser beiden Zeitpunkte sowie eine Klassifizierung unter Berücksichtigung der Zeitpunkte des 05.05.2016, des 21.05.2016, des 16.08.2016 und des 26.08.2016. Die erzielten Klassifizierungsergebnisse sind in Tab. 3 bezüglich der Gesamtgenauigkeit, in Tab. 4 bezüglich des Kappa-Index und in Tab. 5 bezüglich des mittleren F_1-Wertes angegeben, wobei als Klassifizierungsverfahren ein Nearest-Neighbor-Klassifikator (NN), ein naiver Bayes-Klassifikator (NB), eine lineare Diskriminanz-Analyse (LDA), ein Random Forest (RF), eine lineare SVM (SVM-lin) und eine nichtlineare SVM (SVM-rbf) verwendet wurden. Anhand dieser Ergebnisse wird ersichtlich, dass für alle verwendeten Klassifizierungsverfahren die multitemporale Analyse insbesondere unter Berücksichtigung der Sentinel-2 Daten für alle vier Zeitpunkte zu einer deutlichen Verbesserung der Klassifizierungsergebnisse führt, wobei die Verbesserung bezüglich der Gesamtgenauigkeit, des Kappa-Index und des mittleren F_1-Wertes jeweils bis zu >10% beträgt. Die besten Ergebnisse werden jeweils mit dem Nearest-Neighbor-Klassifikator erreicht und das jeweilige Klassifizierungsergebnis für das gesamte betrachtete Gebiet ist in Abb. 3 visualisiert.

Tab. 3: Gesamtgenauigkeit, die mit den verschiedenen Klassifizierungsverfahren erreicht wurde

berücksichtigte Zeitpunkte	NN	NB	LDA	RF	SVM-lin	SVM-rbf
21.05.	73,28%	60,63%	68,86%	71,15%	69,65%	70,85%
16.08.	79,09%	66,72%	70,08%	76,24%	76,67%	74,96%
21.05., 16.08.	80,54%	65,13%	73,55%	74,57%	75,81%	75,68%
05.05., 21.05., 16.08., 26.08.	83,62%	68,28%	75,65%	80,54%	80,98%	79,19%

Tab. 4: Kappa-Index, der mit den verschiedenen Klassifizierungsverfahren erreicht wurde

berücksichtigte Zeitpunkte	NN	NB	LDA	RF	SVM-lin	SVM-rbf
21.05.	68,15%	53,97%	62,93%	65,92%	64,22%	65,54%
16.08.	74,41%	60,28%	63,89%	70,94%	71,48%	69,49%
21.05., 16.08.	76,54%	58,88%	68,25%	69,49%	71,16%	70,95%
05.05., 21.05., 16.08., 26.08.	80,19%	62,38%	70,55%	76,72%	77,24%	74,99%

Tab. 5: Mittlerer F_1-Wert, der mit den verschiedenen Klassifizierungsverfahren erreicht wurde

berücksichtigte Zeitpunkte	NN	NB	LDA	RF	SVM-lin	SVM-rbf
21.05.	76,03%	65,92%	72,00%	75,70%	74,12%	74,65%
16.08.	79,31%	66,42%	72,34%	76,87%	77,68%	75,89%
21.05., 16.08.	82,02%	67,58%	77,06%	77,47%	78,79%	79,17%
05.05., 21.05., 16.08., 26.08.	86,53%	70,42%	80,16%	83,61%	84,65%	82,47%

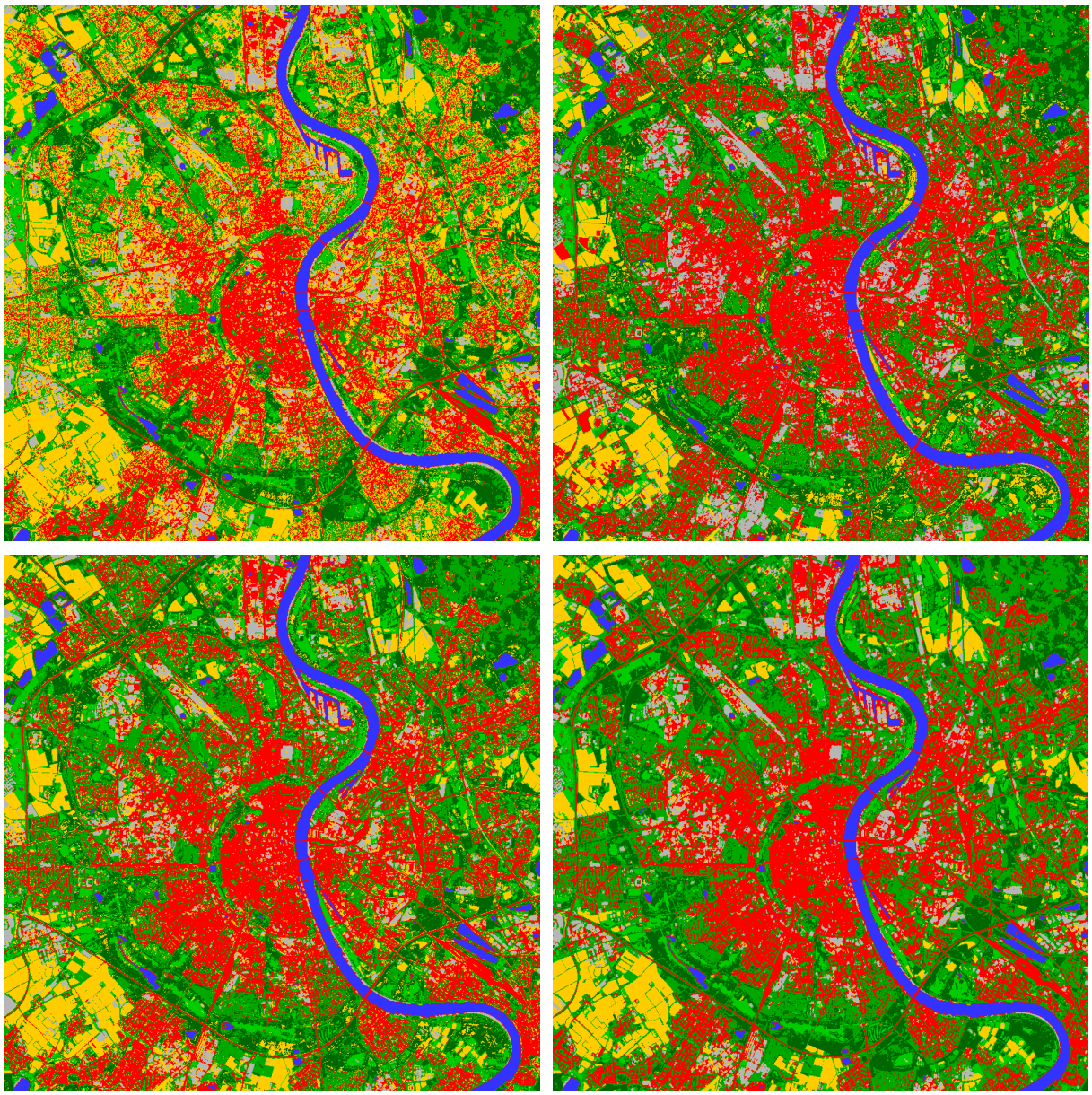


Abb. 3: Klassifizierungsergebnisse, die mit dem Nearest-Neighbor-Klassifikator erreicht wurden unter Berücksichtigung der Sentinel-2 Daten für den Zeitpunkt des 21.05.2016 (oben links), für den Zeitpunkt des 16.08.2016 (oben rechts), für die Zeitpunkte des 21.05.2016 und 16.08.2016 (unten links) und für die Zeitpunkte des 05.05.2016, des 21.05.2016, des 16.08.2016 und des 26.08.2016 (unten rechts). Die Farbcodierung erlaubt eine Unterscheidung gemäß der Klassen *Siedlung* (rot), *Industrie* (grau), *Feld* (gelb), *Wald* (dunkelgrün), *Garten* (grün), *Wiese* (hellgrün) und *Wasser* (blau)

5 Diskussion

Eine genauere Untersuchung der erzielten Klassifizierungsergebnisse (siehe Tab. 3, Tab. 4 und Tab. 5) zeigt, dass im Rahmen der durchgeführten Experimente die für den 21.05.2016 vorliegenden Sentinel-2 Daten weniger geeignet sind als die für den 16.08.2016 vorliegenden Sentinel-

2 Daten. Um die Ursache dafür zu ergründen, sind die jeweiligen Konfusionsmatrizen sowie daraus abgeleitete Evaluierungsmaße für die Ergebnisse des Nearest-Neighbor-Klassifikators in Tab. 6 und Tab. 7 angegeben. In diesen Tabellen wird erkennbar, dass für den 16.08.2016 insbesondere die Klassen *Feld*, *Garten* und *Siedlung* besser identifiziert werden können, während die Klasse *Wald* schlechter erkannt wird.

Tab. 6: Konfusionsmatrix für das Ergebnis des Nearest-Neighbor-Klassifikators unter Berücksichtigung der Sentinel-2 Daten für den Zeitpunkt des 21.05.2016

Klasse		**Schätzung**						
		Siedlung	**Industrie**	**Feld**	**Wald**	**Garten**	**Wiese**	**Wasser**
Referenz	**Siedlung**	3218	27	1455	0	524	1	1
	Industrie	159	1834	876	0	101	2	0
	Feld	0	1	1992	4	23	148	0
	Wald	1	0	4	1408	417	0	0
	Garten	0	0	298	6	1820	25	0
	Wiese	0	0	200	83	148	701	0
	Wasser	0	0	0	0	0	0	1382
Vollständigkeit		0,62	0,62	0,92	0,77	0,85	0,62	1,00
Korrektheit		0,95	0,98	0,41	0,94	0,60	0,80	1,00
F_1-Wert		0,75	0,76	0,57	0,85	0,70	0,70	1,00

Tab. 7: Konfusionsmatrix für das Ergebnis des Nearest-Neighbor-Klassifikators unter Berücksichtigung der Sentinel-2 Daten für den Zeitpunkt des 16.08.2016

Klasse		**Schätzung**						
		Siedlung	**Industrie**	**Feld**	**Wald**	**Garten**	**Wiese**	**Wasser**
Referenz	**Siedlung**	4264	140	51	12	758	0	1
	Industrie	611	2082	144	0	105	30	0
	Feld	382	43	1413	83	60	187	0
	Wald	0	0	242	1361	227	0	0
	Garten	10	4	13	41	2042	39	0
	Wiese	19	0	179	99	49	786	0
	Wasser	0	0	0	0	0	0	1382
Vollständigkeit		0,82	0,70	0,65	0,74	0,95	0,69	1,00
Korrektheit		0,81	0,92	0,69	0,85	0,63	0,75	1,00
F_1-Wert		0,81	0,79	0,67	0,79	0,76	0,72	1,00

Durch die Fusion von Sentinel-2 Daten, die zu verschiedenen Zeitpunkten erfasst wurden, wird besser berücksichtigt, dass das spektrale Verhalten verschiedener Objekte von zahlreichen Einflussfaktoren abhängt. Dadurch entsteht insbesondere bei der multitemporalen Analyse von Sentinel-2 Daten unter Berücksichtigung von vier Zeitpunkten eine deutliche Verbesserung der erzielten Klassifizierungsergebnisse (siehe Tab. 3, Tab. 4 und Tab. 5). Im Gegensatz zur Gesamtgenauigkeit, aus welcher lediglich eine globale Verbesserung der Güte des Klassifizierungsergebnisses ersichtlich wird, kann über den Kappa-Index auf eine bessere Trennbarkeit der einzelnen Klassen geschlossen werden, während der mittlere F_1-Wert hier ein Indikator dafür ist, dass die Erkennungsraten über alle Klassen hinweg relativ gut sind. Zur genaueren Analyse der Verbesserung sind für die Ergebnisse des Nearest-Neighbor-Klassifikators unter Berücksichtigung der beiden Zeitpunkte des 21.05.2016 und des 16.08.2016 sowie unter Berücksichtigung aller vier Zeitpunkte die jeweiligen Konfusionsmatrizen sowie daraus abgeleitete Evaluierungsmaße in Tab. 8 und Tab. 9 angegeben. Beispielsweise zeigen ein Vergleich der Konfusionsmatrizen in Tab. 6 und Tab. 9 sowie ein Vergleich der daraus abgeleiteten Evaluierungsmaße der Vollständigkeit, der Korrektheit und des F_1-Werts, dass durch die Berücksichtigung mehrerer Zeitpunkte insbesondere die Klassen *Feld* (+30% für den F_1-Wert), *Wiese* (+24% für den F_1-Wert) und *Wald* (+9% für den F_1-Wert) besser erkannt werden können. Die durch die Berücksichtigung von vier statt zwei Zeitpunkten resultierende Verbesserung der Klassifizierungsergebnisse ist gemäß Tab. 8 und Tab. 9 durch eine verbesserte Erkennung der Klassen *Wiese* (+14% für den F_1-Wert) und *Wald* (+7% für den F_1-Wert) begründet. Erwartungsgemäß bestätigt die Betrachtung der klassenweisen Evaluierungsmaße auch, dass die Klasse *Wasser* generell einfach zu identifizieren ist, während die Klassen *Garten* und *Industrie* selbst unter der Berücksichtigung von Sentinel-2 Daten für alle vier betrachteten Zeitpunkte im Vergleich zu den anderen Klassen noch relativ schlecht erkannt werden.

Die Gegenüberstellung der Ergebnisse, welche mit verschiedenen Klassifizierungsverfahren erzielt wurden, zeigt auch, dass der naive Bayes-Klassifikator keine zufriedenstellenden Klassifizierungsergebnisse liefert. Dies liegt insbesondere daran, dass die Grundannahme von voneinander statistisch unabhängigen Merkmalen bei den vorliegenden Daten nicht erfüllt ist. Damit ist die Eignung des in der Trainingsphase angepassten probabilistischen Modells nur unzureichend für das betrachtete Klassifizierungsproblem gegeben. Im Vergleich dazu liefert die lineare Diskriminanz-Analyse deutlich bessere Ergebnisse, jedoch ist durch die Annahme von einer zugrundeliegenden multivariaten Gauß-Verteilung mit identischen Kovarianzmatrizen für alle Klassen immer noch eine starke Einschränkung gegeben, welche für die betrachteten Sentinel-2 Daten im Allgemeinen nicht vorliegt. Solche stark einschränkenden Annahmen liegen beim Nearest-Neighbor-Klassifikator, beim Random Forest und bei der Support Vector Machine nicht vor, was in einer besseren Leistungsfähigkeit resultiert. Insbesondere der Nearest-Neighbor-Klassifikator ist für das vorliegende Klassifizierungsproblem geeignet, da die entsprechenden Ergebnisse immer noch besser sind als die eines Random Forest oder einer Support Vector Machine, obwohl für jene Klassifizierungsverfahren eine Optimierung von internen Parametern über einen geeigneten Suchraum erfolgt. Ferner lassen sich auch Vorteile einer linearen Support Vector Machine im Vergleich zur nicht-linearen Support Vector Machine feststellen.

Tab. 8: Konfusionsmatrix für das Ergebnis des Nearest-Neighbor-Klassifikators unter Berücksichtigung der Sentinel-2 Daten für die Zeitpunkte des 21.05.2016 und des 16.08.2016

Klasse		Schätzung						
		Siedlung	**Industrie**	**Feld**	**Wald**	**Garten**	**Wiese**	**Wasser**
Referenz	**Siedlung**	3932	28	285	1	979	0	1
	Industrie	397	1854	516	0	183	22	0
	Feld	0	3	1981	18	55	111	0
	Wald	0	0	55	1480	295	0	0
	Garten	3	0	16	11	2082	37	0
	Wiese	0	0	159	54	51	868	0
	Wasser	0	0	0	0	0	0	1382
Vollständigkeit		0,75	0,62	0,91	0,81	0,97	0,77	1,00
Korrektheit		0,91	0,98	0,66	0,95	0,57	0,84	1,00
F_1-Wert		0,82	0,76	0,76	0,87	0,72	0,80	1,00

Tab. 9: Konfusionsmatrix für das Ergebnis des Nearest-Neighbor-Klassifikators unter Berücksichtigung der Sentinel-2 Daten für alle vier betrachteten Zeitpunkte

Klasse		Schätzung						
		Siedlung	**Industrie**	**Feld**	**Wald**	**Garten**	**Wiese**	**Wasser**
Referenz	**Siedlung**	3973	35	72	0	1145	0	1
	Industrie	554	1957	265	1	192	3	0
	Feld	0	149	1955	0	25	39	0
	Wald	0	0	0	1660	166	4	0
	Garten	0	0	0	27	2092	30	0
	Wiese	0	0	32	19	3	1078	0
	Wasser	0	0	0	0	0	0	1382
Vollständigkeit		0,76	0,66	0,90	0,91	0,97	0,95	1,00
Korrektheit		0,88	0,91	0,84	0,97	0,58	0,93	1,00
F_1-Wert		0,81	0,77	0,87	0,94	0,72	0,94	1,00

6 Zusammenfassung und Ausblick

In diesem Beitrag wurde die multitemporale Analyse von Sentinel-2 Daten bezüglich diverser Landbedeckungs- und Landnutzungsklassen behandelt. Ausgehend von den Reflektanzwerten der einzelnen Spektralkanäle wurde mittels verschiedener Klassifizierungsverfahren eine semantische Interpretation eines betrachteten Gebiets durchgeführt, wobei die für verschiedene Zeitpunkte vorliegenden Sentinel-2 Daten separat und in Kombination betrachtet wurden. Die erzielten Ergebnisse zeigten eine deutliche Verbesserung der Klassifizierungsergebnisse für den Fall

einer multitemporalen Analyse von Sentinel-2 Daten im Gegensatz zur Berücksichtigung von Sentinel-2 Daten für lediglich einen Zeitpunkt. Dabei wurde ersichtlich, dass insbesondere die Klassen *Feld*, *Wiese* und *Wald* durch eine multitemporale Analyse besser erkennbar sind.
Im Rahmen weiterer Untersuchungen wird der Einbezug von Verfahren zur Merkmalsselektion angestrebt (KELLER et al. 2016), was eine Bewertung der Relevanz einzelner Merkmale bezüglich des betrachteten Klassifizierungsproblems erlaubt. Dadurch können die für das untersuchte Klassifizierungsproblem irrelevanten und zum Teil auch die redundanten Spektralbereiche identifiziert und dementsprechend bei der Klassifizierung ignoriert werden, was in der Regel zu einer deutlichen Verbesserung der Klassifizierungsergebnisse führt.

7 Literaturverzeichnis

BREIMAN, L., 2001: Random forests. Machine Learning, **45**(1), 5-32.

BRUZZONE, L., BOVOLO, F., PARIS, C., SOLANO-CORREA, Y.T., ZANETTI, M. & FERNÁNDEZ-PRIETO, D., 2017: Analysis of multitemporal Sentinel-2 images in the framework of the ESA Scientific Exploitation of Operational Missions. Proceedings of the 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 1-4.

CORTES, C. & VAPNIK, V., 1995: Support-vector networks. Machine Learning, **20**(3), 273-297.

GÓMEZ, C., WHITE, J.C. & WULDER, M. A., 2016: Optical remotely sensed time series data for land cover classification: a review. ISPRS Journal of Photogrammetry and Remote Sensing, **116**, 55-72.

KARASIAK, N., SHEEREN, D., FAUVEL, M., WILLM, J., DEJOUX, J.F. & MONTEIL, C., 2017: Mapping tree species of forests in southwest France using Sentinel-2 image time series. Proceedings of the 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 1-4.

KELLER, S., BRAUN, A. C., HINZ, S. & WEINMANN, M., 2016: Investigation of the impact of dimensionality reduction and feature selection on the classification of hyperspectral EnMAP data. Proceedings of the 8th IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-5.

KUKAWSKA, E., LEWINSKI, S., KRUPINSKI, M., MALINOWSKI, R., NOWAKOWSKI, A., RYBICKI, M. & KOTARBA, A., 2017: Multitemporal Sentinel-2 data – Remarks and observations. Proceedings of the 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 1-4.

SOLANO-CORREA, Y.T., BOVOLO, F., BRUZZONE, L. & FERNÁNDEZ-PRIETO, D., 2017: Spatiotemporal evolution of crop fields in Sentinel-2 satellite image time series. Proceedings of the 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 1-4.

SPOTO, F., SY, O., LABERINTI, P., MARTIMORT, P., FERNANDEZ, V., COLIN, O., HOERSCH, B. & MEYGRET, A., 2012: Overview of Sentinel-2. Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 1707-1710.

ZHANG, T., SU, J., LIU, C., CHEN, W.H., LIU, H. & LIU, G., 2017: Band selection in Sentinel-2 satellite for agriculture applications. Proceedings of the 23rd International Conference on Automation and Computing (ICAC), 1-6.

A Service Architecture for Processing Big Earth Data in the Cloud with Geospatial Analytics and Machine Learning

WOLFGANG GLATZ & THOMAS BAHR[1]

Abstract: The Geospatial Services Framework (GSF) brings together data, geospatial analytics, and computing power in the cloud to enable the deployment of applications. The service architecture given by GSF is used by web clients to access and analyze on-demand remotely sensed data as well as for the automated, permanent processing of big geospatial data. GSF can be integrated in any public or internal server environments.
GSF is based on the concept of service engines and their workers. Harris provides the ready-to-use ENVI/IDL/SARscape/Machine Learning service engines. Available ENVI analytics include feature extraction, object identification, change detection, and classification. A specific machine learning algorithm for spectral-based land cover mapping is the Softmax Regression classifier. Harris machine learning contains deep learning capabilities, which focus on object recognition within scenes. They are designed for the unique characteristics of space- and airborne imagery of multiple modalities, and point cloud data sets.

1 Introduction

A continually increasing, massive amount of geospatial data, i.e. Big Earth Data, from different sources (commercial satellite constellations and small satellites, drones) and modalities (optical: Pan, RGB, MSI, HSI; SAR; LiDAR), enforces the automation of data processing. New tools and technologies are needed for hosting and managing distributed data processing in a high-performance computing environment within an enterprise or in the cloud.
The Geospatial Services Framework (GSF) brings together data, geospatial analytics, and computing power in the cloud to enable the deployment of applications, which solve problems at scale across industries.

2 Geospatial Services Framework (GSF)

GSF is based on the concept of service engines and their workers (HARRIS GEOSPATIAL SOLUTIONS 2017). Each worker uses multiple CPUs for parallel processing, the workers are processes executed in a dynamically scalable cluster of machines. Harris provides the ready-to-use ENVI/IDL/SARscape/Machine Learning service engines (BAHR & OKUBO 2013). In addition, customers may implement their own engines. The service architecture given by GSF is used by web clients to access and analyze on-demand remotely sensed data as well as for the automated, permanent processing of big geospatial data (Fig. 1).

[1] Harris Geospatial Solutions GmbH, Talhofstraße 32a, D-82205 Gilching,
E-Mail: [Wolfgang.Glatz, Thomas.Bahr]@harris.com

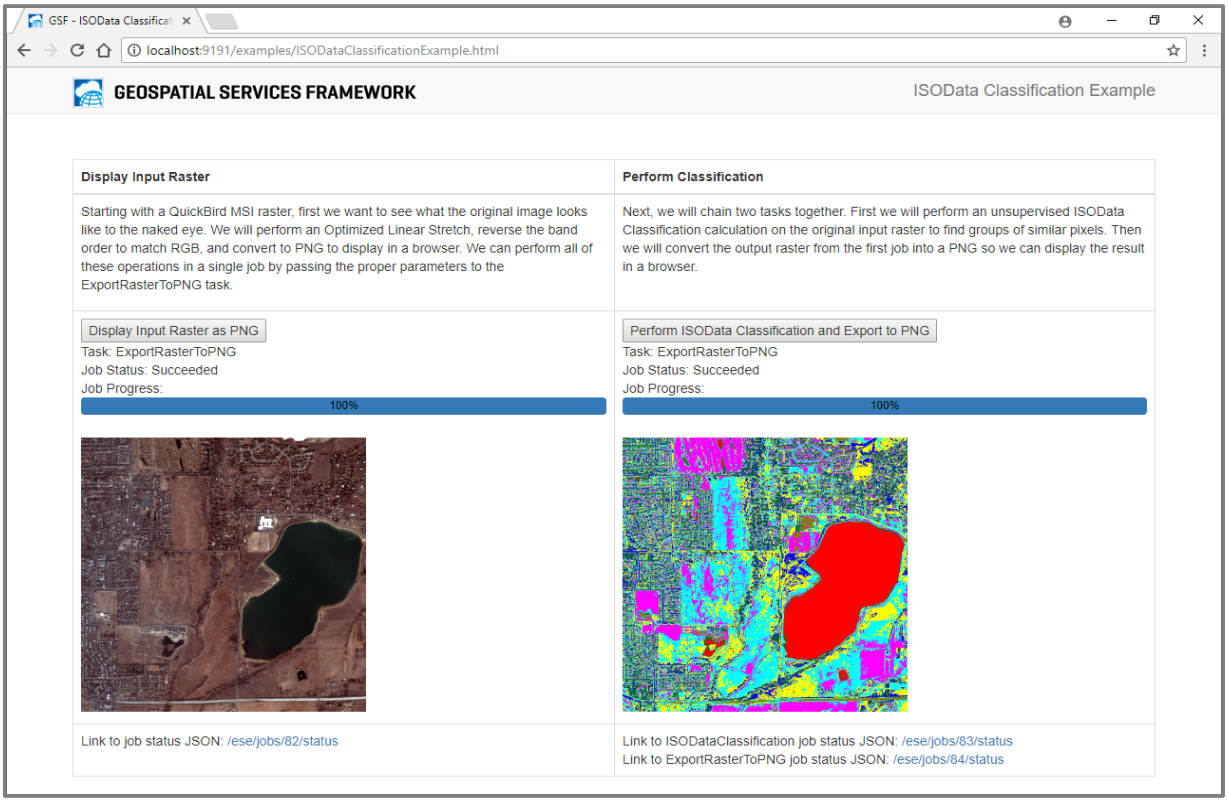


Fig. 1: GSF example: Web client showing the result of a land use classification processed in the cloud

GSF can be integrated in any public or internal server environments, such as Amazon Web Services (AWS), Microsoft Azure, or the Google Cloud Platform (Fig. 2).

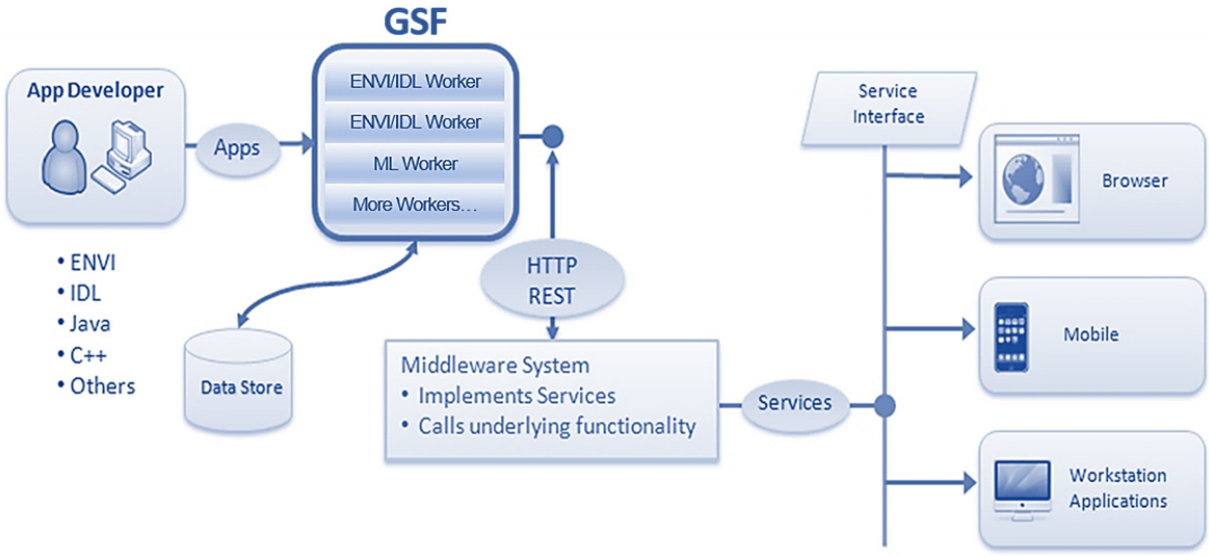


Fig. 2: GSF: A scalable framework for geospatial web services, using highly parallelized clusters of ENVI/IDL workers, Machine Learning (ML) workers, and other workers (processors)

Developers may use GSF to easily publish custom algorithms for the hosted service engines. These processing workflows can then be shared across the enterprise or cloud. For analysis of remotely sensed data, developers can resort to the full width of ENVI software analytics.
One of the essential features of GSF is the adaptation to varying utilization. On demand, the operator of the service architecture can add additional parallel workers (scalability) and additional geospatial workflows, if the amount of data and the number of service requests from any clients increases (Fig. 3).

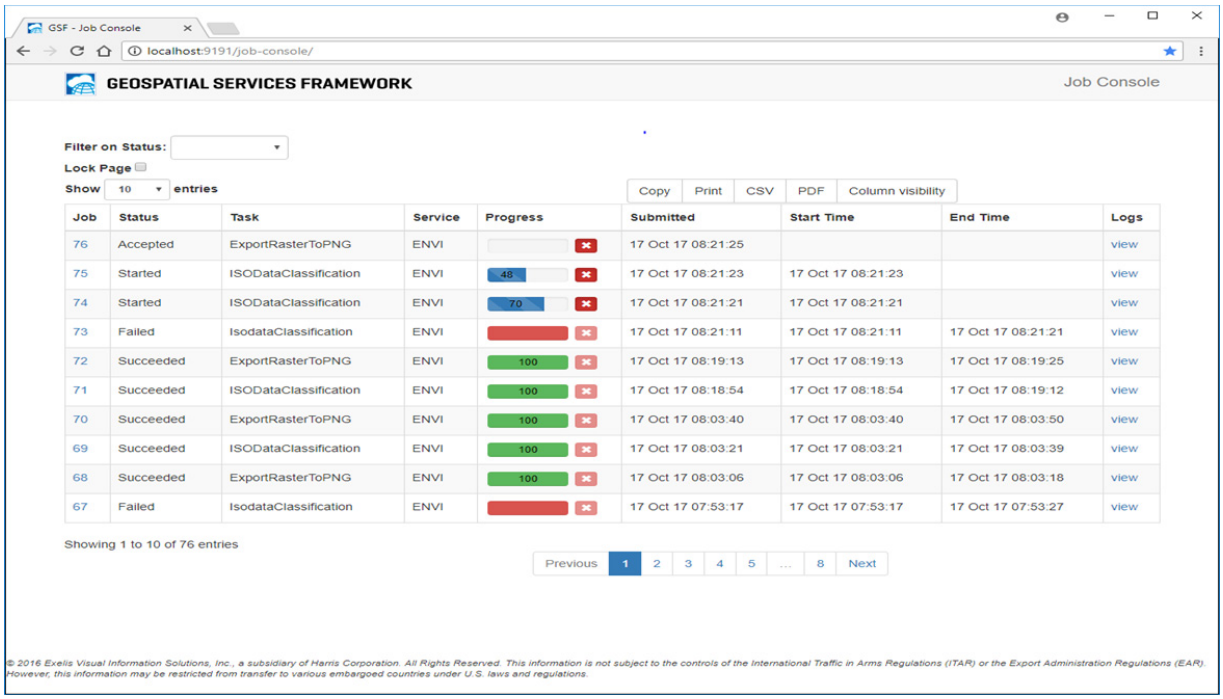


Fig. 3: GSF Job Console: State list of the geospatial analytics jobs executed by the GSF

3 ENVI Software Analytics

ENVI combines image processing and analysis technology to derive detailed information from all geospatial data, i.e. optical imagery, SAR, and LiDAR. Available analytics include feature extraction, classification, object identification, change detection, and more (HARRIS GEOSPATIAL SOLUTIONS 2018). A specific machine learning algorithm for spectral-based land cover mapping is the Softmax Regression classifier (WOLFE et al. 2017). It can be created and trained on a reference dataset using spectral and spatial information and then be applied to similar data multiple times (Fig. 4). Implemented in a classification framework, it provides a flexible approach to customize a classification process.
All described geoprocessing tools are capable of being fully integrated with ArcGIS® for Server from ESRI via a Python client library.

Fig. 4: Softmax Regression classifier: Merge of two classification images (bottom). The classifier was trained on one attribute image (top right) and then applied to a similar attribute image (top left)

4 Deep Learning Capabilities

Harris machine learning contains deep learning capabilities, which focus on object recognition within scenes.

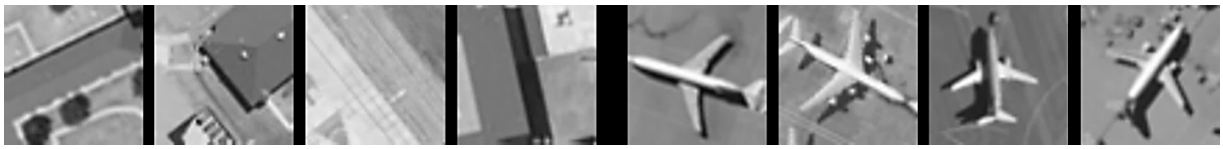


Fig. 5: Training of the Deep Learning network for recognition of commercial airplanes with negative (left) and positive (right) training chips

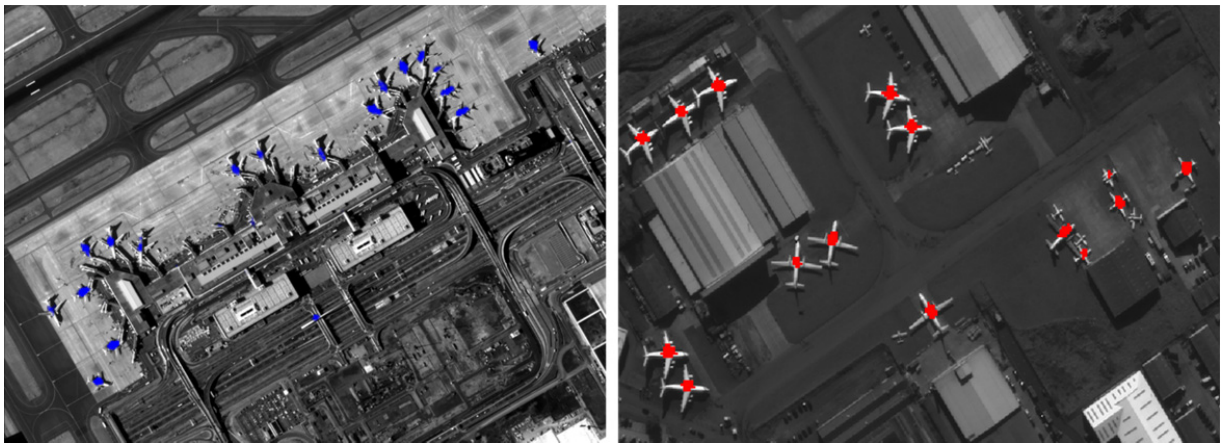


Fig. 6: Application of the trained network to different data for plane recognition: Ikonos (left), WorldView-3 (right)

They are designed for the unique characteristics of space- and airborne imagery of multiple modalities, as well as point cloud data sets. Successful sample applications, for instance on Pan imagery, included the detection of airplanes, storage tanks, cooling towers, athletic fields, paved roads, overpasses, and tollbooths (Fig. 5, 6).

5 Conclusion

The Geospatial Services Framework (GSF) brings together data, geospatial analytics, and computing power in the cloud for processing remotely sensed data with geospatial analytics and machine learning. In particular, the Harris modules for Deep Learning can be applied for automated object recognition processes and executed as GSF services.
Overall, GSF is a substantial contribution to operational, spatio-temporal analytics of Big Earth Data.

6 References

BAHR, T. & OKUBO, B., 2013b: A New Cloud-based Deployment of Image Analysis Functionality. GI_Forum 2013, Jekel, T., Car, A., Strobl, J. & Griesebner, G. (Eds.), Herbert Wichmann Verlag, VDE VERLAG GMBH, Berlin/Offenbach, ÖAW Verlag, Wien, 243-250.

HARRIS GEOSPATIAL SOLUTIONS (Ed.), 2017: Geospatial Services Framework Documentation. Geospatial Services Framework v2.1.

HARRIS GEOSPATIAL SOLUTIONS (Ed.), 2018: ENVI Documentation, ENVI v5.5.

STANFORD UNIVERSITY (Ed.), 2013: Unsupervised Feature Learning and Deep Learning. http://ufldl.stanford.edu/wiki/ index.php/Softmax_Regression (updated April 2013, accessed September 2016).

WOLFE, J., JIN, X., BAHR, T. & HOLZER, N., 2017: Application of Softmax Regression and its Validation for Spectral-Based Land Cover Mapping. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **42**(1/W1), 455-459.

Registrierung von Flugzeug getragenen Kameraaufnahmen und UAV-Aufnahmen zur Anreicherung von 3D Daten

RICHARD BOERNER[1], YUSHENG XU[1], LUDWIG HOEGNER[1] & UWE STILLA[1]

Zusammenfassung: Dieser Beitrag beschreibt ein Verfahren zur automatischen Bestimmung eines initialen Modells für eine anschließende Feinregistrierung von gemeinsam aufgenommenen ALS-Punktwolken und Luftbildern mit einer weiteren, von einem UAV nicht zeitgleich aufgenommenen Bildsequenz. Dabei wird der Ansatz verfolgt, geometrische Features ohne radiometrische Information zur Koregistrierung zu verwenden. Hier soll sowohl auf Passpunkte als auch auf manuell festgelegte Verknüpfungspunkte verzichtet werden. Als geometrische Elemente dienen Geradenstücke. Die Geradenstücke werden in den Luftbildern gesucht zu Strukturen von paarweise Parallelen zusammengefasst und anschließend zur Bestimmung der Affintransformation zwischen den Bildern verwendet. Als Testdatensatz dient ein Waldstück des Mangfallgebietes südöstlich von München. Hier ist es mit dem vorgestellten Verfahren möglich eine initiale Orientierung mit einer Genauigkeit von 20 Pixeln in einem Gebiet zu schätzen, das nur natürliche Strukturen aufweist, die sich in den geschätzten Geradenstücken als wiederkehrende Muster generalisieren lassen.

1 Einleitung

Sollen großräumige Flächen 3-dimensional erfasst werden, wird oft eine Befliegung durchgeführt, bei der sowohl Luftbilder als auch Laserscans gemeinsam mit einem GPS/INS-System aufgenommen werden. Die Kamera und der Laserscanner liegen kalibriert in einem gemeinsamen Koordinatensystem vor. Aus diesen Daten wird eine kolorierte 3D-Punktwolke mit Georeferenzierung erzeugt. Eine weitere Variante ist die Befliegung mit einem unbemannten Luftfahrzeug (unmaned aerial vehicle: UAV). Die Verwendung eines UAVs ist kostengünstiger und unabhängiger von Witterungsbedingungen. Des Weiteren wird durch eine geringere Flughöhe eine ebenfalls geringere Bodenpixelgröße erzielt. Eine aus einer 3D-Rekonstruktion resultierende Punktwolke der vom UAV aufgenommenen Bilder besitzt also eine höhere Punktdichte als die Punktwolke aus der Flugzeug-Befliegung. Auf der anderen Seite sind in UAVs zumeist kleinere und günstigere GNSS-Antennen und günstigere IMUs verbaut, wodurch die Orientierungsparameter im Vergleich zur Flugzeug-Befliegung ungenauer sind und lediglich als Näherungswerte angenommen werden können. Die Schwierigkeit einer Registrierung von UAV-basierten und flugzeugbasierten Aufnahmen besteht darin, dass die Daten nicht zwangsläufig zeitgleich aufgenommen werden und dadurch sowohl geometrische als auch vor allem radiometrische Unterschiede in der Szene vorliegen können. Des Weiteren existieren große Maßstabsunterschiede im Fall einer Bild-zu-Bild-Zuordnung und Unterschiede in der Geometrie im Fall einer Punktwolke-Bild-Zuordnung. Eine Kombination beider Sensoren würde folglich eine lokal höhere zeitliche und räumliche Auflösung von großräumigen Referenzdaten erzielen.

[1] Technische Universität München, Photogrammetrie und Fernerkundung , Arcisstraße 21, D-80333 München, E-Mail: [richard.boerner, yusheng.xu, ludwig.hoegner, stilla]@tum.de

Eine Zuordnung auf Grundlage von radiometrischen Verteilungen gestaltet sich folglich schwierig. Stattdessen wird die geometrische Verteilung von markanten Punkten in beiden Bildern für die Suche nach Korrespondenzen verwendet. Das vorgestellte Verfahren dient dazu, den Suchraum bei solch einer rein geometrischen Zuordnung durch eine grobe Vororientierung zu reduzieren. Zur Reduktion der Komplexität bei der Zuordnung wird hierfür auf Geradenstücke zurückgegriffen. Es wird davon ausgegangen, dass Geradenstücke markante Strukturen in der Szene auch für natürliche Umgebungen zumindest grob annähern und diese markanten Geradenstücke sowohl in den Luftbildern als auch in den vom UAV aus aufgenommenen Bildern auftreten und vergleichbar sind. Die Geradenstücke werden zu einer Struktur von paarweise parallelen Geraden zusammengefasst und über deren Abstand den Maßstab abzuschätzen. Anschließend werden in einem RANSAC-basierten Verfahren die Geradenpaare aus beiden Bildern zugeordnet. Hier wird eine räumliche Abgrenzung von potentiell zuordenbaren Geraden durch eine Segmentierung von Linienelementen vorgenommen. Mit einer Homographie wird die Affintransformation zwischen den Bildern auf Grundlage der zugeordneten Geraden bestimmt. Die resultierenden Parameter der groben Orientierung können in einer folgenden Feinregistrierung mit einem stark eingeschränkten verbleibenden Suchraum für eine geometrische Verknüpfung von markanten Punkten verwendet werden. In zukünftigen Arbeiten sollen die so berechneten Affinparameter dazu verwendet werden, um aus der ALS-Punktewolke ein Tiefenbild aus Sicht der Luftbildkamera zu erzeugen und dieses Tiefenbild in das Bild der UAV-Kamera zu übertragen. Die übertragenen Entfernungswerte des Tiefenbildes liefern somit 3D-Informationen zu den Pixeln des Bildes der UAV-Kamera, welche mittels dieser 3D-Informationen im Raum neu orientiert werden kann, um anschließend die vorhandene 3-Punktwolke zu verdichten.
Abschnitt 2 stellt einen kurzen Überblick über vorangegangen Arbeiten im Bereich der Registrierung von 3D- und 2D-Daten dar. Abschnitt 3 beschreibt die vorgestellte Methode und Abschnitt 4 zeigt die verwendeten Daten. Im Anschluss werden die Ergebnisse unter Abschnitt 5 vorgestellt und unter Abschnitt 6 folgt eine Zusammenfassung und Ausblick auf zukünftige Arbeiten.

2 Vorangegangene Arbeiten

Die Thematik der Multi-Sensor-Registrierung wird aktuell auf zwei unterschiedliche Wegen gelöst: im 2D-Raum durch Rückprojektion der 3D-Information in eine Ebene und Registrierung 2er rückprojizierter Bilder oder direkt im 3D-Raum. In beiden Fällen gestaltet sich die Verwendung von radiometrischer Information aufgrund der unterschiedlichen radiometrischen Eigenschaften der unterschiedlichen Sensoren schwierig. Somit müssen geometrische Eigenschaften über die räumliche Verteilung der markanten Punkte für die Registrierung herangezogen werden.
Im 3D-Fall finden punktbasierte als auch auf geometrischen Primitiven basierende Methoden Anwendung. Der verbreitetste punkbasierte Algorithmus ist hier der Iterative Closest Point (ICP) (BESL & MCKAY 1992), welcher mehrfach adaptiert wurde. Hier werden Punkt-zu-Punkt-Distanzen in überlappenden Gebieten minimiert. Hierfür ist eine gute Vororientierung der Datensätze notwendig, damit der Algorithmus die richtigen 3D-Korrespondenzen für die

Minimierung des Abstandes findet. Alternativ gibt sich auch Algorithmen, die gezielt nach markanten Punkten suchen und die diese über einen Deskriptor beschreiben. Somit wird die Anzahl der möglichen Korrespondenzen auf diejenigen mit ähnlichem Deskriptor drastisch reduziert. Zur Beschreibung von markanten Punkten finden sich beispielsweise SIFT-basierte (WEINMANN et al. 2011) Ansätze als auch Schnittpunkt-basierte (THEILER & SCHINDLER 2012; GE 2017). Als geometrische Primitive finden Geraden (HABIB et al. 2005), Raumkurven (YANG & ZANG 2014) oder Ebenen (XIAO et al. 2012; XU et al. 2017) Anwendung. Diese auf geometrischen Primitiven basierenden Methoden hängen stark vom Inhalt der aufgenommenen Scene ab. Finden sich in der Szene keine urbanen Objekte wie zum Beispiel Gebäude, ist sowohl der Erfolg als auch die Genauigkeit dieser Methoden limitiert.
Sollen 2D-Daten registriert werden, ist das am meisten verwendete geometrische Primitiv ebenfalls der Punkt. Markante Punkte werden im 2D Fall über Algorithmen gefunden, die nach starken Grauwert-Variationen in unterschiedlichen Bildmaßstäben suchen. Als Beschreibung eines markanten Punktes finden beispielsweise der etablierte SIFT (LOWE 2004), Surf (BAY et al. 2006) oder mit der Verwendung eines nicht linearen Skalenraumes: KAZE (ALCANTARILLA et al. 2012). Diese sogenannten Deskriptoren benötigen aber eine annähernd gleiche radiometrische Beschreibung der lokalen Umgebung eines markanten Punktes, da sich sonst die Deskriptoren zu stark unterscheiden und die Punkte einander nicht mehr zugeordnet werden können. Dies ist auch der Fall, wenn die Bilder starke Rotations- und Maßstabsunterschiede aufweisen, wie im besonderen Fall der Registrierung von Aufnahmen einer UAV-Kamera und Luftbildern (KOCH et al. 2016). Hier wird auf andere geometrische Primitiven wie Konturen oder Geraden zurückgegriffen. Konturen benötigen allerdings vorab eine aufwendige Segmentierung und Vektorisierung. Geraden sind einfach zu segmentieren und lassen sich über die mathematische Repräsentation mit wenigen Parametern eindeutig beschreiben. Aus diesem Grund fokussiert sich dieser Beitrag auf die Verwendung von Geraden welche in einer Registrierung von Smartphone Aufnahmen zu einer Referenz Punktwolke bereits verwendet wurden (BOERNER & KRÖHNERT 2016).

3 Methodik

Die Vorgestellte Methode gliedert sich in die folgenden 3 Teile: erstens: die Extraktion von Geraden und die Segmentierung von Linien entlang dieser Geraden; zweitens: die Definition der Zuordnungsstruktur und die Zuordnung dieser Strukturen in einem RANSAC basierten Ansatz; drittens: die Bestimmung der Affinparameter für die Transformation der Bilder in ein gemeinsames Koordinatensystem.

3.1 Extraktion von Geraden und Segmentierung

Die Geraden in einem Bild werden mit Hilfe des Sobel Filters und einer Hough Transformation extrahiert. Die Hough Transformation verwendet die Geradengleichung:

$$d = r \cdot \sin\alpha + c \cdot \cos\alpha \tag{1}$$

mit:

d = Distanz vom Ursprung des Bildes

α = Winkel der Gerade zur Horizontalen Achse des Bildes (gleich der Kantenrichtung resultierend aus dem Sobel-Operators)

Der Hough-Raum ist auf einen Prozentsatz zur Anzahl der Pixel im Bild normiert. Ausgehend von einem Prozentsatz als Schwellwert werden dann die Hauptgeraden extrahiert. Die so extrahierten Hauptgeraden werden für eine Liniensegmentierung verwendet und auf Grundlage der entstandenen Segmente weiter gefiltert.
Bei der Segmentierung (schematisch in Abbildung 1 gezeigt) wird die Hough-Gerade traversiert und alle Pixel mit der gleichen Gradientenrichtung wie die Hough-Gerade, die einander benachbart sind, zu einem Segment zusammengefasst. Hat das nächste Pixel entlang der Geraden eine andere Richtung, zählt das aktuelle Segment als abgeschlossen und ein neues Segment entsteht am nächsten Pixel. Anschließend werden Geraden eliminiert die kein Segment mit minimaler Größe besitzen. So erfolgt die Filterung der Hauptgeraden nicht nur auf Grundlage eines Schwellenwertes im Hough-Raum, sondern zusätzlich auch auf Grundlage der Konnektivität der Pixel die die Gerade bilden.

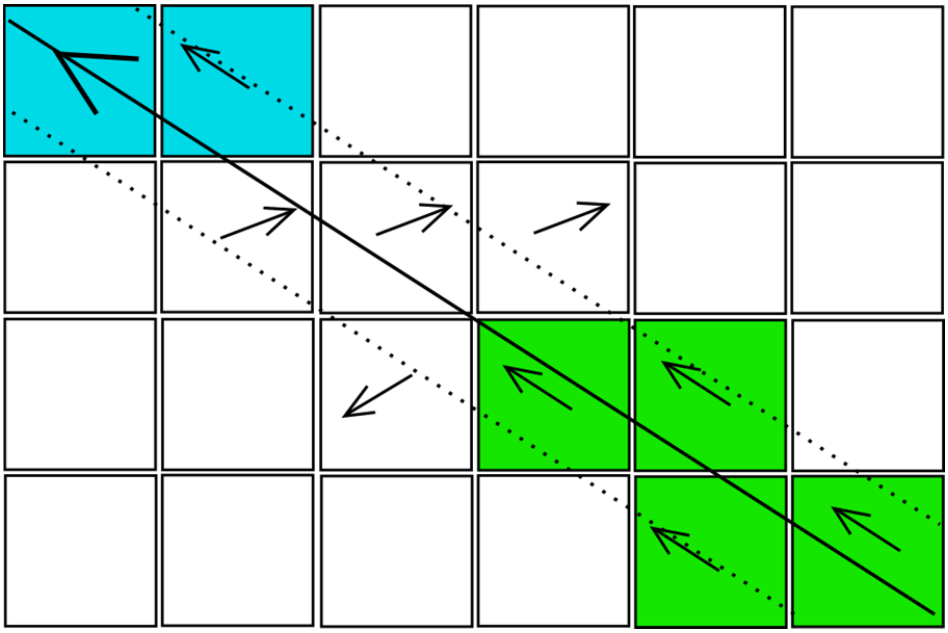


Abb. 1: Schema der Segmentierung. Die Pfeile verdeutlichen die Gradientenrichtung der einzelnen Pixel und die Farben symbolisieren die unterschiedlichen gefundenen Segmente.

3.2 Ransac Zuordnung

Nachdem die Geraden gefunden und Segmente gebildet sind, werden diese für eine Zuordnung von Korrespondenzen verwendet. Hierbei erfolgt zunächst die Zuordnung der Geraden in einem RANSAC-basierten Ansatz. Alle Geradenpaare, die in ein Initialmodel passen, zählen zunächst als zugeordnete Paare. Aus dieser Gesamtheit aller möglichen Paare werden nun zufällig 4 Paare ausgesucht. Diese 4 Paare müssen zwei Bedingungen erfüllen: es müssen paarweise parallele Geraden in einem Bild sein und der Winkel zwischen den nicht parallelen Geraden muss größer als 45 Grad sein. Die Erfüllung dieser Bedingungen garantiert, dass auf Grundlage der gefundenen Korrespondenzen eine Affintransformation bestimmbar ist. Ein optimaler

Schnittwinkel wären dabei 90 Grad, aber um den Suchraum nicht so stark einzugrenzen, dass in manchen Fällen keine Korrespondenzen zu finden sind, ist die Grenze für den Schnittwinkel herabgesetzt. Die parallelen Geraden der Struktur liefern einen Abstandswert der Linien zueinander. Da es sich um paarweise parallele Geraden handelt, können entsprechend zwei Abstände in einem Bild berechnet werden (Abbildung 2). Das Verhältnis eines Abstandes im Bild *B* zum Abstand der korrespondierenden Geraden im Bild *B‘* liefert eine Schätzung des Maßstabes der Transformation. Folglich muss für eine valide Transformation das Verhältnis der anderen Abstände gleich sein, andernfalls werden die Korrespondenzen erneut zufällig gezogen. Korrespondierende Geraden, die die Randbedingungen der Struktur erfüllen, werden zur Bestimmung der Affintransformation verwendet (siehe 3.3). Die Parameter der Affintransformation stellen dann das Model der aktuellen RANSAC-Iteration dar.

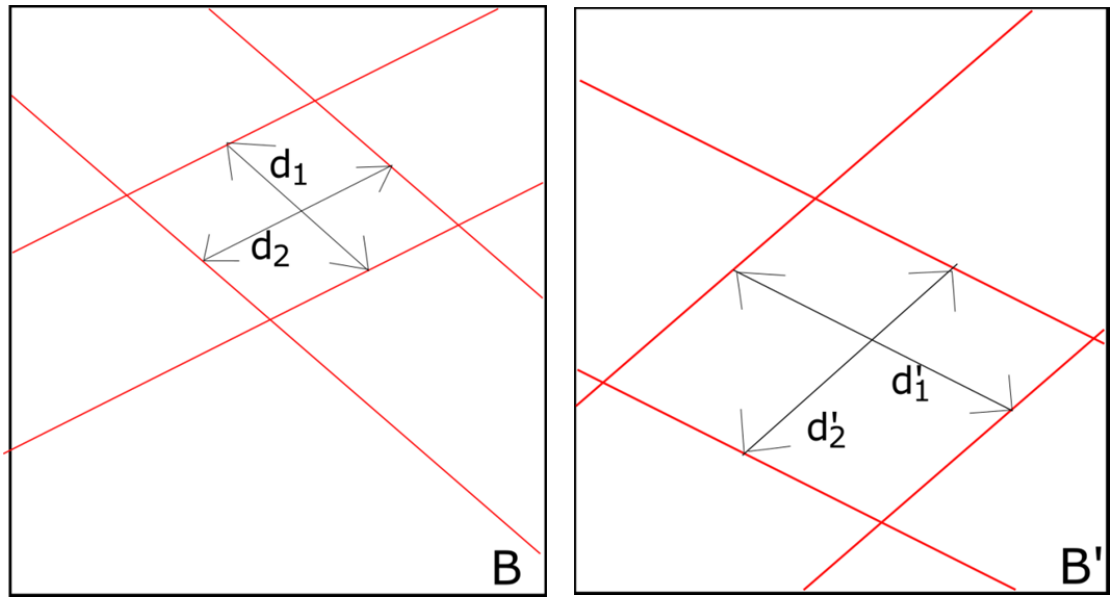


Abb. 2: Prinzipskizze der Struktur-Bedingung. Die roten Linien symbolisieren die Hough-Geraden im jeweiligen Bild B bzw. B‘. Die Distanz zwischen den parallelen Geraden wird mit d dargestellt.

Zur Evaluierung des aktuellen Modells erfolgt eine Zählung der Geraden, die zu diesem Modell passen (Inlier). Dafür erfolgt auf Grundlage der Affinparameter eine Projektion der Geraden vom Bild *B‘* in das Bild *B.* Ist der Abstand zwischen der Geraden in *B* und der korrespondierenden transformierten Geraden aus *B‘* kleiner als ein Grenzwert, zählt die Gerade zunächst als zum Modell passend. Um die Gerade vollständig als zum Modell passend zu deklarieren, erfolgt im zweiten Schritt eine Untersuchung der Segmente der projizierten und der tatsächlichen Geraden in *B.* Liegt ein Abstand der projizierten Segmente zu allen Segmenten der Gerade im Bild unterhalb einer Schwelle, wird die Gerade vollständig als zum Modell passend definiert, andernfalls wird diese wieder aus der Anzahl der zum Modell passenden Geraden herausgenommen. Das Modell mit den meisten passenden Geraden gilt als das beste Modell am Ende aller RANSAC-Iterationen.

3.3 Bestimmung der Parameter

Die Bestimmung der Affintransformation auf Grundlage von korrespondierenden Geraden wird wie in Hartey and Zisserman (2004) beschrieben vorgenommen. Ausgangspunkt ist die homologe Geradenpräsentation.

$$l = \begin{pmatrix} x \\ y \\ w \end{pmatrix} \tag{2}$$

mit

$$l \cdot \begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = x \cdot c + y \cdot r + w = 0 \tag{3}$$

Diese Repräsentation lässt sich aus der Hough-Geradengleichung (Gleichung 1) extrahieren. Mittels Vergleich der Koeffizienten ergibt sich:

$$l = \begin{pmatrix} \cos\alpha \\ \sin\alpha \\ -d \end{pmatrix} \tag{4}$$

Die Affintransformation zwischen den Geraden l und l' wird mit der Homographie

$$l = H^T l' \tag{5}$$

definiert mit:

$$H = \begin{pmatrix} \boldsymbol{h}_1^T \\ \boldsymbol{h}_2^T \\ \boldsymbol{h}_3^T \end{pmatrix} = \begin{pmatrix} h1 & h2 & h3 \\ h4 & h5 & h6 \\ h7 & h8 & h9 \end{pmatrix} \tag{6}$$

Die homologen Beobachtungsgleichungen ergeben sich zu:

$$\begin{pmatrix} \mathbf{0}^T & -w_i \boldsymbol{x}_i'^T & y_i \boldsymbol{x}_i'^T \\ w_i \boldsymbol{x}_i'^T & \mathbf{0}^T & -x_i \boldsymbol{x}_i'^T \end{pmatrix} \begin{pmatrix} \boldsymbol{h}_1 \\ \boldsymbol{h}_2 \\ \boldsymbol{h}_3 \end{pmatrix} = A_i \boldsymbol{h} = \mathbf{0} \tag{7}$$

wobei der Index i eine Gerade bestimmt. Die Funktionalmatrix A besitzt damit doppelt so viele Zeilen wie Beobachtungen (Korrespondierende Geraden). Die Lösung für **h** ist der Eigenvektor zum kleinsten Eigenwert der Eigenwertzerlegung von A^TA.

4 Daten

Als Testdatensatz dient ein Waldgebiet des Mangfall-Gebietes südöstlich von München. Hier liegen Daten einer Befliegung der Mangfall im April 2017 vor, wobei eine Haselblad-Kamera zur Bildaufnahme und ein Riegl VQ880G Scanning System zur Aufnahme der ALS-Punktwolke zum Einsatz kamen. Des Weiteren erfolgte eine UAV-Befliegung des genannten Waldgebietes im Mai 2016, wobei eine Nikon Nex-7 zur Bildaufnahme verwendet wurde. Das Gebiet zeigt einen Seitenarm der Mangfall in einem Waldstück (Abbildung 3). Somit ist von einem nicht urbanen Gebiet auszugehen. Hier finden sich keine von Menschenhand geschaffenen Elemente die für eine Registrierung der Aufnahmen herangezogen werden könnten. Die Bodenpixelgröße

beträgt in den Bilddaten ca. 8 cm bei der Flugzeug-Befliegung und bei der UAV-Befliegung ca. 1 cm. Im Hinblick auf eine Registrierung der Daten sind zwei Fehlerquellen ersichtlich: Erstens kann sich im Fall einer nicht zeitgleichen Aufnahme der Flusslauf ändern und zweitens hängen die Wasserpixel durch unterschiedliche Refraktion beim Luft-Wasser Übergang stark von der Aufnahmerichtung und anderen Faktoren ab. Die mögliche Änderung des Flusslaufes wird im Rahmen dieses Beitrages ignoriert, da im Untersuchungszeitraum keine starke Änderung stattgefunden hat. Die Korrektur der Refraktion wird ebenfalls ignoriert, da sich das Verfahren auf detektierte Kanten beschränkt und diese hauptsächlich an Land und am Übergang vom Boden zum Wasser erwartet werden. Diese Annahmen können im Einzelfall zu fehlerhaften Geradenzuordnungen führen, die jedoch durch die RANSAC-Schätzung als Ausreißer eliminiert werden.

Abb. 3: Die Testdaten: links das Luftbild aus der Luftbefliegung, rechts eine Aufnahme aus der UAV-Befliegung des gleichen Gebietes.

5 Ergebnisse

Die vorgestellte Methode liefert die Möglichkeit mit Hilfe eines Hough-Raumes Geraden zu generalisieren. Für die Segmentierung der Geraden wurde eine Distanz von bis zu 4 Pixel zu der Hough-Gerade zugelassen. Dieser Distanzschwellwert wurde kleiner gewählt, als die Distanz der Hough-Geraden zueinander (Abhängig vom Generalisierungsgrad), damit die Liniensegmente nicht zu stark in die Breite wachsen. Der Schwellwert zur Bestimmung von markanten Kanten im Hough-Raum ist auf einen Wert von $2 \cdot 10^{-2}$ % gesetzt, womit die Filterung nach Hauptgeraden hauptsächlich durch die Segmentgröße erfolgt. Abbildung 4 zeigt die nach dem Segmentfilter verbliebenen Geraden. Es ist zu sehen, dass die Segmente eine geometrische Abgrenzung der sonst kontinuierlichen Geraden in einem wiederkehrenden Muster ermöglichen.

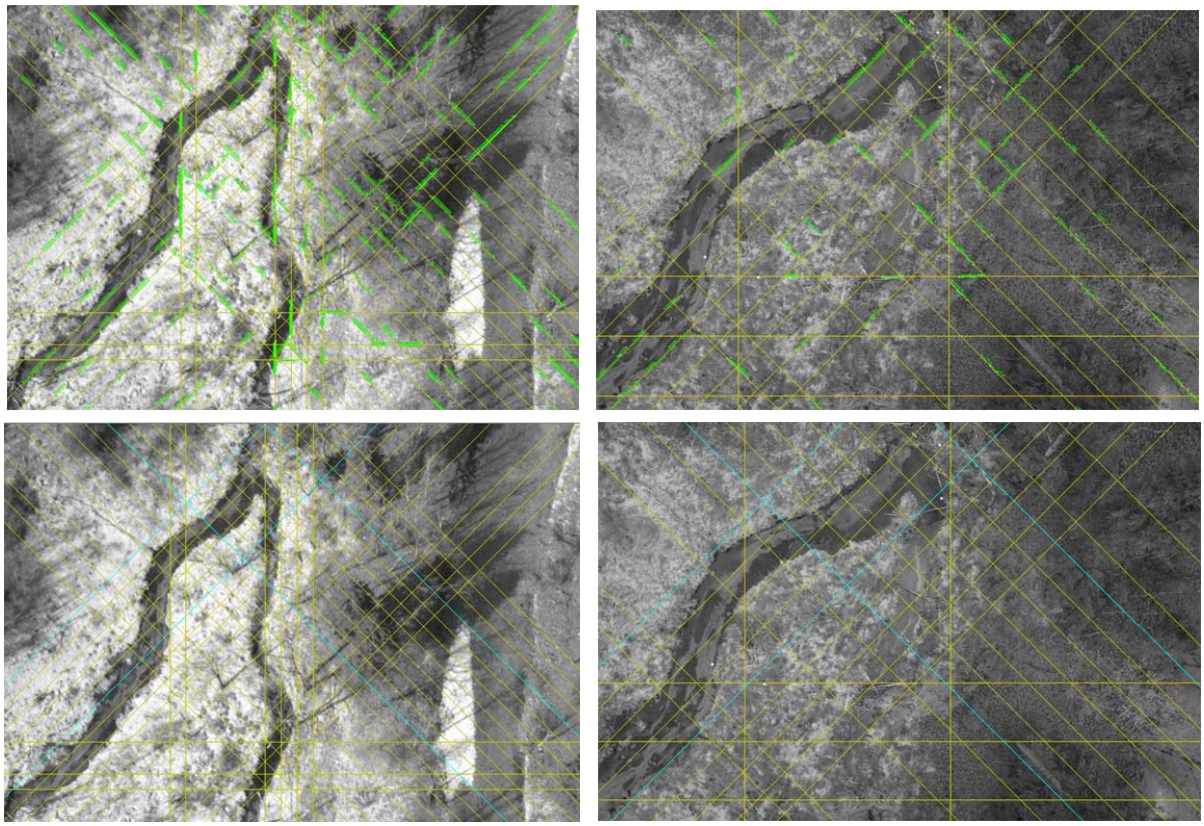


Abb. 4: Ergebnisse der Segmentierung (oben) und der Geradenzuordnung (unten):in Gelb die Hough-Geraden, in Grün: gefundene Segmente, in Blau: gefundene Korrespondierende Strukturen.

Beim Testen der Registrierung hat sich gezeigt, dass eine feinere Diskretisierung des Hough- (höhere Auflösung in der Distanz als auch in der Richtung) dazu führen kann, dass es vermehrt zu wiederkehrenden Strukturen kommt, die eine falsche Modellannahme des RANSAC-Verfahren bewirken. Mit einer Einheitsmatrix als initialer Homographie und einer zulässigen Distanz der zugeordneten Geraden von 15 m (aus der Pixeldistanz und der Bodenpixel Größe berechnet) konnte das dargestellte Ergebnis der Koregistrierung (Abbildung 4) erzeugt werden. Mit einem groben Hough-Raum mit einer Richtungsauflösung von 45 Grad und einer Distanzauflösung von 1/25 der Bilddiagonale wurde eine Genauigkeit von umgerechnet 3 m der Bild-zu-Bild-Zuordnung erzielt.

Beim Betrachten der Hough-Geraden lassen sich folgende Vor- und Nachteile des Verfahrens erkennen: Zwar muss der Suchraum mit Hilfe der Hough-Generalisierung stark eingeschränkt werden, wenn kein hinreichend genaues Modell vorliegt, aber die Verwendung von Liniensegmenten und die Bedingung der geometrischen Struktur ermöglichen auch, dass Korrespondenzen in hochfrequenten Musterns zu finden sind. Des Weiteren sind die Hough-Geraden nicht auf tatsächlich vorhandene Linienelemente beschränkt, da eine Generalisierung stattfindet. So lässt sich das Verfahren auch in nicht urbanen Gebieten verwenden. Da geometrische Primitive verwendet werden besteht keine Beschränkung der Zuordnung durch

radiometrische Unterschiede in den Aufnahmen. Somit lassen sich unterschiedlich stark belichtete Szenen als auch Bilddaten verschiedener Spektralbereiche koregistrieren.

6 Diskussion

In diesem Beitrag wurde ein Verfahren zur Registrierung von 2D-Bilddaten verschiedener Sensoren und Aufnahmezeitpunkte mit Hilfe von extrahierten Geradenstücken vorgestellt. Diese Geradenstücke werden über einen Hough-Raum generalisiert und nachfolgend segmentiert. Die so entstehenden Liniensegmente tragen dazu bei, auch Korrespondenzen in wiederkehrenden Strukturen von Hough-Geraden zu finden. Durch die Verwendung von geometrischen Primitiven kann das Verfahren auch zur Registrierung von Bilddaten mit stark unterschiedlichen radiometrischen Eigenschaften verwendet werden, wie beispielsweise multitemporale Daten und Luftbilder mit unterschiedlichen Blickrichtungen und Maßstäben. Beides ist im vorgestellten Datensatz der Fall. Das Ergebnis des Verfahrens ist eine grobe Vororientierung der Bilddaten, welche für nachfolgende Feinorientierungen mit eingeschränktem Suchraum verwendet werden können. So ist das gefundene grobe Model verwendbar, um markante Punkt rein durch deren geometrische Verteilung im Bild zu verknüpfen. Außerdem kann die Generealisierungsstufe der Hough-Transformation iterativ angepasst werden, um ein grobes Modell der vorherigen Stufe zu verfeinern und so eine Feinregistrierung auf Kantenbasis zu erzielen.
Zukünftige Arbeiten werden das Ziel verfolgen Multisensordaten zu fusionieren und die Orientierung der UAV-Daten auf Grundlage der Luftbefliegungsdaten zu verbessern. Hierfür soll die ALS-Punktwolke der Befliegung in die Luftbilder rückprojiziert und von Luftbildern über die Affintransformation in die korrespondierenden Bilder der UAV-Kamera übertragen werden. Somit können zu jedem Pixel des vom UAV aufgenommenen Bildes 3D-Koordinaten interpoliert werden, welche für eine Lokalisierung der Kamera verwendet werden können. Des Weiteren soll untersucht werden, wie die ALS-Punktwolke um einer höher aufgelöste photogrammetrische Punktwolke aus den Bildern der UAV-Befliegung lokal verdichtet werden kann. Außerdem sollen über die kürzere Wiederholrate der UAV-Befliegungen 3D-Punktwolken zur Änderungsdetektion verwendet werden.

7 Danksagung

Diese Arbeit ist gefördert von der Bayerischen Forschungsstiftung, Projekt “Schritthaltende 3D-Rekonstruktion und -Analyse (AZ-1184-15)”, Teilprojekt „Änderungsdetektion in Punktwolken“. Die Luftbefliegung wurde von der Firma „AHM- Airborne Hydro Mapping“ durchgeführt.

8 Literatur

ALCANTARILLA, P. F., BARTOLI, A., & DAVISON, A. J., 2012: KAZE features. European Conference on Computer Vision, 214-227.

BAY, H., TUYTELAARS, T. & VAN GOOL, L., 2006: Surf: Speeded up robust features. Computer vision–ECCV 2006, 404-417.

BESL, P.J. & MCKAY, N.D., 1992: Method for registration of 3-d shapes. Robotics-DL tentative, International Society for Optics and Photonics, 586-606.

BOERNER, R. & KRÖHNERT, M., 2016: Brute force matching between camera shots and synthetic images from point clouds. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 771-777.

GE, X., 2017: Automatic markerless registration of point clouds with semantic-keypoint-based 4-points congruent sets. ISPRS Journal of Photogrammetry and Remote Sensing, **130**, 344-357.

HABIB, A., GHANMA, M., MORGAN, M. & AL-RUZOUQ, R., 2005: Photogrammetric and lidar data registration using linear features. Photogrammetric Engineering & Remote Sensing, **71(6)**, 699-707.

HARTLEY, R. & ZISSERMAN, A., 2003: Multiple View geometry in computer vision. Cambridge University Press.

KOCH, T., ZHUO, X., REINARTZ, P. & FRAUNDORFER, F., 2016: A new paradigm for matching uav-and aerial images. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **3**, 83-90.

LOWE, D.G., 2004: Distinctive image features from scale-invariant keypoints. International journal of computer vision, **60(2)**, 91-110.

STAMOS, I. & LEORDEANU, M., 2003: Automated feature-based range registration of urban scenes of large scale. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, **2**, II–Ii.

THEILER, P. & SCHINDLER, K., 2012: Automatic registration of terrestrial laser scanner point clouds using natural planar surfaces. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **3**, 173-178.

WEINMANN, M., WEINMANN, M., HINZ, S. & JUTZI, B., 2011: Fast and automatic image-based registration of tls data. ISPRS Journal of Photogrammetry and Remote Sensing, **66(6)**, 62-70.

XIAO, J., ADLER, B. & ZHANG, H., 2012: 3d point cloud registration based on planar surfaces. IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), 40-45.

XIAO, J., ADLER, B., ZHANG, J. & ZHANG, H., 2013: Planar segment based three-dimensional point cloud registration in outdoor environments. Journal of Field Robotics, **30(4)**, 552-582.

XU, Y., BOERNER, R., HOEGNER, L. & STILLA, U., 2017: Automated coarse registration of point clouds in 3d urban scenes using voxel based plane constraint. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **IV-2/W4,** 185-191.

YANG, B. & ZANG, Y., 2014: Automated registration of dense terrestrial laser-scanning point clouds using curves. ISPRS Journal of Photogrammetry and Remote Sensing, **95**, 109-121.

Smart Phone Accuracy of Multi-Camera Pedestrian Tracking in Overlapping Fields of View

STEFFEN BUSCH[1]

Abstract: In this paper, we analyze the suitability of event mapping via smart phones by evaluating the accuracy of pedestrian tracking. We employ a multiple view tracking and bundle adjustment and recover the scale by an alignment to reference points measured via a total station. The use of multiple perspectives by loosely time synchronized smart phones, enables a least squares assignment of detections while improving the pose estimation of pedestrians and reducing occlusions. Finally, we used a total station to track a pedestrian to evaluate our results and show that the photogrammetric smart phone tracking accuracy of 20 cm is suitable for event mapping in the daily traffic situations.

1 INTRODUCTION

Nowadays, with the explosive development of technology for the internet of things, various sensors record lots of information in our world. This work is about the pedestrian tracking accuracy examination by exploiting wide spread smart phone sensors. By use of smart phones we will be able to comprehensively map the movement of pedestrians with multiple mobile devices, we call agents. As we demonstrated in our previous work (BUSCH et al. 2016), the movements of objects can be utilized to generate dynamic maps. These maps provide information about the real paths (trajectories) of pedestrians, when and where events occur and, similar to floating car data, traffic information.

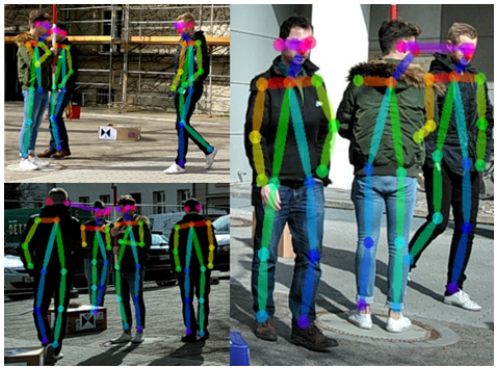


Fig. 1: Three time synchronized images of a scene, taken from three different camera viewpoints. Estimated person pose is overlaid. The bottom left and right images show a mismatch in the head association to illustrate that this detection, too, fails sometimes.

The resulting data can be used to generate information about periodic events and especially to detect anomalies such as traffic accidents or crowded scenes. Such dynamic maps can improve the navigation in cities and be a useful tool for city planning (FAYAZI et al. 2015). The more detailed mapping of the spatiotemporal traffic behavior will play a significant role in future mobility solutions. Thus, pedestrian trajectories are a valuable source for such new kinds of maps. The tracking information could be obtained with high frequency by the daily traffic of the future, when more traffic participants are equipped with appropriate sensors. As of now, a comprehensive data collection is possible by traffic control or smart phone cameras, for instance. This is achieved by following the concept of holding maps

[1] Leibniz Universität Hannover, Institute of Cartography and Geoinformatics, Appelstraße 9a, D-30167 Hannover, E-Mail: Steffen.Busch@ikg.uni-hannover.de

up-to-date by using crowd-sensing information of the daily traffic. For that reason we use smart phones as ”all-in-one” sensors and analyze their accuracy. In order to investigate the pedestrian tracking from multiple viewpoints, we used a multi-camera network consisting of smart phone cameras, to analyze the tracking accuracy. This is relevant in maps due to the necessity for reliable 3D information like trajectory planning and obstacle handling. The real world tracking results rely on the position and orientation of the cameras, which we determined by a point cloud alignment, since this method can locate independent cameras and enables a crowd sensed tracking. During the last decade, we have witnessed an explosion in Computer Vision techniques for pedestrian detection and tracking (LEAL-TAIXÉ et al. 2016; LINDER et al. 2016; ZHANG et al. 2016). This task is critical for surveillance, autonomous driving and robotics applications. Even though pedestrian detection and tracking have attracted a lot of researchers, it still poses many challenges due to the alteration of pedestrians' pose, scale, or changes in the scene illumination. Additionally, in a crowded scene, pedestrians are usually occluded by others, which is a problem that cannot be solved effectively by single camera views. Combining images from multiple viewpoints can provide a more comprehensive knowledge of a scene, in which missing information in a particular view can be supplemented by the others (MITTAL & DAVIS 2003). This method not only helps to overcome the limitations of monocular camera views, but also improves the robustness of the 3D position estimation during the pedestrian tracking procedure. However, it is difficult to fuse information from different camera views because of uncertainties in detection and pedestrian correspondence ambiguities. To tackle this problem, we propose a framework to track pedestrians in 3D on a ground plane using images from multiple cameras. Our method extends the recursive dynamic Bayes network multi-person tracker, which was introduced by KLINGER et al. (2015) to a scenario of three cameras. Besides that, we propose a method for finding the correct combination of detections from different views. For this purpose, we combine detections of different perspectives by spatial intersection and a Gauss-Markov method. Then we calculate the probability of the resulting combinations and filter outliers. Afterwards, we select the most likely combinations and utilize the prediction information of the Kalman Filter for the association process. Thus, we strive to enhance the reliability and precision of trajectories. We test our approach on a database produced by ourselves, in which areas of interest were captured by three smart phone cameras placed at three different viewpoints. One pedestrian tracking trajectory is compared against the ground truth trajectory determined using a total station (tachymeter). Figure 2 illustrates the alignment problem of multiple detections from different perspectives. The derivation from the detectors leads to several combination possibilities by spatial intersection between the viewpoints. Our approach is looking for the correct combination of detections by filtering the set of all possible combinations.

The rest of the paper is structured as follows: in Section 2 we discuss previous relevant research. Section 3 describes our method in detail. Section 4 illustrates our experiment with the captured scene. After an evaluation of our approach in Section 5 we sum up with conclusions in Section 6.

2 Related Work

This section starts with an overview of (event) mapping approaches, since the comprehensive mapping of events has inspired our work. Afterwards, different detection approaches are presented due to their high influence on the tracking approaches. Finally, various tracking approaches are discussed.

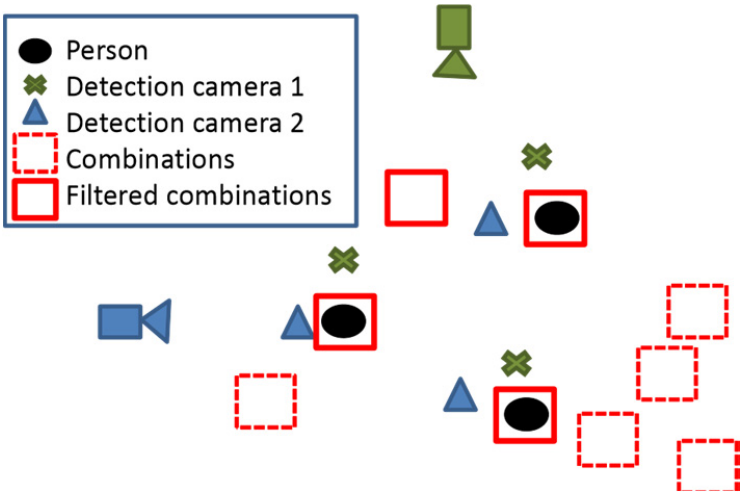


Fig. 2: Filtering of all possible combinations generated by the least square assignment of spatial intersection of different angles.

Today floating car data (FCD) is used for more effective path planning, avoiding congestions and optimizing transportation systems. Various works (RAZA & ZHONG 2017; TREIBER & KESTING 2013) predict traffic flow information based on trajectory analysis to make traffic faster, safer and more environmentally friendly. These approaches analyze trajectories generated by camera observations or by ordinary Global Navigation Satellite System (GNSS) data matched to a map. In addition to FCD, even maps and much more information are generated and updated by crowd-sensing data (HAKLAY & WEBER, 2008; HU et al., 2017). LANDSIEDEL & WOLLHERR (2017) calculated a root mean square error of around 1.5m between open street map, which is generated by using ordinary GNSS receivers, and a laser scanner reference map. GNSS trajectories generated by crowd sourcing are a widespread source to map much more information. For example ROETH et al. (2017), RUHHAMMER et al. (2017) and DURAN et al. (2016) automatically mapped road network graphs by trajectory analysis. ZOURLIDOU & SESTER (2015) and EFENTAKIS et al. (2017) used GPS tracks to enrich network graphs with information about turning behavior. Furthermore RUHHAMMER et al., (2017) and FAYAZI et al. (2015) generate higher detailed information about the spatial-temporal behavior at the network graph through traffic light sequences (TLS). In addition, they identify TLS as valuable information for traffic flow optimization. Furthermore, WANG et al. (2016) show that with the knowledge of TLS the travel time could be reduced by about 36%. All of the trajectory analysis approaches depend on trajectory precision. ATIA et al. (2017) showed that at least an accuracy of around 0.90 m is required for match detection to lane accurate maps. The precision of "GNSS only solutions" are not suitable for crowd sourcing because the low cost sensors' error varies between 5 m and 22 m (DURAN et al., 2016), but high cost solutions are expensive for a comprehensive and high frequency use. (KNOOP et al., 2017) used GPS Precise Point Positioning to adjust low cost measurements and reached a precision of 1.2 m with a reliability of 95%. Another more accurate low cost approach is presented by VIVACQUA et al. (2017), their vision based self-localization reached a precision of 0.06 m average and 0.54 m maximal error. We use these quality measures to evaluate smart phones as a valuable

crowd sensing tracking sensor. Next, we discuss the previous studies concerning pedestrian detection and tracking.

Pedestrian detection and tracking have been studied intensively during the last 20 years and their performances have been boosted continuously. Currently, there are three principal families of methods for detecting and localizing pedestrians in images. The first group of algorithms extracts hand-crafted features such as histogram oriented gradient (HOG), integral channel feature (ICF) or deformable part model (DMP) (DOLLÁR et al., 2009) of relevant objects, followed by a supervised classifier like SVM (DALAL & TRIGGS, 2005; VIOLA & JONES, 2001; FREUND & SCHAPIRE, 1997). The other method trend utilizes the Convolutional Neural Network (CNN) which is able to automatically obtain the optimum and high level features from raw pixels by training a deep network with great amounts of positive and negative training samples (CAO et al., 2016; CORDTS et al., 2016; DENG et al., 2009; LIN et al., 2014; REDMON & FARHADI, 2016). These two classes of methods deliver comparable results on different benchmark databases, such as KITTI Vision Benchmark Suite (GEIGER et al., 2012), ETH (ESS et al., 2008), MOTChallenge (LEAL-TAIXÉ et al., 2015). Another class of detector approaches for static cameras uses background modelling (ST-CHARLES et al., 2015) to detect a person by detecting differences between dynamic and static background layers and a current frame.

Using detections in each image frame, the tracking can be accomplished by associating pedestrians between two consecutive image frames. Most of the state-of-the-art methods employ the tracking-by-detection pipeline to track pedestrians, both in 2D and 3D spaces, and achieved impressive results (CHOI et al., 2013; KLINGER et al., 2017; MILAN et al., 2013). KLINGER et al., (2017) utilized the HOG-SVM to detect pedestrians in an image sequence. Next, the pedestrian association was performed using linear programming. They employed the Kalman filter and a dynamic Bayesian network to predict and correct the pedestrian trajectories as well. Moreover, they did not consider each pedestrian movement independently, but took into account interactions and affectations of other movements in the scene as well.

Our tracking method for multiple viewpoints is an extension of KLINGER'S (2017) approach. However, we take it a step further by replacing the HOG-SVM with deep learning detectors to obtain improved detection accuracy. We also include a multi view alignment module to determine the corresponding pedestrian across all viewpoints. Despite many techniques have been developed to solve the pedestrian tracking challenges, tracking from a single camera viewpoint usually fails in case of occlusion and is less accurate due to lacking information concerning the whole scene. Several approaches have been studied to leverage multi-view object tracking. To track pedestrians in crowded scenes, KHAN & SHAH (2006) performed a planar homography constraint to clarify the occlusion and localize positions of pedestrians based on the measured foot blobs. Both BERCLAZ et al. (2011) and LEAL-TAIXE et al. (2012) treated the multi-view tracking problem as a flow optimization task. Utilizing the relationship of knowledge obtained from multiple view point images by joining spare learning, HONG et al. (2013) could extract more sufficient information to boost the performance of his particle-filter-based tracker. XU et al. (2016) adopted a hierarchical composition model to generate the object trajectories through maximizing a posterior. While the existing approaches achieved remarkable improvements employing different approaches, they are more or less imposed on appearance resemblance, movement smoothness, sparsity, 3D position coincidence, etc., of pedestrians over temporal and spatial domains. In

the same manner, we take those cues into the tracker. Our tracking approach differs from high precision multi-camera tracking approaches such as used for motion capturing systems for film studios since we do not apply our approach in a closed room with defined background for Chroma keying with high precision time synchronized cameras, a full bundle adjustment over all cameras with unlimited observations and controlled light conditions.
We focus on the multi view alignment task to reduce the effect of occlusion, in which the corresponding pedestrians are determined with assumption of 3D position coincident over multiple viewpoints. Moreover, we perform our tracking in 3D world space and compare our results to a ground truth trajectory to get a quantitative measurement of accuracy we can achieve in realistic applications.

3 Multi-View Tracking

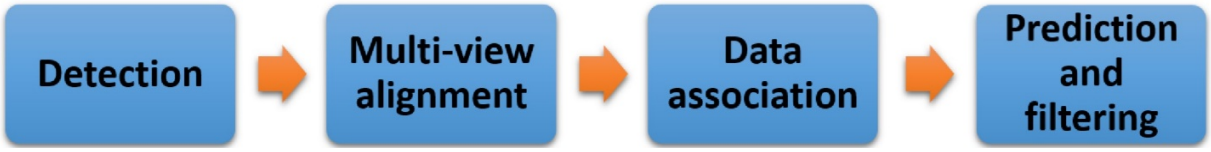


Fig. 3: The framework of our proposed tracker

We extend the conventional tracking approach from KLINGER et al. (2017) by use of different perspectives from a multi-camera network. The different viewpoints enable a more precise 3D position estimation in comparison to a single mono camera. Moreover, they make the tracking much more robust against occlusion. The framework of our approach includes four primary phases (see Figure 3): first, the detection, in which pedestrians are recognized and localized separately in each image. Secondly, the alignment task is executed to find the corresponding pedestrians between different views. Next, the detections are concatenated across frames in the data association phase. Finally, the trajectories of pedestrians are smoothed by applying a Kalman filter for the position of detections and predicted locations inferred from previous image frames. This section will first explain the localization methods since the tracking is based on the known camera positions. Afterwards, each step is explained in more detail in the corresponding chapter.

3.1 Pose Estimation via Point Cloud Alignment

To determine the camera positions, we use a point cloud alignment because this method allows for an independent localization for each camera. The poses only depend on a point cloud map and a short image sequence of a moving camera. For the alignment we calculate a point cloud via bundle adjustment and align it to the reference points from the total station. For this we pick corresponding points manually for a Four Point Congruent Sets (4PCS) (AIGER et al. 2008) registration algorithm. Thus, a transformation including the scale between the coordinate frame of the bundle adjustment and the coordinate frame of the total station is calculated. With this pose information we identify the ground plane. This ground plane is used to project every detection onto the ground and approximate the real world position for each detection.

3.2 Pedestrian Detection

For the detection, we adapt two different deep learning based methods which were recently published and can efficiently perform the pedestrian detection task. On one side, YOLO: Real-Time Object Detection (REDMON & FARHADI 2016) produces bounding boxes of pedestrians; we interpret the center of the bottom bounding box line as a foot point for the tracking. On the other side, the 2D pose estimation approach (2DPE) (CAO et al. 2016; LIN et al. 2014) delivers the skeletons of detected pedestrians, whereby we combine the neck x-position and lowest ankle y-position in order to precisely detect the foot point of pedestrians. Particularly, only the pre-trained networks provided by the authors are used and no domain adaptation to our newly generated data set is done since the models generalized well. Three images from three different viewpoints with overlaid pose skeletons are depicted in Figure 1.

3.3 Multi-View Alignment

In our case, the problem of data association expands by the use of multiple cameras. To simplify the association process, we combine the detections of the same persons from different views before starting to process. Thus, we create all possible combinations from the detections of different views and assume that each combination describes the same person. We then optimize the detected 3D position of that person by correcting the image coordinates of the detections so that they describe the same point in 3D. Based on this spatial intersection by using the Gauss-Markov model (2), with the design matrix A (1), we identify the correct combinations, which have the smallest residuals. To find that smallest residual independent of camera count, we use the square improvements (3) and the chi square function (4). In more detail, we use the Jacobian matrix of the object (world) coordinates X, Y and Z of the collinearity equation (KRAUS 2004) and use the assumption that Z is constant to determine the design matrix

$$A = \begin{bmatrix} -\frac{f_0(N_0 r_{0,0,0} - \xi_0 r_{0,2,0})}{N_0^2} & -\frac{f_0(N_0 r_{0,0,1} - \xi_0 r_{0,2,1})}{N_0^2} & -\frac{f_0(N_0 r_{0,0,2} - \xi_0 r_{0,2,2})}{N_0^2} \\ -\frac{f_0(N_0 r_{0,1,0} - \eta_0 r_{0,2,0})}{N_0^2} & -\frac{f_0(N_0 r_{0,1,1} - \eta_0 r_{0,2,1})}{N_0^2} & -\frac{f_0(N_0 r_{0,2,1} - \eta_0 r_{0,2,2})}{N_0^2} \\ \vdots & \vdots & \vdots \\ -\frac{f_i(N_i r_{i,0,0} - \xi_i r_{i,2,0})}{N_i^2} & -\frac{f_i(\&N_i r_{i,0,1} - \xi_i r_{i,2,1})}{N_i^2} & -\frac{f_i(N_i r_{i,0,2} - \eta_i r_{i,2,2})}{N_i^2} \\ -\frac{f_i(N_i r_{i,1,0} - \eta_i r_{i,2,0})}{N_i^2} & -\frac{f_i(N_i r_{i,1,1} - \eta_i r_{i,2,1})}{N_i^2} & -\frac{f_i(N_i r_{i,2,1} - \eta_i r_{i,2,2})}{N_i^2} \\ 0 & 0 & 1 \end{bmatrix}. \quad (1)$$

Where: f_i = focal length of camera i

$$N_i = r_{i,2,0}(X - T_{i,x}) + r_{i,2,1}(Y - T_{i,y}) + r_{i,2,2}(Z - T_{i,z})$$

$$\xi i = r_{i,0,0}(X - T_{i,x}) + r_{i,0,1}(Y - T_{i,y}) + r_{i,0,2}(Z - T_{i,z})$$

$$\eta_i = r_{i,1,0}(X - T_{(i,x)}) + r_{i,1,1}(Y - T_{i,y}) + r_{i,1,2}(Z - T_{i,z})$$

$T_{i,x|y|z}$ = world coordinat of camera i

r_i = orientation of camera i

We iterate the Gauss-Markov model to optimize the 3D position x' of each combination,

$$x' = (\mathrm{A^T PA})^{-1}\mathrm{A^T Pl}\,. \tag{2}$$

Where:

$\mathrm{l} = (\mathrm{x_0\ y_0 \ldots x_i\ y_i 0})^T$

$\mathrm{x}_i, \mathrm{y_i}$ = image coordinates of detection

$$P = \begin{bmatrix} \sigma_{x_0}^2 & 0 & \ldots & 0 & 0 & 0 \\ 0 & \sigma_{y_0}^2 & \ldots & 0 & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & \ldots & \sigma_{x_i}^2 & 0 & 0 \\ 0 & 0 & \ldots & 0 & \sigma_{y_i}^2 & 0 \\ 0 & 0 & \ldots & 0 & 0 & \sigma_P \end{bmatrix}^{-1}$$

$\sigma_{\mathrm{y_i}}, \sigma_{\mathrm{x_i}}$ = standard derivation of detector i

σ_p = standard derivation for ground Z coordinate

Thus, we improve the image observation by finding the smallest residuals v = Ax−l and use the variance propagation to find the right matches (NIEMEIER 2001). Then the sum of the squared improvements is calculated as follows:

$$\Omega = \mathrm{vTPv} \tag{3}$$

Afterwards, the chi square function is used to calculate a comparable measurement independent of the count of perspectives a. Inspired by the global test, we used the cumulative distribution of the chi square function to get a probability for the detections belonging to the same object:

$$\mathrm{P_D} = 1 - \int_0^{\Omega} \chi^2_{2a-2} \tag{4}$$

Where:

a = count of different perspectives.

Finally, we filter the unlikely combinations which are smaller than a defined threshold and ensure that no detection is used twice.

3.4 Data Association

For the association from detections to trajectories we use the Mahalanobis distance (5) to weight how well a detection belongs to a track, which actual position is prediction-based on the previous walking direction:

$$\delta\mathrm{M} = \begin{pmatrix} \mathrm{x_D} - \mathrm{x_T} \\ \mathrm{y_D} - y_T \end{pmatrix}^{\mathrm{T}} \cdot (\Sigma_{\mathrm{D}} + \Sigma_{\mathrm{T}})^{-1} \cdot \begin{pmatrix} \mathrm{x_D} - \mathrm{x_T} \\ \mathrm{y_D} - \mathrm{y_T} \end{pmatrix} \tag{5}$$

Where:

x_D, y_D= coordinates of detected 2D position

x_T, y_T= coordinates of predicted 2D object position

Σ_D, Σ_T= the covariance matrices of detection and prediction

Thus, δM is used to calculate the weight $w_{D,T}$ (6) for the detections belonging to a tracked pedestrian:

$$w_{D,T} = e^{-1/2\delta M} \tag{6}$$

Afterwards, we used mixed integer linear programming from BERKELAAR et al. (2004) to assign the right detections by use of a Branch-and-bound method for solving the optimization problem. Thus, we employ the constraints that every detection and trajectory could have one assignment at most.

3.5 Prediction and Filtering (Kalman Filter)

For tracking we use a dynamic Bayes Network of KLINGER et al. (2015). From each detection we calculate the two positions foot p_f and head p_h, which both have the same x-image coordinate of the neck, while y-coordinates are estimated corresponding to the coordinate of the lower ankle and the upper eye detection. The tracking is performed by an extended Kalman filter in the real world coordinate frame by tracking X, Y, Z of the foot pF and the head pH, whereby only Z differs. In addition the hidden variables for the speed [vx, vy] along X and Y are part of the tracking state 7:

$$wi,t = [X_{i,t}, Y_{i,t}, Z_{f,i,t}, Z_{h,i,t}, v_{x,i,t}, v_{y,i,t}] \tag{7}$$

Where:

i = object index

t = time index

Z_f, Z_h = Z coordinate of food and head

The measurement matrix of the Kalman filter is similar to the design matrix equation 1 but is extended by the observed head position. For this, the Z coordinates of the heads are given by the height of the detected bounding boxes.

4 Experiment

We tracked pedestrians in a network of three smartphone cameras. The cameras were placed in a triangle with an edge length of around 20 m and with a horizontal orientation. The floor plan of our setup is depicted in Figure 4.

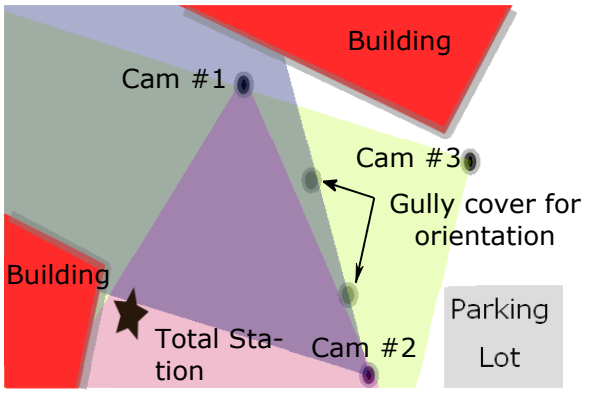


Fig. 4: Floor plan for our recorded sequences. In regions of overlapping field of views, pedestrians are visible in more than one camera.

In our experiment the agents were mounted on tripods at a height of 1.3 m above ground, a person's usual smartphone carrying height. In addition we measured a precise ground truth with 192 total station (Leicar LSM50) measurements of one pedestrian. This person carried a support rod mounted with a 360-degree prism at a height of about 3 m to prevent occlusions. For the calculation of the real world transformation, including the missing scale information, we measured 10 points with the total station and manually aligned the point cloud of the bundle adjustment. In order to focus on the accuracy of the detection and the tracking, we used only one static pose for each smartphone. The point cloud alignment could be used to localize the position of an agent based on 3D point clouds of mobile mapping systems, e.g. the 4PCS performs with a root mean square error (RMSE) of 0.027 m. Furthermore, we analyzed the influence of the transformation error in a mean pixel error of 11.3 by projecting the real world marks into the images. This relatively big mean pixel error distorted the results of less than 1 pixel standard derivation of the bundle adjustment and was added to the standard derivation of the detectors. Thus, we used a standard derivation of 40 pixels for the detected y and 30 pixels for the x image coordinate. We captured a crowded scene including 10 pedestrians for a period of two minutes. All trajectories of the tracked persons with the total station are shown in Figure 5. The pedestrians walked back and forth in the scene. We recorded many occlusions, as expertly anticipated in Figure 1. In more detail, we used a notebook to record the tracking data of the total station and time stamped the position with the processor time as soon as they arrived along the serial connection. We captured the images via action listener and time stamped this image at the arrival with the processor time of the smart phone. After we saved the image in the internal storage of the smart phone we captured the next image. Thus, the recording with full HD resolution (1920×1080 pixels) was not synchronized and the frame rate of each smart phone was not constant, but no variations were lower than 5fps. The images from different perspectives allowed for a higher tracking precision by triangulation of synchronized detections. The time synchronization was applied by the Wi-Fi internet connection of the smart phones and the notebook via the Network Time Protocol (NTP) to a global NTP sever. For the evaluation we identified a time lag of one second for the connection between total station and the notebook. Due to the irregular image capturing we used a 100 millisecond time window to synchronize the images.

5 Evaluation

The evaluation is structured in two parts, starting with the evaluation of the detection precision, including a comparison between bounding box and skeleton detectors, and the analysis of the improvements by combining multiple perspectives. Afterwards, the analysis of the tracking results is presented.

5.1 Detection Statistics

In this paragraph we discuss the detection results that we achieved with the pre-trained detectors. We evaluate every detection from our image sequence by calculating precision, recall and f-score of both YOLO and 2DPE detectors. Table 1 shows the precision, recall and f-score for the YOLO bounding box detector, while in Table 2 the results for the 2DPE are shown. The performance of the 2DPE skeleton detector on the sequences of camera 1 and camera 3 is better, while

the YOLO approach delivers the higher f-score for the sequence of camera 2. Secondly, we analyze the enhancement of spatial intersection for detections. We assume that the nearest position, calculated by projecting the detection onto the ground (GP) or least square assignment of spatial intersection (OP), is the position of the reference person, called closest position assumption in the following. The next section (Tracking 5.2) will indicate that this assumption holds true in most cases. Nevertheless, the analysis will show some outliers, which appear because of an assignment of an OP of another person to the reference pedestrian. Based on the closest position assumption we compare the accuracy of the position GP and OP by their difference to the reference.

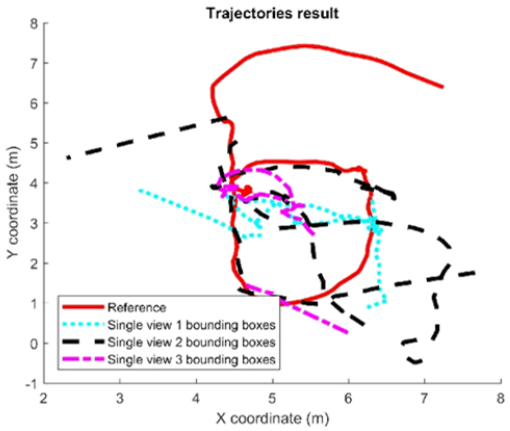


Fig. 5a: Trajectories from mono camera tracking. All trajectories belong to the pedestrian tracked by the total station

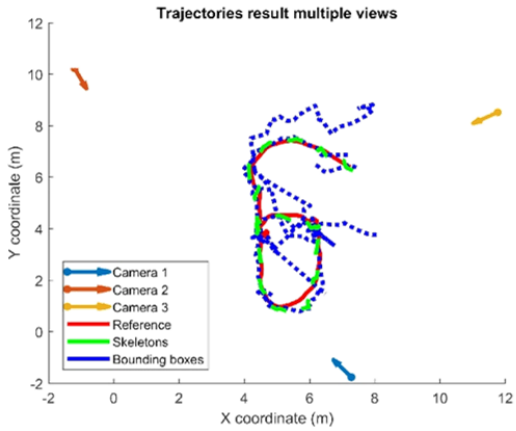


Fig. 5b: Trajectories from multi-camera tracking. All trajectories belong to the pedestrian tracked by the total station

Tab. 1: Detection results for the Yolo V2 bounding box detector

	Precision	Recall	F1-score
Camera 1	0.98	0.91	0.95
Camera 2	0.96	0.82	0.88
Camera 3	0.98	0.81	0.89

Tab. 2: Detection results for the CMU pose estimation

	Precision	Recall	F1-score
Camera 1	1	0.94	0.97
Camera 2	1	0.77	0.87
Camera 3	0.96	0.92	0.94

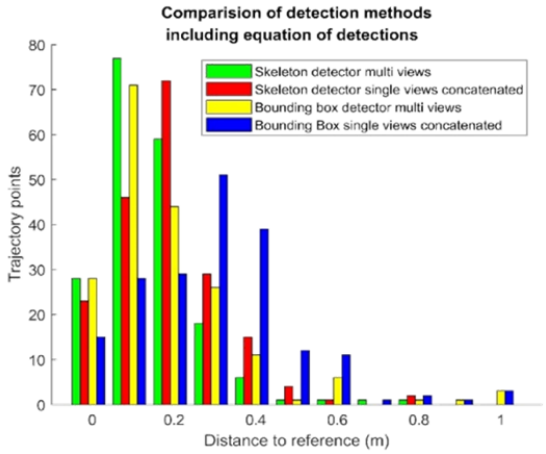


Fig. 6: Histogram of detection differences of detectors. Objective is a high number of trajectory points by a small distance to the reference

The higher accuracy of the skeleton detector is shown in Figure 6. The average precision of positions is raised from 0.31 m to 0.2 m by the spatial intersection of multiple views for the bonding box positions. This improvement is higher in contrast to the skeleton detector with an improvement of the mean accuracy from 0.2 m to 0.16 m. This difference occurs because of the higher bounding box height variation of the bounding box detector. This result shows that less accurate detections benefit more from the spatial intersection, but more variations have a fatal influence on the tracking, as will be shown in the next section (5.2 Tracking). Nevertheless, this higher accuracy shows the benefits of spatial intersection for determining positions from detections. A closer look at the positional improvement by spatial intersection is given in Figure 7 by calculating the difference between the distance to the reference of the two positions GP and OP for both detectors. A positive value means that the OP was closer to the reference and a negative value shows that the GP was closer to the reference. High negative values are explicable by comparing a spatial intersection of different persons because of the single observability of the reference pedestrian and depicted outlier. Another uncertainty of the analysis is given by the difference between the position of the camera tracked foot point and tripod point tracked by the total station. Nevertheless, Subfigure 7a shows that most of the improvements are positive with the mean of 0.101 m for the bounding box detector. Subfigure 7b shows more negative values for the skeleton detector, but the decline varies around ± 20 cm and thus is more strongly influenced by the different tracking points from cameras and the total station. Finally, Subfigure 7c shows the differences between the automatically chosen OPs and the GPs for the tracking process, which are aligned to the reference person. The fact that the outliers are missing in this case indicates that they were caused by violation of the nearest position assumption. The average change by optimization of -0.01 for the automatically chosen positions indicates enhancement potential of our filtering step. However, at least in 73 out of 186 situations an improvement at the automatically chosen positions was achieved. Besides the improvements of the estimated positions, our approach was able to summarize the detections of different views and thus performs an end-to-end tracking despite the occlusion scenarios. For the tracking, the huge set of all possible combinations was filtered to identify the right pairs of detections.

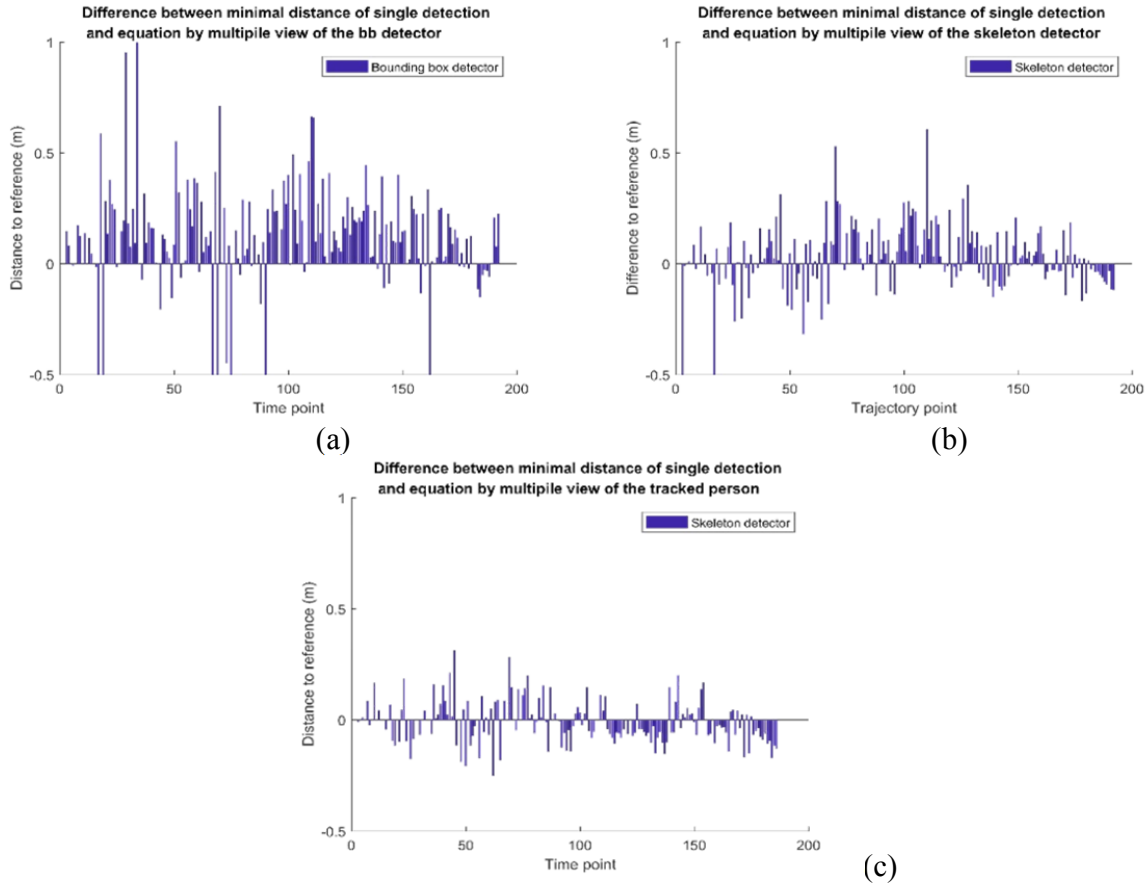


Fig. 7: Differences of detection and ground truth

Figure 8 shows the fusion of the detections by summarizing the detections of different perspectives which belong to the same person. This snapshot visualizes our way from several detections through the exponential growing set of combinations to the small set of filtered positions, which reflects the real number of pedestrians at the scene.

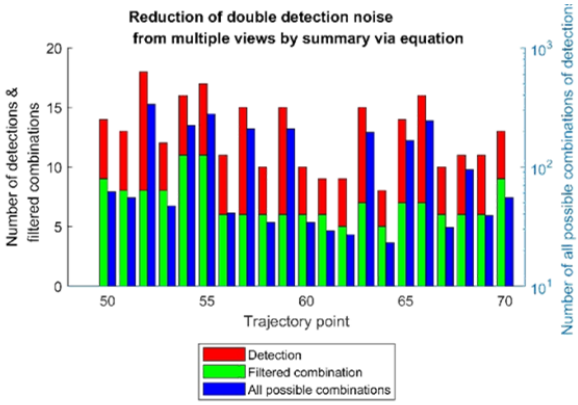


Fig. 8: Summary of detections by filtering all possible combinations of different views

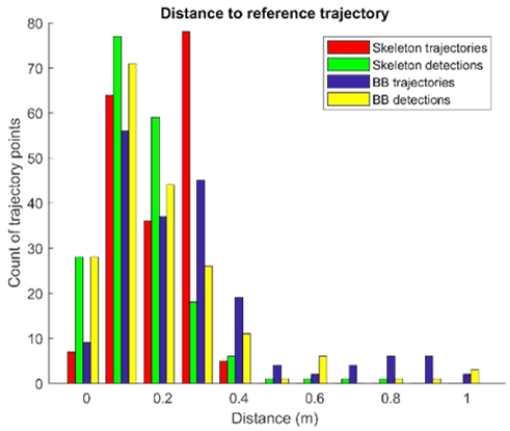


Fig. 9: Histogram of the distance of tracked trajectories and nearest detected position to the reference trajectory assigned by time stamp

5.2 Tracking

After the analysis of the detection precision we analyze the tracking accuracy. The different short trajectories in Figure 5 a show that no accurate tracking is possible because of many occlusions in case of a mono view. The influence of the bounding box size variance is shown in Figure 5b, where inaccurate detections break tracks and thus new tracks start farther from the reference. That is why no end-to-end tracking was possible with the bounding box detector. For that reason, we show every trajectory by this detector, which was once assigned to the reference pedestrian in Figure 5b as dotted blue lines. In contrast to the bounding box tracking, our approach by using the skeleton detector was able to continuously track the reference pedestrian. Even in the case of occlusion situations, where the distance of pedestrians was only 30 cm - as depicted in Figure 1- we received no identity switches. Thus, an end-to-end tracking was possible and the track is shown in Figure 5b with a dashed green line. To make the end-to-end tracking result of the skeleton detector comparable to the bounding box detector with many identity switches, we used the nearest trajectory to the reference at each time step, which was assigned to the reference person once and compared their distances. Figure 9 shows the higher accuracy and the outlier resistance of tracking with use of the skeleton detector in contrast to the use of the bounding box detector. This figure shows the amount of positions over their distance to the reference. Our final result for our tracking approach via smart phones and a skeleton detector is a RMSE of 0.23 with a variance of 0.01. Besides this accuracy the small difference between the RMSE for the positions from the skeleton tracking and detector (0.23 m and 0.2 m) showed the high assignment precision of our approach. The Kalman filtering reduces the outliers (maximum position error) and shows results that are more robust. Moreover, the same RMSE analysis for the position from the bounding box detector (0.35 m and 0.31 m) demonstrates the robustness of our summary of different perspectives, even in cases of high variation at the detections. However, the variation of the bounding box detections led to more outliers and identity switches. Since skeleton detections are more precise, the identity switches are reduced from six for the bounding box detector to zero switches for the skeleton detector.

6 Conclusion

We presented a multi-view tracking approach with the overarching goal of event mapping, which was able to track a pedestrian on our dataset without identity switches despite occlusion scenarios in a crowded scene. We were able to ensure a maximal failure of 42 cm and a mean precision of 20 cm by using simple time-synchronized smartphones, a neural network detector and the spatial intersection of different viewpoints. We showed that the positions of cameras, estimated by a bundle adjustment and a point cloud alignment can be used to track pedestrians with this accuracy. In future work we will extend this approach to mobile camera networks to build a dynamic map with the long term objective of anomaly detection in dynamic maps.

7 ACKNOWLEDGEMENTS

Steffen Busch was funded by the German Science Foundation DFG within the priority program SPP1835, "Cooperative Interacting Automobiles".

8 References

AIGER, D., MITRA, N.J. & COHEN-OR, D., 2008: 4-points Congruent Sets for Robust Surface Registration. ACM Trans. Graph., **27**(85), 1-10.

ATIA, M.M., HILAL, A.R., STELLINGS, C., HARTWELL, E., TOONSTRA, J., MINERS, W.B. & BASIR, O.A., 2017: A Low-Cost Lane-Determination System Using GNSS/IMU Fusion and HMM-Based Multistage Map Matching. IEEE Trans. Intell. Transp. Syst. **18**, 3027-3037. https://doi.org/10.1109/TITS.2017.2672541

BERCLAZ, J., FLEURET, F., TURETKEN, E. & FUA, P., 2011: Multiple object tracking using k-shortest paths optimization. IEEE Trans. Pattern Anal. Mach. Intell., **33**, 1806-1819.

BERKELAAR, M., EIKLAND, K. & NOTEBAERT, P., 2004: lpsolve: Open source (mixed-integer) linear programming system. Eindh. U Technol.

BUSCH, S., SCHINDLER, T., KLINGER, T. & BRENNER, C., 2016: Analysis of Spatio-Temporal Traffic Patterns Based on Pedestrian Trajectories. Int Arch Photogramm Remote Sens Spat. Inf Sci XLI-B2.

CAO, Z., SIMON, T., WEI, S.-E. & SHEIKH, Y., 2016: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. ArXiv Prepr. ArXiv161108050.

CHOI, W., PANTOFARU, C. & SAVARESE, S., 2013: A general framework for tracking multiple people from a moving camera. IEEE Trans. Pattern Anal. Mach. Intell., **35**, 1577-1591.

CORDTS, M., OMRAN, M., RAMOS, S., REHFELD, T., ENZWEILER, M., BENENSON, R., FRANKE, U., ROTH, S. & SCHIELE, B., 2016: The Cityscapes Dataset for Semantic Urban Scene Understanding. in: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

DALAL, N. & TRIGGS, B., 2005: Histograms of Oriented Gradients for Human Detection. in: CVPR, 886-893.

DENG, J., DONG, W., SOCHER, R., LI, L.-J., LI, K. & FEI-FEI, L., 2009: ImageNet: A Large-Scale Hierarchical Image Database. in: CVPR09.

DOLLÁR, P., TU, Z., PERONA, P. & BELONGIE, S., 2009: Integral channel features.

DURAN, D., SACRISTÁN, V. & SILVEIRA, R.I., 2016: Map Construction Algorithms: An Evaluation Through Hiking Data. in: Proceedings of the 5th ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems, MobiGIS '16. ACM, New York, NY, USA, 74-83. https://doi.org/10.1145/3004725.3004734

EFENTAKIS, A., GRIVAS, N., PFOSER, D. & VASSILIOU, Y., 2017: Crowdsourcing turning-restrictions from map-matched trajectories. Inf. Syst., **64**, 221-236. https://doi.org/https://doi.org/10.1016/j.is.2016.04.004

ESS, A., LEIBE, B., SCHINDLER, K., GOOL, & L. VAN, 2008: A Mobile Vision System for Robust Multi-Person Tracking, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR'08). IEEE Press.

FAYAZI, S.A., VAHIDI, A., MAHLER, G. & WINCKLER, A., 2015. Traffic signal phase and timing estimation from low-frequency transit bus data. IEEE Transactions on Intelligent Transportation Systems, **16**(1), 19-28.

FREUND, Y. & SCHAPIRE, R.E., 1997: A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting.

GEIGER, A., LENZ, P. & URTASUN, R., 2012: Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite, in: Conference on Computer Vision and Pattern Recognition (CVPR).

HAKLAY, M. & WEBER, P., 2008: Openstreetmap: User-generated street maps. IEEE Pervasive Comput., **7**, 12–18.

Hong, Z., Mei, X., Prokhorov, D. & Tao, D., 2013. Tracking via robust multi-task multi-view joint sparse representation, in: Proceedings of the IEEE International Conference on Computer Vision, 649-656.

HU, M., ZHONG, Z., NIU, Y. & NI, M., 2017: Duration-Variable Participant Recruitment for Urban Crowdsourcing With Indeterministic Trajectories. IEEE Trans. Veh. Technol., **66**, 10271-10282. https://doi.org/10.1109/TVT.2017.2718043

KHAN, S. & SHAH, M., 2006: A multiview approach to tracking people in crowded scenes using a planar homography constraint. Comput. Vision–ECCV 2006 133-146.

KLINGER, T., ROTTENSTEINER, F. & HEIPKE, C., 2017: Probabilistic multi-person localisation and tracking in image sequences. ISPRS J. Photogramm. Remote Sens.

KLINGER, T., ROTTENSTEINER, F. & HEIPKE, C., 2015: PROBABILISTIC MULTI-PERSON TRACKING USING DYNAMIC BAYES NETWORKS. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. II-3/W5, 435-442. https://doi.org/10.5194/isprsannals-II-3-W5-435-2015

KNOOP, V.L., DE BAKKER, P.F., TIBERIUS, C.C. & VAN AREM, B., 2017: Lane Determination With GPS Precise Point Positioning. IEEE Transactions on Intelligent Transportation Systems, **18**(9), 2503-2513.

KRAUS, K., 2004: Photogrammetrie, 3rd ed, 7. Walter de Gruyter, 10785 Berlin.

LANDSIEDEL, C. & WOLLHERR, D., 2017: Road geometry estimation for urban semantic maps using open data. Adv. Robot., **31**, 282-290. https://doi.org/10.1080/01691864.2016.1250675

LEAL-TAIXÉ, L., CANTON-FERRER, C. & SCHINDLER, K., 2016: Learning by tracking: Siamese CNN for robust target association, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 33-40.

LEAL-TAIXÉ, L., MILAN, A., REID, I., ROTH, S. & SCHINDLER, K., 2015: MOTChallenge 2015: Towards a Benchmark for Multi-Target Tracking. ArXiv150401942 Cs.

LIN, T.-Y., MAIRE, M., BELONGIE, S.J., BOURDEV, L.D., GIRSHICK, R.B., HAYS, J., PERONA, P., RAMANAN, D., DOLLÁR, P.& ZITNICK, C.L., 2014: Microsoft COCO: Common Objects in Context. CoRR abs/1405.0312.

LINDER, T., BREUERS, S., LEIBE, B. & ARRAS, K.O., 2016: On multi-modal people tracking from mobile platforms in very crowded and dynamic environments. in: IEEE International Conference on Robotics and Automation (ICRA), 5512-5519.

MILAN, A., SCHINDLER & K., ROTH, S., 2013: Detection-and trajectory-level exclusion in multiple object tracking. in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3682-3689.

MITTAL, A. & DAVIS, L.S., 2003: M2tracker: A multi-view approach to segmenting and tracking people in a cluttered scene. Int. J. Comput. Vis., **51**, 189-203.

NIEMEIER, W., 2001: Ausgleichsrechnung, 3rd ed, 1. Walter de Gruyter, 10785 berlin.

RAZA, A. & ZHONG, M., 2017: Hybrid lane-based short-term urban traffic speed forecasting: A genetic approach. in: 4th International Conference on Transportation Information and Safety (ICTIS), 271-279. https://doi.org/10.1109/ICTIS.2017.8047776

REDMON, J. & FARHADI, A., 2016: YOLO9000: Better, Faster, Stronger. ArXiv Prepr. ArXiv161208242.

ROETH, O., ZAUM, D. & BRENNER, C., 2017: Extracting Lane Geometry and Topology Information from Vehicle Fleet Trajectories in Complex Urban Scenarios Using a Reversible Jump Mcmc Method. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. **4**, 51.

RUHHAMMER, C., BAUMANN, M., PROTSCHKY, V., KLOEDEN, H., KLANNER, F. & STILLER, C., 2017: Automated Intersection Mapping From Crowd Trajectory Data. IEEE Trans. Intell. Transp. Syst., **18**, 666-677.

ST-CHARLES, P.-L., BILODEAU, G.-A. & BERGEVIN, R., 2015: Subsense: A universal change detection method with local adaptive sensitivity. IEEE Trans. Image Process. **24**, 359-373.

TREIBER, M. & KESTING, A., 2013: Trajectory and Floating-Car Data, in: Traffic Flow Dynamics: Data, Models and Simulation. Springer Berlin Heidelberg, Berlin, Heidelberg, 7-12. https://doi.org/10.1007/978-3-642-32460-4_2

VIOLA, P. & JONES, M., 2001: Robust Real-time Object Detection, in: International Journal of Computer Vision.

VIVACQUA, R.P.D., BERTOZZI, M., CERRI, P., MARTINS, F.N. & VASSALLO, R.F., 2017: Self-Localization Based on Visual Lane Marking Maps: An Accurate Low-Cost Approach for Autonomous Driving. IEEE Transactions on Intelligent Transportation Systems, 1-16. https://doi.org/10.1109/TITS.2017.2752461

WANG, S., DJAHEL, S., ZHANG, Z. & MCMANIS, J., 2016: Next Road Rerouting: A Multiagent System for Mitigating Unexpected Urban Traffic Congestion. IEEE Trans. Intell. Transp. Syst., **17**, 2888-2899. https://doi.org/10.1109/TITS.2016.2531425

XU, Y., LIU, X., LIU, Y. & ZHU, S.-C., 2016: Multi-view people tracking via hierarchical trajectory composition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4256-4265.

ZHANG, S., BENENSON, R., OMRAN, M., HOSANG, J. & SCHIELE, B., 2016: How far are we from solving pedestrian detection? in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1259-1267.

ZOURLIDOU, S. & SESTER, M., 2015. Road regulation sensing with in-vehicle sensors. in: AGILE PhD School.

Bestimmung der Korrespondenz zwischen historischen Gebäudeaufnahmen basierend auf der Zuordnung geometrischer Merkmale

FERDINAND MAIWALD[1], DANILO SCHNEIDER[1] & FRANK HENZE[1]

Zusammenfassung: Dieser Beitrag befasst sich mit der Korrespondenzbestimmung historischer Gebäudeaufnahmen aus der Fotothek der Sächsischen Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB). Im Gegensatz zu aktuellem Bildmaterial sind bei historischen Fotografien verschiedene Faktoren ausschlaggebend, die eine automatische Bildanalyse (Merkmalssuche, Merkmalszuordnung und relative Orientierung) deutlich erschweren. Aufgrund dessen werden in dem hier präsentierten Ansatz ausschließlich geometrische Merkmale verwendet, um die Verknüpfung zwischen zwei oder mehreren historischen Aufnahmen zu ermöglichen. Als robuste Merkmale haben sich vor allem Rechtecke im Objektraum herausgestellt. Diese können zum einen relativ leicht und schnell detektiert werden und zum anderen zeigen Rechtecke bei historischen Gebäuden oft eindeutige Formationen (Lage und Anordnung von Fenstern, Steinen, Verkleidungen) und können so in anderen Bildern wiedererkannt werden.

1 Einleitung

Im Rahmen des interdisziplinären Forschungsprojektes „UrbanHistory4D“ soll der Zugang zu digitalen Bildrepositorien für die Erstellung räumlich-zeitlicher Modelle (4D) sowie deren Nutzung in der stadtgeschichtlichen Forschung und Wissensvermittlung untersucht werden. Ein wichtiges Teilprojekt in diesem Forschungsvorhaben ist die automatisierte Erschließung größerer digitaler Bildbestände für eine photogrammetrische 3D-Auswertung.

Dieser Beitrag befasst sich mit der Verknüpfung und anschließender relativen Orientierung historischer Aufnahmen im Kontext eines historischen dreidimensionalen Stadtmodells der Stadt Dresden. Das Ziel ist eine Darstellung von historischen Bildern und Plänen in einem dreidimensionalen Modell. Notwendig für die Überlagerung von Bild und Modell ist die Bestimmung der Kameraposition und -orientierung. Grundlage dafür bildet das Bildarchiv der Fotothek der Sächsischen Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB), das zum jetzigen Zeitpunkt ca. 2 Millionen Bilder aus 88 Institutionen enthält. Die Aufnahmen wurden größtenteils zwischen 1940 und 1990 gemacht (deutschefotothek.de). Die Untersuchungen konzentrieren sich zunächst auf Gebäudeaufnahmen des historischen Stadtzentrums von Dresden. Im Gegensatz zu aktuellem Bildmaterial sind bei historischen Fotografien verschiedene Faktoren ausschlaggebend, die eine automatische Bildanalyse (Merkmalssuche, Merkmalszuordnung und relative Orientierung) deutlich erschweren:

[1] Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstraße 10, D-01062 Dresden, E-Mail: [ferdinand.maiwald, danilo.schneider, frank.henze]@tu-dresden.de

- Durch Digitalisierung der Aufnahmen in geringer Auflösung gehen Bildinformationen verloren.
- Über die Digitalisierung liegen keine Informationen vor. Beim Scan von Bildausschnitten kann der Bildhauptpunkt am Rand oder sogar außerhalb des digitalen Bildes liegen.
- Für die verwendeten Kameras liegen keine Kameraparameter und Kamerastandpunkte vor. Innere und äußere Orientierung der Kameras sind somit teilweise oder gänzlich unbekannt.
- Aufgrund verschiedener Ursachen (Filmkörnung, Staubpartikel, Digitalisierungsprozess…) weisen die digitalisierten Aufnahmen ein Bildrauschen (Noise) auf, das die Objekttextur (Signal) überlagern kann. In einigen Bildern entsteht ein ungünstiges Signal-Rausch-Verhältnis (signal-to-noise ratio, SNR), das eine automatisierte Merkmalssuche erschwert.
- Die radiometrischen Differenzen zwischen zwei Bildern derselben Epoche sind in den meisten Fällen sehr groß (Abbildung 1).

Abb. 1: Beispiel für radiometrische Differenzen zwischen drei historischen Aufnahmen des Georgentors in Dresden

Herkömmliche Matching-Verfahren wie Brute-Force-Matching und Flann-based-Matching (MUJA & LOWE 2009) führten mit Deskriptoren wie SIFT (LOWE 2004), SURF (BAY 2006) und ORB (RUBLEE et al. 2011) zu keinen bzw. wenigen korrekten Zuordnungen bei Anwendung auf zwei historische Aufnahmen.

In dem hier präsentierten merkmalsbasierten Ansatz werden daher ausschließlich *geometrische* Merkmale verwendet, um die Verknüpfung zwischen zwei oder mehreren historischen Aufnahmen zu ermöglichen. Als robuste Merkmale haben sich dabei zunächst rechteckige Strukturen im Objektraum herausgestellt. Diese können zum einen relativ leicht und schnell als allgemeine Vierecke im Bildraum detektiert werden und zum anderen zeigen Rechtecke bei historischen Gebäuden oft eindeutige Formationen (Lage und Anordnung von Fenstern, Steinen, Verkleidungen) und können so in anderen Bildern über semantisch-topologische Beziehungen wiedererkannt werden.

2 Stand der Forschung

Die Arbeit mit historischen Fotografien wird durch die zunehmende Erschließung und Digitalisierung von Datensätzen ein immer größer werdendes Forschungsfeld. Sowohl Forschungsplattformen wie Europeana (europeana.eu), prometheus (prometheus-bildarchiv.de) oder die deutsche Fotothek (deutschefotothek.de) als auch Seiten auf denen private Nutzer historische Bilder verwalten können wie rephotos (re.photos) oder Mapillary (mapillary.com) vergrößern jährlich ihren Bildbestand. Diese Bilddaten werden meist in einer klassischen Bilddatenbank abgelegt, in der mittels Stichwortsuche gesucht werden kann. Um diese Datenmengen besser zugänglich zu machen, kann zusätzlich eine räumliche Darstellung hilfreich sein. Für diese Art der Darstellung ist die Position des Bildes im Modell und damit eine Bestimmung der Kameraposition und Aufnahmerichtung (äußere Orientierung) notwendig. SCHINDLER & DELLAERT (2012) entwickelten mit „4D Cities“ bereits eine Plattform die mittels Structure-from-Motion große Bilddatenbestände in einem vierdimensionalen Modell darstellt und lokalisiert. Allerdings erreichte die automatisierte Variante mittels SIFT-Merkmalen deutlich weniger korrekte Zuordnungen als die manuelle Punktzuordnung in verschiedenen historischen Bildern durch einen Operateur (SCHINDLER & DELLAERT 2012).
In weiteren Publikationen wurden zusätzliche Deskriptoren und Matching-Verfahren auf historische Bilder angewendet. Eine Kombination aus SIFT/SURF und ORB erreichte die größte Anzahl an korrekten Zuordnungen, wobei der Digitalisierungsprozess stets als Grund für große Unsicherheiten in Merkmalssuche und -zuordnung angegeben wurde (WOLFE 2015), (ALI & WHITEHEAD 2015). Eine Weiterverarbeitung der homologen Punkte zur Bestimmung der relativen Orientierung von zwei Bildern ist daher kritisch zu sehen. Vor allem periodische Strukturen und ein ungünstiges Verhältnis zwischen Bildrauschen und Objekttextur sorgen bei der Arbeit mit historischen Gebäudeaufnahmen für wenige robuste Merkmalspunkte und fehlerhafte Punktzuordnungen. Weitere Fehlerquellen sind Informationsverlust, Bilddeformation und Artefaktbildung hervorgerufen durch den Digitalisierungsvorgang (ALI & WHITEHEAD 2015).
Deshalb sind für die Arbeit mit historischen Bildern intensitäts- bzw. flächenbasierte Deskriptoren weitestgehend ungeeignet (STYLIANOU et al. 2015). Geometriebasierte Matching-Verfahren wie Linien-Matching basierend auf speziellen Signaturen und Semantik (WANG et al. 2009), oder Rechteck-Matching basierend auf einzelnen Liniensegmenten (MIČUŠÍK et al. 2008), können bei historischen Aufnahmen zu besseren Ergebnissen führen. Im Folgenden soll ein auf diesen Anwendungsfall entwickelter Viereck-Deskriptor sowie die damit erreichten Zuordnungsergebnisse vorgestellt werden.

3 Arbeitsablauf

Der Arbeitsablauf wird in den folgenden Unterkapiteln in verschiedene Schritte aufgeteilt. Es wird dabei zunächst von der Bearbeitung eines Bildpaares ausgegangen. Zuerst werden die zu untersuchenden Bilder mit unterschiedlichen Filtern bearbeitet, um die Kantendetektion zu erleichtern. Anschließend werden allgemeine Vierecke in den verschiedenen Bildern erkannt und abgespeichert. Die Schwerpunkte der Vierecke werden berechnet und diesen wird ein Deskriptor zugeordnet, der verschiedene Merkmale des Vierecks sowie der umliegenden Vierecke be-

schreibt. Abschließend erfolgt eine Zuordnung der Merkmalspunkte über die berechneten Deskriptoren. Die folgenden Schritte wurden in C++ unter Verwendung von Funktionen aus der OpenCV Bibliothek (opencv 3.2.0, opencv.org) programmiert.

3.1 Angleichung und Filterung der Bilder

Um die Aufnahmen zu vergleichen und die Vierecke besser detektieren zu können, werden die Bilder vorbehandelt. Ziel ist es die Kanten zu verstärken, Bildrauschen abzuschwächen und die Bilder radiometrisch auf ein Niveau zu bringen, um möglichst viele Vierecke in den historischen Aufnahmen zu detektieren. Der Deskriptor (Kapitel 3.3) wird robuster wenn eine große Anzahl an Vierecken erkannt wird. Um dies zu erreichen, wurden verschiedene Filter an unterschiedlichen Bildern empirisch getestet. Eine Automatisierung der Filterung ohne manuelle Schwellwertsetzung stand dabei im Vordergrund. Die Variante, bei der die meisten Strukturen erkannt wurden, wird im Folgenden erläutert.

Zuerst wird jeweils ein Histogrammausgleich auf beide Bilder angewandt, um den Kontrast in den meist kontrastarmen Fotografien zu verstärken. Nach diesem Schritt wird das Bild bilateral gefiltert, um Kanten zu erhalten aber Texturen (z. B. Vorhänge in Fenstern, Steinkörnung) zu vermindern (TOMASI & MANDUCHI 1998). Anschließend werden in den Aufnahmen mit Hilfe des Canny-Algorithmus‘ (CANNY 1986) und automatisch bestimmten Parametern mittels Otsu-Schwellwertbildung (OTSU 1979) Kanten detektiert. Diese Kanten werden abschließend mittels Closing verstärkt und geschlossen (Abbildung 2).

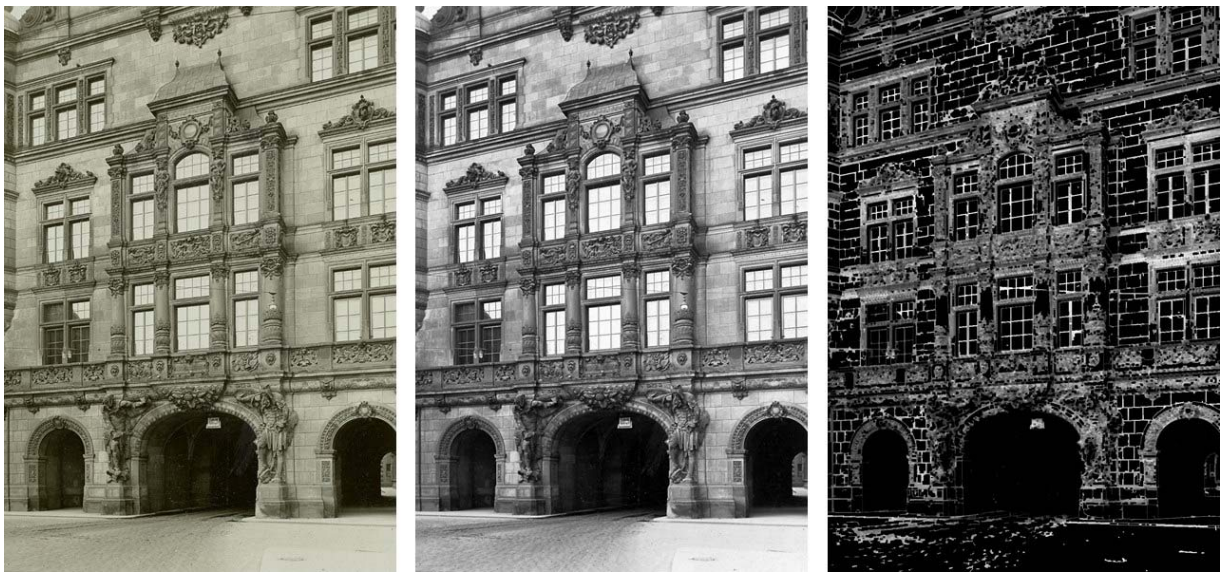


Abb. 2: Originalbild (links), mit Histogrammausgleich und bilateralem Filter bearbeitetes Bild (Mitte) und mit Canny-Algorithmus und Closing bearbeitetes Bild (rechts)

3.2 Detektion der Rechtecke

Als besonders robuste Geometrien für die Verknüpfung der Bilder haben sich Rechtecke im Objektraum herausgestellt (häufiges Vorkommen in Gebäudeaufnahmen, charakteristische Eigenschaften). Aber auch Liniengeometrien werden häufig für merkmalsbasierte Zuordnungen verwendet. Problematisch ist hierbei allerdings die Wahl definierter Punkte, da detektierte Geraden

erst in Liniensegmente geteilt werden müssen. Die Liniensegmente können in zwei Bildern durch fehlerhafte Segmentierung unterschiedlich fragmentiert sein. Eine Wahl von Fixpunkten zur Beschreibung der Linien (z. B. Endpunkte der Linie) ist erschwert (ZHANG & KOCH 2013). Zusätzlich ist die eindeutige Beschreibung von Liniensegmenten für die Zuordnung in einem zweiten Bild deutlich komplexer (WANG et al. 2009). Rechtecke bzw. Vierecke können hingegen bei eindeutiger Detektion auf ihren Schwerpunkt oder ihre vier Eckpunkte reduziert werden.
Bei ersten Versuchen wurden die Linien bzw. Geraden und anschließend die Vierecke mittels Hough-Transformation detektiert. Bei diesem Ansatz sind allerdings die lange Berechnungszeit und die zu definierenden Parameter problematisch. Diese lassen sich schlecht automatisch bestimmen und sind besonders für sich stark unterscheidende Bilder sehr unterschiedlich. Für die Liniendetektion wird in Zukunft noch der sogenannte Line Segment Detektor (LSD) getestet, der etwas schneller als das jetzige Verfahren ist und die Rechtecke möglicherweise robuster detektiert (VON GIOI et al. 2012).
Zum jetzigen Zeitpunkt werden die Vierecke im Canny-Bild über ein Template-Matching gefunden. Zuerst werden Konturen mittels „border following“ (SUZUKI 1985) erkannt und hierarchisch abgespeichert. Anschließend werden Konturen, deren Figur grob mit dem Template eines Rechtecks übereinstimmt, detektiert. Alle anderen vorhandenen Konturen werden verworfen. Die verbliebenen Konturen sind zu diesem Zeitpunkt jeweils als Vektor von vielen zweidimensionalen Punkten gespeichert. Eine Vereinfachung der Konturen erfolgt rekursiv durch den Douglas-Peucker-Algorithmus (DOUGLAS & PEUCKER 1973). Alle Konturen die anschließend aus genau vier Punkten bestehen, werden abgespeichert (Abbildung 3).

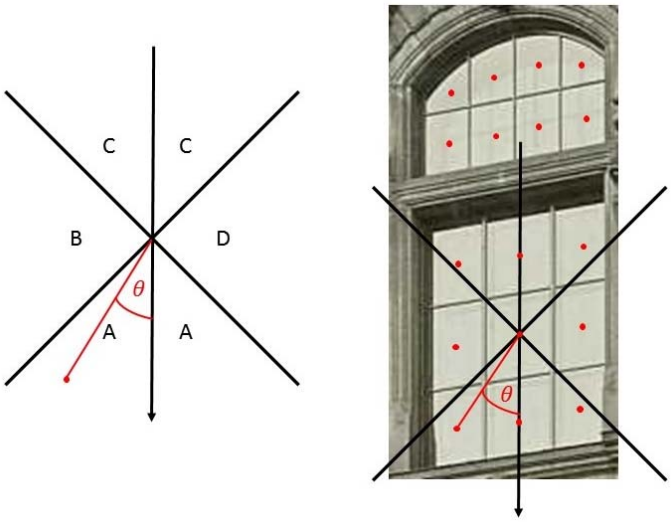
Abb. 3: Canny-Bild (links), detektierte Konturen nach Template-Matching (Mitte) und Vereinfachung der Konturen auf allgemeine Vierecke (rechts)

In der Abbildung kann man erkennen, dass besonders fehlerhafte Strukturen entfernt werden und die Konturen effizient als allgemeine Vierecke durch vier Eckpunkte beschrieben werden können.
In weiteren Untersuchungen soll aus den ermittelten Vierecken die projektive Verzerrung des Bildes ermittelt werden. Eine entzerrte Darstellung würde die im Objektraum rechteckigen

Strukturen auch im Bildraum unter Berücksichtigung der Verzeichnung als rechteckige Strukturen darstellen. Für den Deskriptor könnten weitere Annahmen (z. B. Seitenverhältnis der Rechtecke in zwei Bildern sollte gleich sein) getroffen werden, wodurch die Zuordnung von Punkten zuverlässiger werden sollte.

3.3 Beschreibung der Vierecke durch einen Deskriptor

Um die allgemeinen Vierecke eindeutig zu beschreiben, wird für jedes einzelne ein Deskriptor definiert. Andere Verfahren verwenden auch Deskriptoren, welche sich allerdings ähnlich wie bei SIFT auf Grauwertumgebungen um besondere Merkmalspunkte beschränken (MIČUŠÍK et al. 2008, WANG* et al. 2009). Der im Folgenden vorgestellte Deskriptor hat eine ähnliche Datenstruktur wie der SIFT-Deskriptor, ist jedoch statt mit vielen Gradientenrichtungen mit Werten gefüllt, die die Vierecke sowohl geometrisch als auch semantisch-topologisch beschreiben. Die geometrischen Eigenschaften und Nachbarschaftsbeziehungen auf die der Deskriptor beruht, sind weitestgehend invariant zu projektiven Transformationen (HARTLEY & ZISSERMAN 2003).
Acht verschiedene Werte beschreiben jedes gefundene Viereck V_i (wobei $i \in 1..n$ Anzahl gefundener Vierecke). Zunächst wird für jeden Schwerpunkt S_i des Vierecks V_i (Ursprungsviereck) die Distanz zu allen Schwerpunkten S_j benachbarter Vierecke V_j berechnet (wobei $j \in 1..n-1$). Diese Distanzen werden aufsteigend sortiert und die jeweils k ganzzahligen nächsten Nachbarn eines einzelnen Vierecks können betrachtet werden.
Die ersten vier Werte des Deskriptors beschreiben die Lage benachbarter Vierecke zum Ursprungsviereck. Es wird davon ausgegangen, dass bei erfolgreicher Viereckdetektion gleiche Vierecke in zwei Bildern ähnlich viele Nachbarn in vier verschiedenen Richtungen (Quadranten) haben sollten. Für die Berechnung wird der Winkel θ_j zwischen der negativen y-Achse und der Strecke zwischen den Schwerpunkten $S_i,\ i \in 1..n$ und $S_j,\ j \in 1..k$ bestimmt und einer von vier verschiedenen Richtungen (A, B, C, D) zugewiesen. Die Summe der Treffer in den verschiedenen Richtungen wird anschließend mit k normiert (Abbildung 4):



$$j \in 1..k; \quad k = 16$$

$$A: Deskriptor_i[1] = \frac{\sum_{j=1}^{k} \theta_j \in 315°..45°}{k} = \frac{3}{16}$$

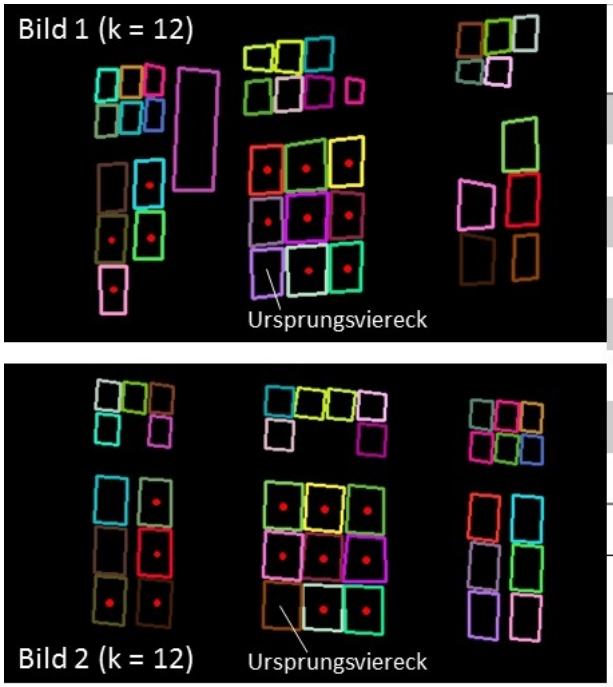
$$B: Deskriptor_i[2] = \frac{\sum_{j=1}^{k} \theta_j \in 45°..135°}{k} = \frac{1}{16}$$

$$C: Deskriptor_i[3] = \frac{\sum_{j=1}^{k} \theta_j \in 135°..225°}{k} = \frac{11}{16}$$

$$D: Deskriptor_i[4] = \frac{\sum_{j=1}^{k} \theta_j \in 225°..115°}{k} = \frac{1}{16}$$

Abb. 4: Beschreibung der ersten vier Werte des Deskriptors für jedes Viereck. Benachbarte Vierecke werden je nach Winkel den verschiedenen Regionen A, B, C und D zugeordnet und anschließend mittels der Gesamtzahl der betrachteten Vierecke normiert.

Der Deskriptor ist somit nur bedingt rotationsinvariant. Da historische Aufnahmen nach der Digitalisierung in der Regel „aufrecht“ orientiert sind, fallen die umliegenden Vierecke meist in den gleichen Quadranten. Um vollständige Rotationsinvarianz zu erreichen, ist bereits geplant die Hauptrichtung der Vierecke zu berechnen und die Winkel zwischen den Hauptachsen (in Abbildung 4 etwa bei 0° und 85°) und der Strecke zwischen den Schwerpunkten zu berechnen.
Die nächsten vier Werte des Deskriptors beschreiben Eigenschaften des Ursprungsvierecks, die mit den k benachbarten Vierecken verglichen werden. Dazu werden verschiedene Verhältnisse miteinander verglichen, die nach einer beliebigen projektiven Transformation von einem Bild in ein zweites Bild erhalten bleiben. Die Werte im Deskriptor sind dann bei der Detektion aller (oder fast aller) Rechtecke in der näheren Umgebung gleich (bzw. fast gleich). Die für das Brute-Force-Matching berechnete L2-Norm wird annähernd 0 und zwei homologe Vierecke in zwei Bildern werden einander zugeordnet (Abbildung 5).



Deskriptor Bild 1	Deskriptor Bild 2	$(D1 - D2)^2$
0	0	0
0,25	0,25	0
0,42	0,42	0
0,25	0,25	0
0,75	1,00	0,06
0,83	1,00	0,03
1,00	1,00	0
0,71	0,79	0,01
Summe $\triangleq$ L2 − Norm		**0,10**

$$L2 - Norm = \sum_{i=1}^{8} (Desk_1[i] - Desk_2[i])^2$$

Abb. 5: Zwei homologe Vierecke in zwei verschiedenen Bildausschnitten und deren Nachbarschaft für k=12 (markiert mit rotem Punkt). Für jedes Viereck werden die acht Deskriptorwerte berechnet, die in der rechten Tabelle zu sehen sind. Deskriptoren werden über Brute-Force-Matching einander zugeordet. Dazu wird die L2-Norm zwischen verschiedenen Kombinationen zweier Deskriptorvektoren berechnet. Bei geringen Werten werden die Deskriptoren und somit die Vierecke einander zugeordnet

Für den fünften Wert des Deskriptors werden die Flächen der benachbarten Vierecke A_j mit der Fläche des Ursprungsvierecks A_i verglichen. Bei annähernd gleich großen Flächen wird der Deskriptor um 1 inkrementiert und nach Prüfung aller benachbarten Flächen wieder mit k normiert, um zur besseren Vergleichbarkeit im gleichen Wertebereich der ersten vier Werte zu bleiben.

$$Flächenverhältnis_i[j] = \begin{cases} \forall\ (A_i \approx A_j) = 1 \\ \forall\ \overline{(A_\iota \not\approx A_j)} = 0 \end{cases};\ j \in 1..k;\ i \in 1..n$$

$$Deskriptor_i[5] = \frac{\sum_{j=1}^{k} Flächenverhältnis[j]}{k};\ i \in 1..n$$

Der sechste Wert beschreibt die Parallelität der Seiten benachbarter Vierecke. Dafür wird der Winkel der Seiten $a_i,\ b_i$ des Ursprungsvierecks mit den Winkeln der Seiten $a_j,\ b_j$ des benachbarten Vierecks verglichen. Bei Parallelität wird der Wert des Deskriptors um 1 inkrementiert und schließlich wieder mit k normiert.

$$Parallelität_i[j] = \begin{cases} \forall\ (a_i \parallel a_j) \cap (b_i \parallel b_j) = 1 \\ \forall\ \overline{(a_\iota \parallel a_\jmath) \cap (b_\iota \parallel b_\jmath)} = 0 \end{cases};\ j \in 1..k;\ i \in 1..n$$

$$Deskriptor_i[6] = \frac{\sum_{j=1}^{k} Parallelität[j]}{k};\ i \in 1..n$$

Der siebte Wert des Deskriptors zeigt das Verhältnis der Seitenverhältnisse benachbarter Vierecke. Es wird das Seitenverhältnis $\frac{a_i}{b_i}$ des Ursprungsvierecks mit dem des benachbarten Vierecks $\frac{a_j}{b_j}$ verglichen und bei annähernd gleichen Verhältnissen wird der Deskriptorwert um 1 inkrementiert.

$$Seitenverhältnis_i[j] = \begin{cases} \forall \left(\frac{a_i}{b_i} \approx \frac{a_j}{b_j}\right) = 1 \\ \forall \left(\frac{a_i}{b_i} \not\approx \frac{a_j}{b_j}\right) = 0 \end{cases};\ j \in 1..k;\ i \in 1..n$$

$$Deskriptor_i[7] = \frac{\sum_{j=1}^{k} Seitenverhältnis[j]}{k};\ i \in 1..n$$

Wert 8 zeigt die Hausdorff-Distanz zwischen Ursprungsviereck und benachbartem Viereck und wird zuerst mit k und dann mit 100 normiert um den Wertebereich der anderen Zahlen des Deskriptors zu erhalten. Die Hausdorff-Distanz kann als Maß für die Ähnlichkeit zweier Punktmengen gesehen werden.

$$Deskriptor_i[8] = \frac{\sum_{j=1}^{k} Hausdorff(R_i, R_j)}{k} \cdot \frac{1}{100};\ i \in 1..n$$

Der vollständige Deskriptor besteht damit aus 8 verschiedenen Werten, die sowohl maßstabs- als auch translationsinvariant sind. Eine gewisse Rotationsinvarianz ist bereits gegeben. Diese soll aber noch verbessert werden um einen affininvarianten Deskriptor zu erhalten. Die Werte des Deskriptors wurden alle auf Fließkommazahlen zwischen 0 und 1 normiert, um vergleichbare Ergebnisse bei den Abweichungen zwischen zwei Deskriptoren zu erhalten. Eine Korrelation der letzten vier Werte des Deskriptors muss noch geprüft werden.

3.4 Zuordnung der Merkmalspunkte

Die Schwerpunkte der Vierecke, die über den bereits erläuterten Deskriptor beschrieben werden, werden mittels Brute-Force-Matching einander zugeordnet. Das bedeutet, dass jeder Deskriptor in Bild 1 mit jedem Deskriptor in Bild 2 verglichen und das beste Ergebnis (geringster Wert der L2-Norm zwischen zwei Deskriptoren) abgespeichert wird. Die Ergebnisse werden weiter mit Hilfe eines Symmetrietests gefiltert. Das bedeutet, dass Zuordnungen von Bild 1 nach Bild 2 nur akzeptiert werden, wenn auch eine Zuordnung von Bild 2 nach Bild 1 der gleichen Punkte existiert (Abbildung 6).

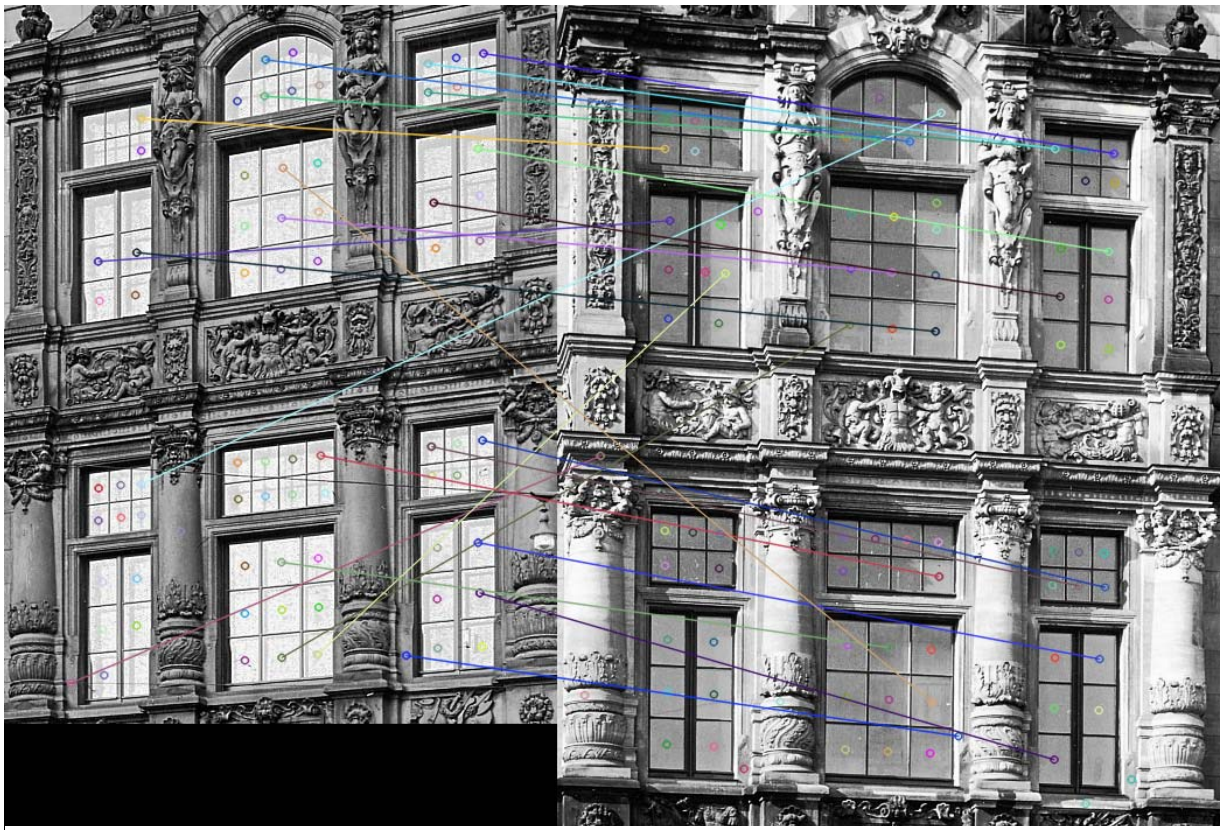


Abb. 6: Zuordnung von Punkten zwischen zwei unterschiedlichen Bildausschnitten nach Brute-Force-Matching und Symmetrietest für k = 20 (Korrekte Matches: 8, Inkorrekte Matches: 14)

Die verbliebenen Zuordnungen werden für die Berechnung einer Fundamentalmatrix verwendet. Fehlerhafte Matches werden mittels RANSAC (FISCHLER & BOLLES 1981) eliminiert. Für die untersuchten Aufnahmen bleiben je nach Größe des Bildes und Anzahl der gefundenen Vierecke 7-15 korrekte Zuordnungen übrig. Diese Ergebnisse werden vor allem bei Bildern von Fassaden mit vielen Fenstern oder besonderen Steinformationen erreicht. Bei Bildern mit weniger als sieben detektierten Vierecken (Kapitel 3.2) ist eine Berechnung der Fundamentalmatrix nicht möglich. Fehlerhafte Zuordnungen die zusätzlich bestehen bleiben, sind meist Vierecke, die nur ein Viereck vom korrekten Match entfernt sind (Abbildung 7). Diese entstehen, wenn bereits in Schritt 3.2 das korrekte Viereck nicht detektiert wurde oder mehrere Vierecke in der Nachbarschaft nicht zu sehen sind. Auch bei starken perspektivischen Unterschieden oder langer Basisli-

nie sind die detektierten Vierecke bzw. deren Nachbarschaften meist zu unterschiedlich, um sie einander zuordnen zu können.

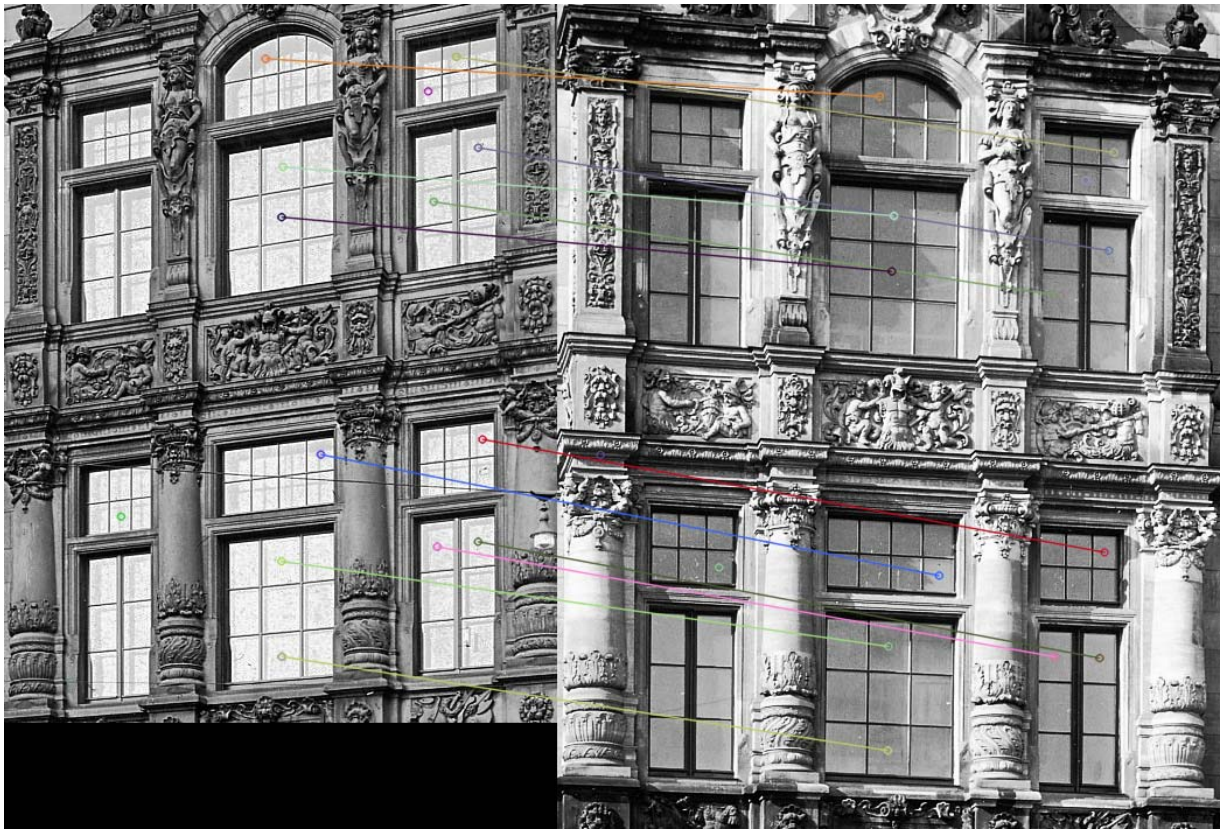
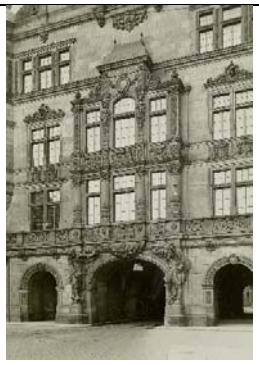


Abb. 7: Zuordnung von Punkten nach Berechnung der Fundamentalmatrix und Filterung mittels RANSAC jeweils für $k \in 7..70$ (Korrekte Matches: 10, um ein Rechteck verschobene Matches: 2)

4 Ergebnisse

Wie in Abbildung 6 und 7 zu sehen ist, sind die Ergebnisse vor allem dort gut, wo viele Vierecke detektiert werden können. Die Zuordnung über den bereits beschriebenen Deskriptor ist robust und ausschließlich über geometrische Merkmale bestimmbar. Radiometrische Differenzen wie Helligkeitsunterschiede oder Kontraständerungen können so vollständig eliminiert werden (außer wenn sie sich bereits stark auf die Viereckdetektion auswirken). Problematisch ist die Zuordnung, wenn insgesamt nur wenige Vierecke detektiert werden können. Dies tritt beispielsweise bei Verdeckungen, kleinen Bilddateien, Gebäudeteilen ohne Fenster/Steine auf. In Tabelle 1 sind verschiedene Bildbeispiele und die korrekte Anzahl an Zuordnungen aufgeführt.

Tab. 1: Anzahl korrekter und fehlerhafter Zuordnungen durch den entwickelten Deskriptor bei verschiedenen historischen Bildpaaren. Zusätzlich werden Ergebnisse bei Bildpaaren des Benchmark-Datensatzes aus LI ET AL. (2016) gezeigt.

Bild			
Beschreibung	Georgentor 1964-69	Dinglinger Haus vor 1945	Zwinger 1952
Zuordnungen	Korrekt: 15 Falsch: 2	Korrekt: 7 Falsch: 2	Korrekt: 5 Falsch: 2
Beschreibung	Benchmarkbild building	Benchmarkbild outdoor	Benchmarkbild bikes
Zuordnungen	Korrekt: 7 Falsch: 4	Korrekt: 7 Falsch: 3	Korrekt: 3 Falsch: 1

In der Tabelle sieht man, dass die Anzahl korrekter Zuordnungen in den meisten Bildern noch zu gering ist. Die gefundenen Matches sind zwar sehr robust aber meist nicht über das gesamte Bild verteilt und somit nicht optimal für die Berechnung einer Fundamentalmatrix. Da jeweils der Schwerpunkt homologer Vierecke nicht zwingend exakt den gleichen Objektpunkt beschreibt, könnte die Reduzierung der Vierecke auf deren vier Ecken zu einer Verbesserung der Fundamentalmatrix führen. Denkbar wäre die Einführung weiterer Geometrien (Kreise, Linien), um den Deskriptor lokal zu verbessern. Über einen iterativen Ansatz könnten über die bereits gefundenen korrekten Zuordnungen weitere Matches ermittelt werden, die in das Modell der berechneten Fundamentalmatrix passen.

Des Weiteren zeigen die Ergebnisse, dass die Auswahl eines fixen Parameters k für die Anzahl betrachteter benachbarter Vierecke schwierig ist. Während für den Parameter $k = 14$ acht korrekte Zuordnungen in den zwei Bildern des Georgentors vorhanden sind, so sind für den Parameter $k = 9$ nur 2 korrekte Zuordnungen vorhanden (Abbildung 8).

Die korrekten Zuordnungen sind zudem meist von unterschiedlichen Vierecken an verschiedenen Positionen im Bild. Tendenziell steigt die Anzahl korrekter Zuordnungen leicht mit steigendem k. Eine Auswahl eines möglichst großen k ist sinnvoll, wenn die Mehrzahl der Vierecke auch im zweiten Bild zu sehen sind.

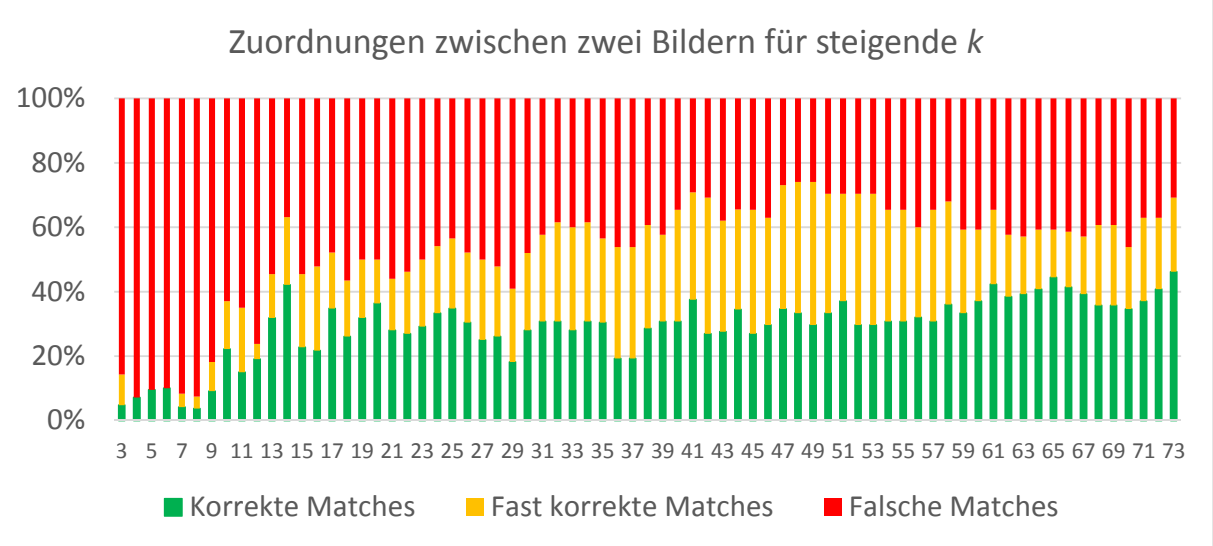


Abb. 8: Anzahl korrekter, fast korrekter (um ein Viereck verschoben) und falscher Zuordnungen von Punkten zwischen den zwei Bildausschnitten aus Abbildung 5 für $k \in 3..73$

Stabiler ist jedoch eine Berechnung über mehrere k und anschließend die Auswahl der am meisten vorkommenden Zuordnungen. Dies hat sich als robuste aber durchaus rechenintensive Lösung herausgestellt. Die Aussagen zum Parameter k sind an dem dargestellten Bildpaar (Abbildung 7) gut zu sehen, zeigen sich aber in anderen Bildpaaren ähnlich.

5 Fazit & Ausblick

Das präsentierte Verfahren zeigt die Zuordnung von Merkmalspunkten über geometrische und semantisch-topologische Eigenschaften. Dabei werden in historischen Bildern zuerst Vierecke gesucht und über einen eigens entwickelten Deskriptor, der auch Nachbarschaftsverhältnisse von Vierecken darstellt, beschrieben. Die Größe der Nachbarschaft kann über einen Parameter frei gewählt werden. Die Deskriptoren werden anschließend zwischen den zwei Bildern mittels Brute-Force-Matching zugeordnet. Schließlich werden die am häufigsten vorkommenden Zuordnungen und deren Fundamentalmatrix als korrekt angenommen.
Die verbliebenen Zuordnungen sind besonders robust und können unabhängig von einer schlechten SNR und periodisch vorkommenden Objektstrukturen berechnet werden.
Da besonders eine große Anzahl an Vierecken hilfreich ist, soll die Viereckdetektion in weiteren Arbeitsschritten verbessert werden. Denkbar ist die Verwendung eines Liniensegment-Detektors, bei dem einzelne Linien zu Vierecken zusammengefügt werden. Zusätzlich ist eine rechentechnische Optimierung geplant, um die Zuordnung für eine große Anzahl an Bildpaaren zu ermöglichen. An einer rotationsinvarianten Form des Deskriptors wird bereits gearbeitet.

6 Danksagung

Dieser Beitrag wurde durch die Förderung des Bundesministeriums für Bildung und Forschung unter der Projektnummer 01UG1630 im Rahmen der Nachwuchsforschergruppe UrbanHistory4D ermöglicht.

7 Literaturverzeichnis

ALI, H. K. & WHITEHEAD, A., 2014: Feature Matching for Aligning Historical and Modern Images. International Journal of Computer Applications, **21**(3), 188-201.

BAY, H., TUYTELAARS, T. & VAN GOOL, L., 2006: Surf: Speeded up robust features. Computer vision – ECCV **2006**, 404-417.

CANNY, J., 1986: A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence, **6**, 679-698.

DOUGLAS, D. H., & PEUCKER, T. K., 1973: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartographica: The International Journal for Geographic Information and Geovisualization, **10**(2), 112-122.

FISCHLER, M. A. & BOLLES, R. C. 1981: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, **24**(6), 381-395.

VON GIOI, R. G., JAKUBOWICZ, J., MOREL, J. M. & RANDALL, G., 2012: LSD: a line segment detector. Image Processing On Line, **2**, 35-55.

HARTLEY, R. & ZISSERMAN, A., 2003: Multiple view geometry in computer vision. Cambridge university press.

LI, K., YAO, J., LU, M., HENG, Y., WU, T. & LI, Y., 2016: Line segment matching: A benchmark. IEEE Winter Conference on Applications of Computer Vision (WACV), 1-9.

LOWE, D. G., 2004: Distinctive image features from scale-invariant keypoints. International journal of computer vision, **60**(2), 91-110.

MIČUŠÍK, B., WILDENAUER, H. & KOSECKA, J., 2008: Detection and matching of rectilinear structures. IEEE Conference on Computer Vision and Pattern Recognition, 1-7.

MUJA, M. & LOWE, D.G., 2012: Fast approximate nearest neighbors with automatic algorithm configuration. International Conference on Computer Vision Theory and Applications (VISAPP'09), 331-340.

OTSU, N., 1979: A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, **9**(1), 62-66.

RUBLEE, E., RABAUD, V., KONOLIGE, K. & BRADSKI, G., 2011: ORB: An efficient alternative to SIFT or SURF. IEEE international conference on Computer Vision (ICCV), 2564-2571.

SCHINDLER, G. & DELLAERT, F., 2012: 4D cities: analyzing, visualizing, and interacting with historical urban photo collections. Journal of Multimedia, **7**(2), 124-131.

STYLIANOU, A., ABRAMS, A. & PLESS, R., 2015: Characterizing Feature Matching Performance over Long Time Periods. IEEE Winter Conference on Applications of Computer Vision, 892-898.

SUZUKI, S. 1985: Topological structural analysis of digitized binary images by border following. Computer vision, graphics, and image processing, **30**(1), 32-46.

WANG, L., NEUMANN, U., & YOU, S., 2009: Wide-baseline image matching using line signatures. IEEE 12th International Conference on Computer Vision, 1311-1318.

WANG*, Z., WU, F., & HU, Z., 2009: MSLD: A robust descriptor for line matching. Pattern Recognition, **42**(5), 941-953.

WOLFE, R., 2015: Modern to historical image feature matching. http://robbiewolfe.ca/programming/honoursproject/report.pdf.

ZHANG, L., & KOCH, R., 2013: An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency. Journal of Visual Communication and Image Representation, **24**(7), 794-805.

Sentinel-Daten für digitale und interaktive Anwendungen im Schulunterricht

CLAUDIA LINDNER[1], HENRYK HODAM[1], ANNETTE ORTWEIN[1], JOHANNES SCHULTZ[1], FABIAN SELG[1] & ANDREAS RIENOW[1]

Zusammenfassung: Das FIS-Projekt entwickelt digitale, wissenschaftspropädeutische, interaktive, intermediale und interdisziplinäre Unterrichtsmaterialien für verschiedene MINT-Fächer, um Schülern die Möglichkeiten der Fernerkundung für Untersuchungen im Bereich der Mensch-Umwelt-Interaktionen aufzuzeigen.
Dank der neuen, frei verfügbaren Sentinel-Daten konnte ein Update der den bereits entwickelten Lernmodulen zugrundeliegender Datensätze durchgeführt werden, deren Vorteile von einer Verbesserung der Erkennbarkeit und Verortung der Untersuchungsgebiete bis hin zu einer effektiveren Anwendbarkeit der eigens für das Projekt entwickelten GIS-Tools reicht. Neue Lernmaterialen in Form von Mini-MOOCs befinden sich in der Entwicklung und werden sowohl Satelliten- als auch ISS-basierte Daten verwenden.

1 Einleitung

Das Projekt „Fernerkundung in Schulen“ (FIS) befasst sich seit 2006 mit der Entwicklung und Weiterentwicklung von digitalen Lernmaterialen für die interaktive Anwendung der Fernerkundung im Schulunterricht. Die satellitenbasierte Erdbeobachtung wird nachhaltig und umfassend in den Schulunterricht integriert. Für die Fächer Mathematik, Physik und Informatik werden vor allem die physikalischen Grundlagen der Fernerkundung thematisiert, wohingegen die Analyse von Erdbeobachtungsdaten bei den anwendungsorientierten Fächern Geographie und Biologie im Vordergrund steht.
Zehn Jahre nach der Entwicklung der ersten Lerneinheiten erhalten diese ein Daten-Update in Form von Sentinel-2-Szenen, die die ursprünglich verwendeten Landsat-5-Szenen und andere ersetzen. Im FIS-Lernportal (http://www.fis.uni-bonn.de/unterrichtsmaterial/sentinel) werden die Unterrichtsmaterialen zum Download bereitgestellt.
Im Folgenden werden fünf Lerneinheiten mit ihren Zielen, Ablauf und Daten-Update vorgestellt, sowie das Konzept der sich in Entwicklung befindlichen Mini-MOOCs.

2 Fernerkundung im Schulunterricht – Konzept und Umsetzung

Satellitenbasierte Erdbeobachtung wurde in den Schulen lange Zeit nur als zusätzliches Anschauungsmaterial lehrplanrelevanter Sachverhalte genutzt, obwohl die Kompetenz, Luft- und Satellitenbilder selbstständig zu interpretieren und zu analysieren in zahlreichen nationalen und internationalen Bildungsstandards gefordert wird (NCGE 2012, DGFG 2014). Vorteile für die

[1] Ruhr-Universität Bochum, Geographisches Institut, AG Geomatik, Universitätsstraße 104, D-44801 Bochum, E-Mail: [claudia.lindner, henryk.hodam, annette.ortwein, johannes.schultz, fabian.selg, andreas.rienow] @rub.de

Schülerinnen und Schüler (SuS) werden v.a. im Bereich der naturwissenschaftlich ausgerichteten Bildung, der hohen Anschaulichkeit und der Stärkung verschiedener Kompetenzen, wie der räumlichen Orientierungs-, der Methoden-, Bewertungs- und Handlungskompetenz, gesehen (VOSS et al. 2011).

Die in FIS entwickelten sequenzierten Lerneinheiten sollen nicht nur die rein visuelle Interpretation von Luftbildern umfassen, sondern sind darauf ausgelegt, wissenschafts-propädeutisches Lernen zu fördern und den SuS den Zugang zu dieser Schlüsseltechnologie im Rahmen ihrer schulischen Ausbildung aufzuzeigen. Mit ihrem interdisziplinären Ansatz vereinen sie das in verschiedenen Schulfächern erlernte Grundwissen in einer realitätsnahen Fragestellung, die durch die Intermedialität den SuS digitale Geräte als wissenschaftliche Werkzeuge näherbringt. Auf Basis des moderaten Konstruktivismus sollen die SuS zum eigenständigen und reflektierten Umgang mit modernen Medien angeregt werden, um die Medienkompetenz zu fördern (GOETZKE et al. 2014). Durch die interaktiven Elemente können viele der Lernmodule zudem selbstständig und mit geringem Vorbereitungsaufwand für die Lehrerinnen und Lehrer (LuL) durchgeführt werden, was die Eigenaktivität und Selbstverantwortung der SuS stärkt.

Innerhalb der digitalen Lernmodule kommen verschiedene vereinfachte GIS-Tools zum Einsatz. Der Einsatz von realer GIS-Software ist für den allgemeinen Unterricht nicht geeignet, da sie einerseits auch durch die benötigte Hardware oft kostenintensiv, in jedem Falle aber für Laien unübersichtlich sind und zunächst eine intensive Schulung der Lehrer notwendig wäre. Aus diesem Grund wurden einzelne GIS-Funktionen in separierten Tools in Flash programmiert. So gibt es nicht nur vorgefertigte Ergebnisse, die freigeschaltet werden, sondern auch die Möglichkeit, selbst Untersuchungen und Berechnungen mit den bereitgestellten Satellitenbildern und Karten durchzuführen – und die Möglichkeit, etwas falsch zu machen und aus den eigenen Fehlern zu lernen.

Zur Verfügung stehen individuelle Tools zur NDVI- und Differenzbild-Berechnung (Change Detection), RGB-Klassifikation, Swipe sowie zwei MeteoViewer. Die programmierten GIS-Tools können in den Modulen wiederverwendet oder auch separat genutzt werden. Zusätzlich zu den o.g. verfügen die Lerneinheiten über Tools zur Streckenberechnung, Pegel- bzw. Schwellenwert-darstellung und zum Auslesen von Punktwerten. Sämtliche Tools bestechen durch ihre Übersichtlichkeit und einfache Bedienbarkeit und sind dazu ausgelegt, auch auf langsamen Rechnern mit niedrigen Spezifikationen ausgeführt zu werden.

Nicht nur die einzelnen Tools, sondern die kompletten digitalen Lernmodule sind vollständig in Flash programmiert. Dies ermöglicht den Download und die spätere Offline-Anwendung der Module. Aufgrund der noch nicht ausreichenden Internet-Kapazitäten an vielen deutschen Schulen, darunter auch die Partnerschulen des Projektes, ist eine Offline-Anwendbarkeit der Lernmodule unbedingt notwendig. Nachteilig sind die geringe Prozessierungsgeschwindigkeit sowie das Ende der Flash-Unterstützung durch den Hersteller Adobe im Jahr 2020.

Jedem Modul liegt eine Musterlösung und ein didaktischer Kommentar bei, der den LuL die Stundenplanungshilfe erleichtert und Hintergrundinformationen zu Aufgaben, Lernzielen und Einordnung in die Lehrpläne der Länder gibt sowie eine kurze inhaltliche Einführung in das Thema und die Fernerkundung im Allgemeinen enthält. Darüber hinaus sind detaillierte Informationen zu jedem Abschnitt des Lernmoduls und Zusatzinformationen zur den Funktionen verfügbar. Im FIS-Webportal selbst wird zudem eine Klassenverwaltung angeboten, in der LuL den

Lernfortschritt ihrer SuS nachvollziehbar bewerten können. Recherchetools für Laien und Fortgeschrittene, die die Hintergrundinformationen der Info-Boxen, eine Bildergalerie mit Beispielen, eine interaktive Einführung in Satellitensysteme und eine Linksammlung enthalten, stehen für jeden Interessierten bereit (RIENOW et al. 2015).
Um den didaktischen Ansatz der schulischen Fernerkundungsvermittlung und die Unterrichtsmaterialien vorzustellen, werden deutschlandweit an Schulen Lehrerfortbildungen durchgeführt. An einer Partnerschule des Projekts, dem Gymnasium Siegburg Alleestraße, wurde ein Differenzierungsfach „Geographie-Physik“ ab der 8. Klasse eingeführt, das die FIS-Anwendungen behandelt. Regelmäßige Expertenbesuche vom Projektteam in den Klassen und zu Lehrerfortbildungen der Schule stellen eine enge Verbindung mit ausführlichem Feedback sicher, dass der Qualität der Module und somit den SuS zugutekommt.
Im Folgenden werden ausgewählte Unterrichtseinheiten, die auf Sentinel-Daten basieren, vorgestellt.

2.1 Modul „Oasen – von nah und fern betrachtet“

Dieses Modul beschäftigt sich mit den Merkmalen und der Bedeutung verschiedener Oasenmodelle, dem Aufbau des Modells der Flussoasen, sowie den Eigenschaften von Satellitenbildern und ihrer diagnostischen Nutzung als allgemeine Grundlage der Fernerkundung. Die SuS sollen später in der Lage sein, grundlegende Modelle von Oasentypen erklären und diese Modelle mittels Fernerkundung mit der Wirklichkeit vergleichen zu können. Weiterhin sollen die angewendeten Fernerkundungsmethoden anhand der zugrundeliegenden allgemeinen physikalischen und methodischen Grundlagen erklärt werden können. Die Zielgruppe sind die 7. und 8. Klasse. Der Zeitaufwand für die SuS wird mit 1-2 Stunden berechnet.

2.1.1 Ablauf der Lerneinheit

Nach einem kurzen Informationstext wird das erste Wissen über Oasenmodelle in einer interaktiven Animation abgefragt. Die SuS werden mit einem georeferenzierten Komposit aus Grün-Rot-Infrarot (GRI) konfrontiert, auf dem Vegetation folglich rot erscheint, und erhalten in einer Infobox Informationen zu Wellenlängen, Reflexion und Komposit-Bildern in altersgerechter Form. Auf dem GRI-Komposit, auf dem eine Niloase 30km südlich von Kairo abgebildet ist, müssen verschiedene Landoberflächen korrekt identifiziert werden: Wüste, Siedlung, Fluss, Kanäle und Felder. Durch ein interaktives Lupentool, dem ein RGB-Komposit des Gebiets hinterlegt ist, kann das Bild vergrößert und in natürlichen Farben betrachtet werden. Durch drag-and-drop der Schilder auf die korrekten Flächen wird nach und nach eine thematische Karte auf dem Satellitenbild sichtbar. Zum Abschluss sollen die SuS die Unterschiede zwischen Modell und Karte erfassen und erklären können.

2.1.2 Einbindung der Sentinel-2-Daten

Für das Update mit Sentinel-2 Daten war es aufgrund der statischen Programmierung des Moduls für das GRI-Komposit notwendig, Bilddaten identischer Größen zu verwenden. Trotz der Reduzierung der Bildgröße auf die erforderlichen 450×450px konnte eine deutliche Verbesserung der Bildqualität erreicht werden (vgl. Abb. 1). Das dahinter gelegte RGB-Komposit im Lupen-Tool bestand zuvor aus einem Google-Earth-Bild. Um die Übereinstimmung zwischen dem

GRI- und dem RGB-Komposit zu erhöhen, wurden beide aus der gleichen Szene entnommen. Durch die geringere Auflösung der Sentinel-Daten gegenüber Google Earth wurde der im Lupen-Tool abgebildete Ausschnitt vergrößert, was zwar die Detailgenauigkeit verringert, die Erkennbarkeit der Verortung jedoch erhöht. Das verbesserte Modul ist auf der FIS-Internetseite bereits veröffentlicht.

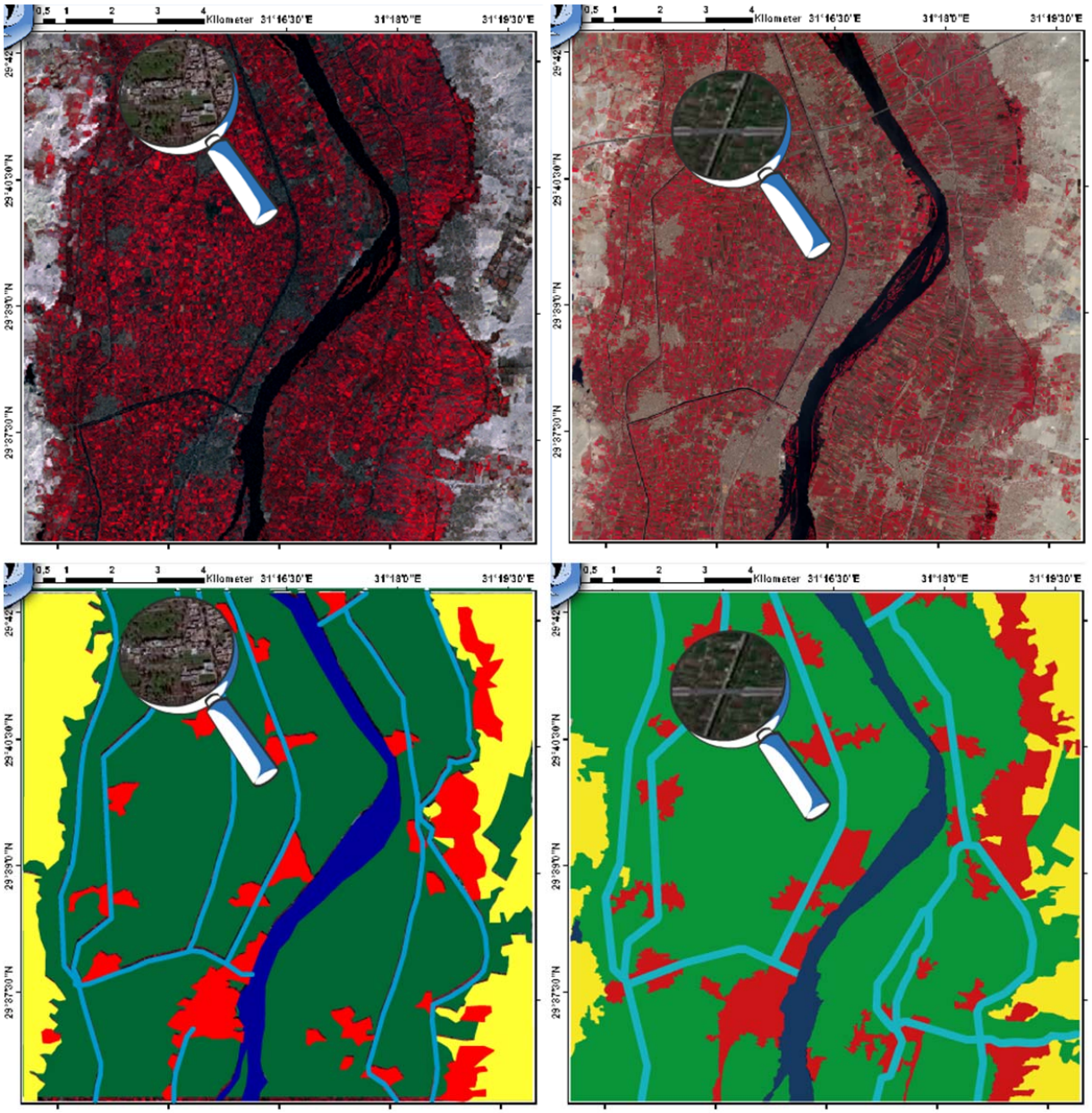


Abb. 1: Vergleich des alten Landsat-GRI-Komposits (oben links) mit dem aktuellen Sentinel-2-GRI-Komposit (oben rechts) und den darauf basierenden thematischen Karten (unten) im Modul „Oasen – von nah und fern betrachtet“.

2.2 Modul „Hochwasser – Umgang mit einer Naturgefahr“

In diesem Modul lernen die SuS über die Naturgefahr Hochwasser und, wie diese durch den Eingriff des Menschen beeinflusst wird. Darüber hinaus wird das Thema Standortentscheidungen thematisiert und erläutert, wie Geographische Informationssysteme diese erleichtern. Ziele des Lernmoduls sind nicht nur das Erkennen und Beschreiben natürlicher und anthropogener Ursachen von Hochwasser, sondern auch die Notwendigkeiten, Möglichkeiten und Grenzen von Schutzmaßnahmen zu begründen. Weiterhin sollen grundlegende GIS-Funktionen, wie das Layer-Prinzip, Bildvergleiche, Entfernungsmessung und Wertebereiche verstanden und in Verbindung mit einem Geländemodell für Standortentscheidungen genutzt werden. Zielgruppe sind die Stufen 7-9, der Zeitaufwand für die SuS wird mit ca. 2 Stunden berechnet.

2.2.1 Ablauf der Lerneinheit

Das Modul beginnt mit einem kurzen Informationstext, der sich mit der Standortsuche für eine Fabrik im Raum Bonn/Rhein-Sieg beschäftigt. Anhand eines Satellitenbildes, einer Landbedeckungskarte, der Positionsmarkierungen potentieller Standorte und der Lage der Autobahnen im Untersuchungsgebiet ermitteln die SuS den bestmöglichen Standort nach den zuvor geforderten Kriterien. Dabei stehen ihnen zunächst nur zwei GIS-Tools zur Verfügung: Eine Anzeige der angeklickten Positionswerte und ein Linienwerkzeug zur Streckenmessung. Nach einem kurzen Tutorial zur Benutzung der Layer und Werkzeuge können die SuS bereits intuitiv Raster- und Vektordaten miteinander verbinden und Daten aus ihnen auslesen.
Danach wird die anhand der Landnutzungsdaten getroffene Standortwahl in Frage gestellt. Zu den bekannten Daten wird ein digitales Geländemodell hinzugefügt, aus dem die jeweiligen Geländehöhen abgefragt werden können. Mit einem Swipe-Tool kann eine Siedlungskarte über das DGM gelegt werden. Anthropogene Strukturen und ihre Auswirkungen auf das Gelände sollen identifiziert werden. Die enthaltene Infobox erklärt, wie ein durch Laserscanning ein Geländemodell erstellt wird und enthält ein interaktives Landschaftsmodell, in dem verschiedene großangelegte Landschaftsveränderungs- bzw. Baumaßnahmen und ihr Einfluss auf das Hochwasserrisiko dargestellt werden.
Basierend auf dem bereits bekannten DGM kann im letzten Teil des Moduls der Pegel verändert werden. Durch eine einfache Klassifizierung, die alle Bereiche des DGMs unter dem gewählten Wert einfärbt, werden anschaulich die Auswirkungen eines Hochwassers an der Siegmündung aufgezeigt (vgl. Abb. 2).

2.2.2 Einbindung der Sentinel-2-Daten

Die aktuellen Sentinel-Daten werden in diesem Modul in Form eines RGB-Komposits eingesetzt, welches das alte Landsat-Bild von 2005 ersetzt. Es kommt im ersten GIS-Tool zum Einsatz, wo es durch die höhere räumliche Auflösung die Erkennbarkeit der einzelnen Elemente verbessert und den SuS so hilft, sich zu orientieren. Das verbesserte Modul ist auf der FIS-Internetseite bereits veröffentlicht.

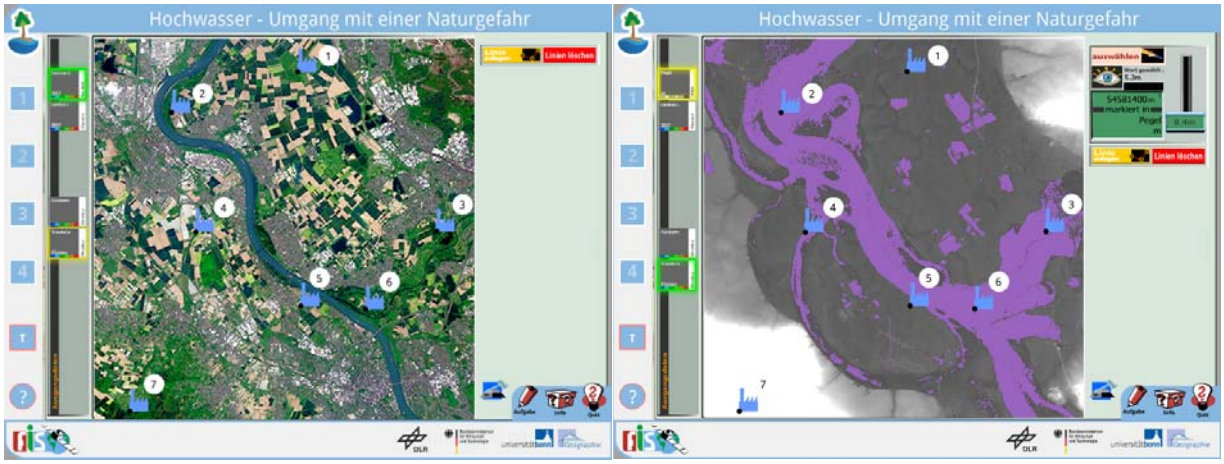


Abb. 2: Ansicht des digitalen Lernmoduls „Hochwasser – Umgang mit einer Naturgefahr“. Rechts: Sentinel-2-RGB-Komposit mit potentiellen industriestandorten. Links: Interaktiver Pegel im Raum Bonn/Siegmündung.

2.3 Modul „Vom Satellitenbild zur Karte“

Eine der grundlegendsten Funktionen der Fernerkundung wird in diesem Modul abgedeckt, indem die SuS eine einfache Klassifikation eines Satellitenbildes durchführen. Hierfür lernen sie, Satellitenbilder über Landbedeckung und Landnutzung zu interpretieren und sich mit ihnen räumlich zu orientieren. Darauf aufbauend lernen sie, wie Klassifikationen in den Grundzügen funktioniert und wie daraus thematische Karten abgeleitet werden. Das Modul ist für die Jahrgangsstufe 5 ausgelegt und benötigt ca. 1-2 Stunden zur Durchführung.

2.3.1 Ablauf der Lerneinheit

Zu Anfang des Moduls wird ein RGB-Komposit von Köln mit Ausschnitten topographischer Karten verschiedener Städte verglichen, um die räumliche Orientierung zu gewährleisten. Die SuS müssen nicht nur die richtige Stadt erkennen, sondern auch korrekt interpretieren, was auf einem Satellitenbild erkennbar ist und was nicht. Danach wird das gezeigte RGB-Komposit anhand eines einfachen Klassifikationstools nach den Anteilen von Rot, Grün, Blau sowie der Ähnlichkeit klassifiziert. Jede Klasse wird einzeln definiert, eingestellt und gespeichert, bis möglichst viele Pixel mit der Klassifikationsmaske bedeckt sind. Zusätzlich wird angezeigt, wie groß die jeder Klasse angehörende Fläche ist.

Das Klassifikationstool wurde speziell für diese Aufgabe entwickelt, ist daher sehr simpel gehalten und erlaubt es den SuS praktisch ohne GIS-Vorwissen, eine anschauliche thematische Karte zu erstellen. Klassen können selbst definiert und beschrieben werden, sodass jede/r SuS zu einem anderen Ergebnis kommen und diese anschließend miteinander und mit der Musterlösung verglichen werden können. Abschließend wird über Schwellenwerte informiert und die Qualität der Klassifikation über die Flächengrößen der einzelnen Klassen abgefragt. Die SuS können Pixel, die nicht klassifiziert wurden, vom Tool auffüllen lassen, um eine vollständige Karte zu erhalten.

2.3.2 Einbindung der Sentinel-2-Daten

Um die Rechenzeit in diesem interaktiven Klassifikationstool auch auf Schulrechnern gering zu halten, wurde die Größe des zu klassifizierenden Bildes auf 500×500px festgelegt. Diese Maße sind auch für den Ausschnitt der neuen Sentinel-Szene einzuhalten. Dennoch wurde eine deutliche Verbesserung, v.a. in der Anwendung des Klassifikationstools, festgestellt. Nicht nur sind Strukturen klarer zu erkennen, sondern durch das auflösungsbedingt verringerte Auftreten von Mischpixeln wird auch das Klassifikationspotential erheblich verbessert und die Anzahl nicht klassifizierbarer Pixel verringert (Abb. 3). Das verbesserte Modul ist auf der FIS-Internetseite bereits veröffentlicht.

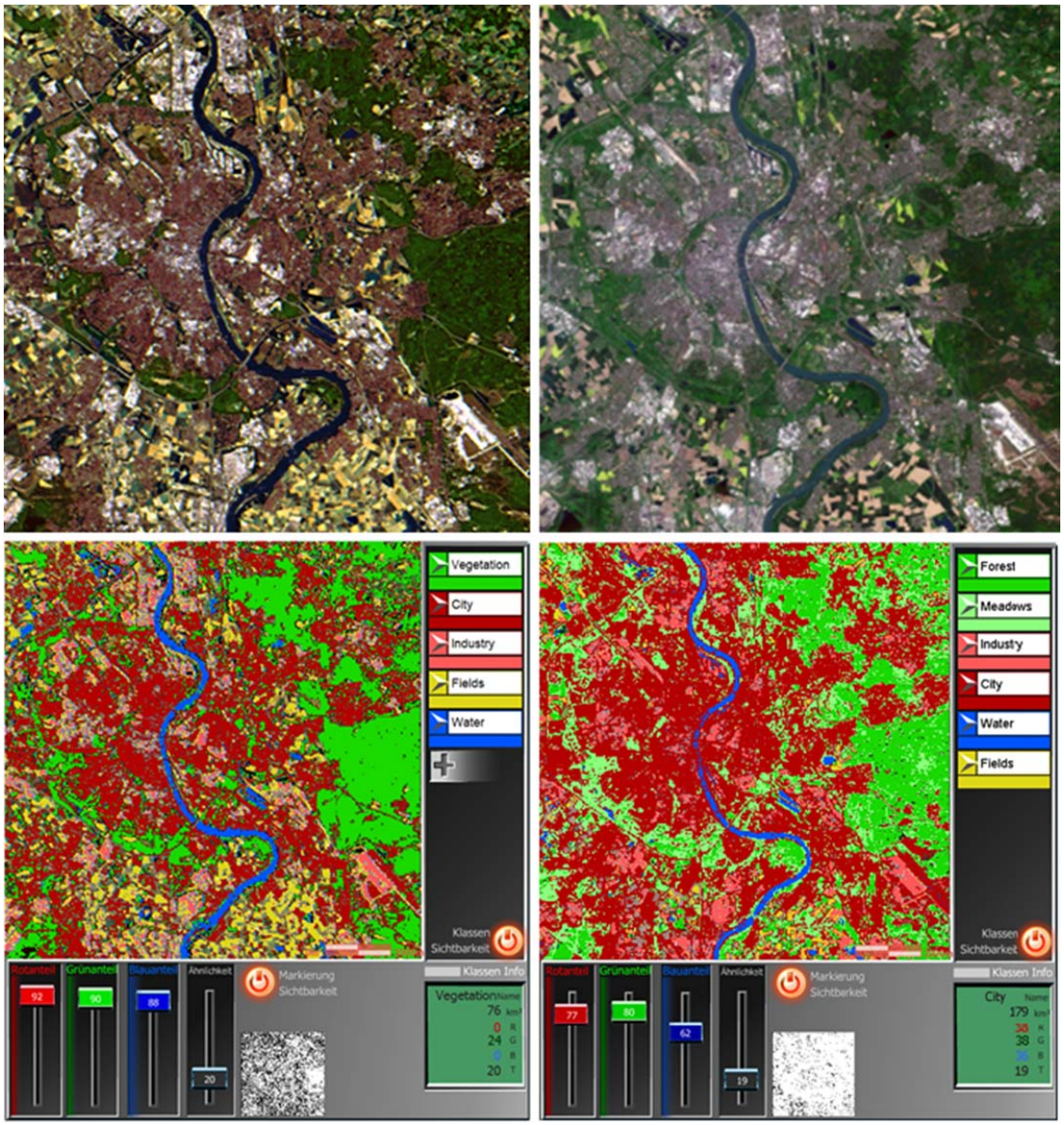


Abb. 3: Vergleich des alten Landsat-RGB-Komposits (links oben) mit dem aktuellen Sentinel-2-RGB-Komposit (rechts oben) und den darauf basierenden Ergebnissen des Klassifikationstools (jeweils darunter) im Modul „Vom Satellitenbild zur Karte“.

2.4 Modul „1, 0 - Spalte, Reihe, Bild“

Für den Informatikunterricht erklärt dieses Modul den Aufbau von digitalen Bildern als Rasterdaten. Die SuS sollen Binäre Zahlen in Form von Bits und Bytes anwenden sowie Histogramme erklären und Bilddaten damit manipulieren können. Da Informatik kein Pflichtfach ist und nicht immer ab der gleichen Jahrgangsstufe unterrichtet wird, reicht die Zielgruppe von Jahrgangsstufe 9 bis 12. Der Zeitaufwand beträgt ca. 2 Stunden.

2.4.1 Ablauf der Lerneinheit

Das Modul startet mit einer interaktiven Erklärung zur Entstehung von Satellitenbildern durch passive Aufnahmeverfahren und Datenübermittlung. Ein Graustufen-Luftbild der Freiheitsstatue in New York dient als Anschauungsmaterial. Das auf 1-Bit Bildtiefe reduzierte Raster wird durch manuelle Einstellung der Spalten- und Reihenanzahl in die richtige Form gebracht und die Orientierung mit der Verortung eines Bildausschnitts abgefragt.
Im nächsten Teil kann das gleiche Bild mit 4-Bit und 8-Bit Bildtiefe dargestellt werden. Die SuS können die Binärwerte der einzelnen Pixel in jeder der Bildtiefen abfragen und damit die Grauwerte verändern. Dabei soll festgestellt werden, dass im vorliegenden Bild nicht der gesamt mögliche Wertebereich der Graustufen ausgenutzt wird. In den zugehörigen Aufgaben und Infotexten wird der Zusammenhang zwischen Dezimal- und Binärwerten durch Umrechnungen weiter verdeutlicht. In einer Infobox werden Häufigkeitsverteilungen und ihre Darstellung in Histogrammen vermittelt, bevor im interaktiven Tool die Möglichkeit besteht, am Beispielbild ein Stretching durchzuführen und den Kontrast des Bildes somit zu verbessern oder zu verschlechtern.

2.4.2 Einbindung der Sentinel-2-Daten

Das im Modul verwendete Luftbild war zwar sehr hochauflösend, aufgrund der Perspektive jedoch für die SuS nur sehr schwer als das Abgebildete zu erkennen und damit korrekt zu verorten. Die hohe Auflösung spielte zudem für die Aufgabenstellung keine Rolle. Es wurde daher entschieden, ein neues, leichter zu verortendes Bild von Sentinel-2 zu verwenden. Das neue Beispielbild zeigt Manhattan auf Höhe des Central Parks. Ein Graustufen-Bild aus dem grünen Kanal in 10m Auflösung wurde aufgrund der guten Kontrasteigenschaften innerhalb des Parkareals ausgewählt. Das verbesserte Modul befindet sich zurzeit in Revision und wird in Kürze veröffentlicht.

2.5 Modul „Der Wald als Klimaretter?!“

Das Modul beschäftigt sich im Kontext der Klimawandelforschung mit Infrarot-Satellitenbildern, aus denen die SuS Waldflächen in den verschiedenen Jahreszeiten ableiten und ihre Bedeutung für den Kohlenstoffkreislauf ableiten sollen. Zielgruppe sind die Jahrgangsstufen 7-9 mit einem Zeitaufwand von ca. 2 Stunden.

2.5.1 Ablauf der Lerneinheit

Das Modul beginnt mit Lesematerial, dem die SuS Hintergrundinformationen zum Thema Klimawandel durch Anstieg des anthropogenen Treibhausausstoßes, v.a. Kohlenstoffdioxid, entnehmen. Es geht um die Frage des Verhältnisses zwischen der von deutschen Wäldern zum

Wachsen benötigten Menge Kohlenstoffdioxids und der anthropogen emittierten Menge. Hierzu wird den SuS der Kohlenstoff-Kreislauf auf interaktive Weise kommuniziert, beginnend mit der Gewinnung und Verbrennung fossiler Treibstoffe. Dabei geht es auch um die Speicherkapazitäten eines einzelnen Baumes gegenüber einem ganzen Wald.
Nach der theoretischen Vorbereitung erhalten die SuS ein Infrarotbild Deutschlands in Graustufen, das sie interpretieren und auf dem sie sich räumlich orientieren. Das hier erlangte Wissen wenden sie nun auf Infrarotbilder des gleichen Raums in allen 4 Jahreszeiten an. In einer Infobox stehen zusätzliche Informationen und Beispiele zur Verfügung: Als Ausschnitt kann der Harz und Umgebung in sichtbarem RGB, Graustufen-Infrarot und als Landnutzungskarte betrachtet werden. Mit einem Analysetool, dass alle Flächen mit einem Wert ähnlich dem ausgewählten markiert, können die SuS die Waldflächen Deutschlands markieren und eine grobe Flächenabschätzung durchführen. Diese reicht aus für die anschließende Bewertung und Diskussion der Frage nach dem Verhältnis von Kohlenstoffbindung und -ausstoß.

2.5.2 Einbindung der Sentinel-2-Daten

Das Zusatzmaterial der Infobox, das den RGB- vs. IR- vs. LULC-Vergleich enthält, wurde mit aktuellen Sentinel-2-Daten aufgewertet. Vier Szenen des Harzes, jeweils im Abstand von 3 Monaten aufgenommen, wurden verarbeitet. Bei der Entwicklung des Moduls konnten aufgrund der häufigen Bewölkung und der nur mittelmäßigen Wiederkehrperiode der Landsat-Satelliten keine echten Herbst- und Winterszenen mit ausreichend geringer Wolkenbedeckung genutzt werden. Dank der hohen Wiederkehrfrequenz der Sentinel-2-Satelliten konnten in der Region nun nahezu wolkenfreie Szenen in gleichmäßigen zeitlichen Abständen gewonnen werden. Das verbesserte Modul befindet sich zurzeit in Revision und wird in Kürze veröffentlicht.

2.6 Mini-MOOCs für Themen der Erdbeobachtung

Bei der Vermittlung von Themen der Erdbeobachtung folgte das Projekt in den vergangenen Jahren maßgeblich dem Ansatz, den SuS über Interaktion und Simulation den Umgang mit Fernerkundungsdaten näher zu bringen. Diese bislang erfolgreiche Strategie soll in der laufenden Projektphase um die Möglichkeiten erweitert werden, welche MOOCs (Massive Open Online Courses) den Lernenden bieten, und welche Videos als hauptsächliches Lernmedium einsetzen.

2.6.1 Struktur und Medien

Eine Lerneinheit ist in inhaltlich abgeschlossene Kapitel eingeteilt. Jedes der Kapitel ist als aufbauendes Kapitel zum vorherigen einsetzbar, von diesem jedoch nicht zwangsläufig semantisch abhängig (vgl. Abb. 4).

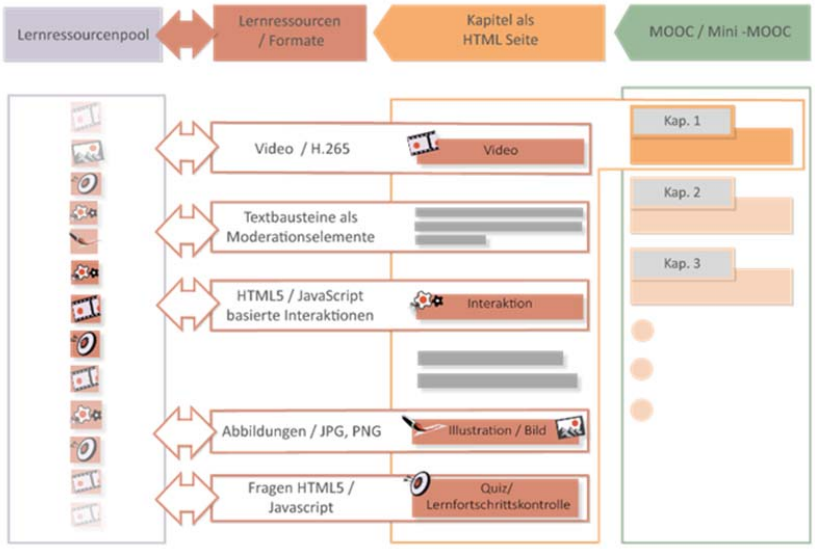


Abb. 4: Struktur der Lerneinheiten

Um diese inhaltliche Modularität zu gewährleisten, wird auf Bezüge zwischen den Lernressourcen verzichtet. Diese könnte lediglich durch Textbausteine erfolgen, welche die Videos und Interaktionen miteinander verbinden und mehr der Moderation als der Inhaltsvermittlung dienen. Die Inhaltsvermittlung soll sich auf die verwendeten Lernressourcen Video, Interaktion und Illustration beschränken (vgl. Abb. 4).
Jedes Kapitel ist als einzelnes HTML5-Dokument verfügbar und steht so für eine erneute Sequenzierung in einem anderen thematischen Zusammenhang zu Verfügung. Die Kapitel werden durch Tests zur Selbstkontrolle des Lernfortschrittes abgeschlossen. Die Ergebnisse dieser Tests können bei Bedarf auch dem Lehrer als Lernfortschrittskontrolle dienen.
Jedes erstellte Lernmaterial, ob Video, Interaktion oder Selbstkontrolle, wird in einen Lernressourcen-Pool (vgl. Abb. 5) eingebracht, der in die Bereiche Bildstruktur, Auflösung, Bildvorverarbeitung, naturwissenschaftliche Grundlagen, Methoden und Systeme eingeteilt ist. Damit können sie entweder für nachfolgende Mini-MOOCs weiterverwendet, oder auch in einem anderen Zusammenhang in der Lehre eigesetzt werden.

2.6.2 Themen

Als thematische Klammer für die Einheiten wird die Beobachtung und Auswertung des globalen Wandels und klimatischer Veränderungen dienen. Eine besondere Rolle wird den Copernicus-Kerndiensten zukommen. Fernerkundliche Methoden und Anwendungsgebiete bilden dabei den Ausgangspunkt für die mediale Aufbereitung jedes Mini-MOOCs. So können Aspekte des globalen Wandels anhand von Fernerkundungsdaten illustriert und erklärt sowie Grundlagen zum Verständnis von Satellitendaten geschaffen werden (vgl. Abb. 5).

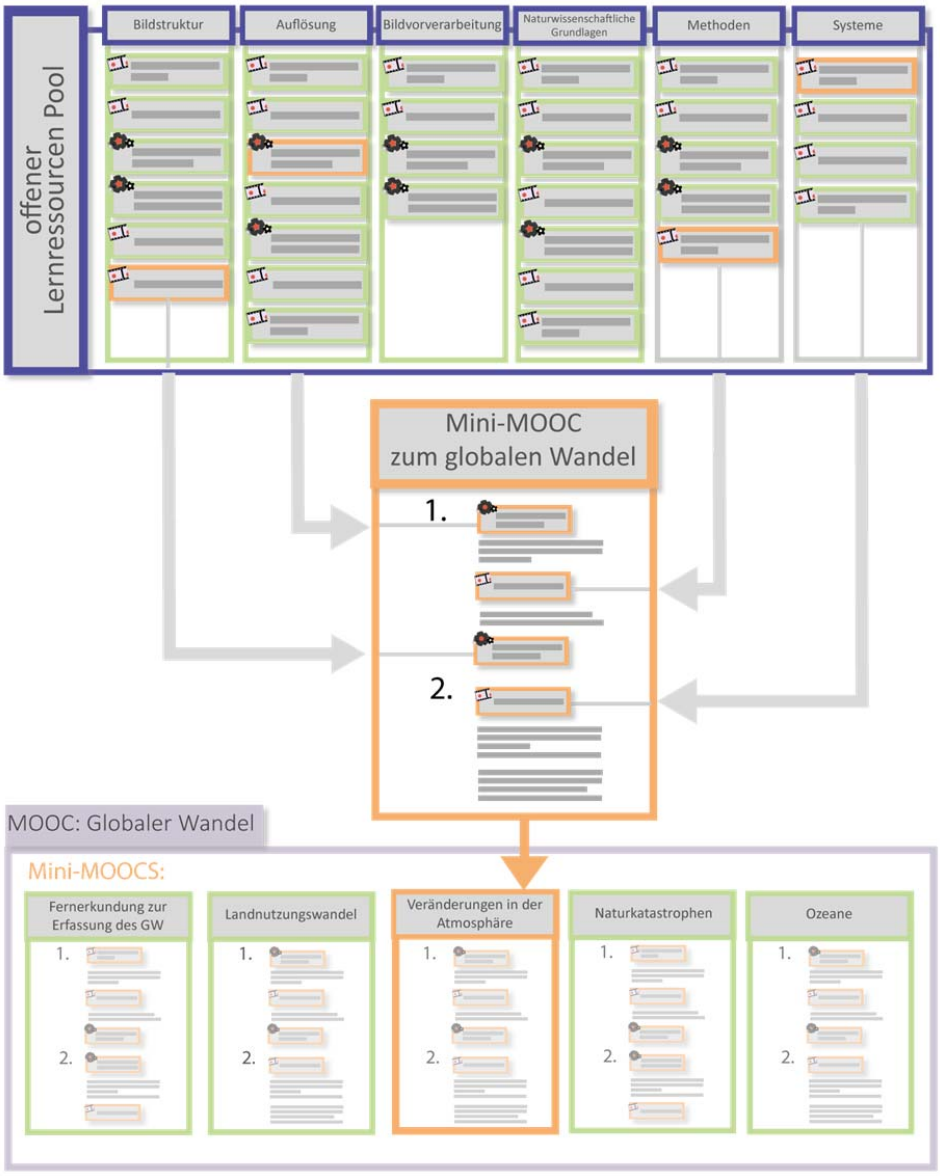


Abb. 5: Themen der geplanten Mini-MOOCs und Einbindung der Lernressourcen.

3 Fazit & Ausblick

Im Projekt „Fernerkundung in Schulen“ (FIS) wurden 21 Lernmodule entwickelt, mit denen SuS der Mittel- und Oberstufe selbstständig, interaktiv, interdisziplinär, intermedial und propädeutisch Wissen über Mensch-Umwelt-Interaktionen anhand von Fernerkundungsdaten und –Methoden erlangen und anwenden können. Von den vorhandenen Lernmodulen wurden fünf mit aktuellen Daten von Sentinel-2 überarbeitet und aufgebessert. Weitere Module werden nach Notwendigkeit um neue Sentinel-Daten erweitert. Dazu werden auch Sentinel-1-Radardaten und thermale Infrarot-Daten von Sentinel-3 gehören.

Die Lernmodule sind momentan nur im Flash-Format verfügbar, was Einschränkungen im Ressourcenverbrauch und in der Wartung bedeutet. In naher Zukunft wird eine Übertragung der Module in das HTML5/JavaScript-Format und gleichzeitige Anpassung an moderne Didaktik-Standards durchgeführt. Es liegt inhaltliches, technisches und didaktisches Feedback von LuL und SuS aus zehn Jahren Praxis vor, das in die Überarbeitung der alten und die Entwicklung neuer Module einfließt.

4 Literaturverzeichnis

DGFG DEUTSCHE GESELLSCHAFT FÜR GEOGRAPHIE (Hrsg.), 2014: Bildungsstandards im Fach Geographie für den Mittleren Schulabschluss mit Aufgabenbeispielen. Bonn.

GOETZKE, R., HODAM, H., RIENOW, A. & VOSS, K., 2014: Teaching Materials, Encyclopedia, Easy-To-Use Image Processing – The FIS Learning Portal on Remote Sensing. In: EARSeL eProceedings, **12**(2), 164-173.

NCGE NATIONAL COUNCIL FOR GEOGRAPHIC EDUCATION (Hrsg.), 2012: Geography for Life. 2nd Edition National geography Standards.

RIENOW, A., HODAM, H., MENZ, G. & VOSS, K., 2015: Wissensvermittlung aus der Vogelperspektive – Fernerkundung und Raumfahrt in Schulen. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 24, T. Kersten (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 16.-18. März 2015 an der Universität zu Köln, Tagungsband.

VOSS, K., GOETZKE, R., HODAM, H., & RIENOW, A., 2011: Remote Sensing, New Media and Scientific Literacy – A New Integrated Learning Portal for Schools Using Satellite Images. In: Jekel, T.; Koller, A.; Donert, K. & Vogler, R. (Hrsg.): Learning with GI 2011- Implementing Digital Earth in Education. Wichmann, Berlin, 172-180.

Ableitung von 3D-Modellen aus Daten des High Definition Earth Viewing-Experiments (ISS) – Anwendungen für den Schulunterricht

JOHANNES A. SCHULTZ[1], HENRYK HODAM[1], CLAUDIA LINDNER[1], ANNETTE ORTWEIN[1], FABIAN SELG[1] & ANDREAS RIENOW[1]

Zusammenfassung: „KEPLER ISS“ ist das Nachfolgeprojekt von „Columbus Eye – Live Bilder von der ISS im Schulunterricht“ und wird vom Deutschen Luft- und Raumfahrtzentrum (DLR) gefördert (Förderkennzeichen 50JR1701). Ziel ist es, nachhaltig die Fernerkundung von der ISS in den Schulunterricht zu integrieren. Das Projekt ist exklusiver europäischer Partner des “High Definition Earth Viewing Experiment“ (HDEV) der NASA. Im Rahmen dieses Experiments wird die Erde mit HD-Kameras rund um die Uhr von der Internationalen Raumstation (ISS) aus beobachtet. Aus diesen Videos entwickelt das Projekt Materialien für den Schulunterricht. Der Schwerpunkt des Beitrages liegt auf der Erstellung von 3D-Modellen aus Erdbeobachtungsdaten von der ISS.

1 Einleitung

Es ist relativ schwierig, Schülerinnen und Schüler (SuS) nachhaltig für die Naturwissenschaft zu begeistern. Hierfür ist es sinnvoll, Medien, Arbeitsmaterialien und Lernwerkzeuge zu entwickeln, die den Forschergeist und die Faszination für die Naturwissenschaft wecken.
Das „High Definition Earth Viewing“ (HDEV) Experiment der NASA (RUNCO 2015) bietet die einzigartige Möglichkeit, die Faszination für Raumfahrt und für die Erforschung des Weltraumes dazu zu verwenden, die SuS für die naturwissenschaftlichen Schulfächer zu begeistern. HDEV wurde im April 2014 am external payload adapter des Columbus-Labors der ISS befestigt. Die Kameras des Experiments beobachten die Erde rund um die Uhr aus verschiedenen Blickwinkeln.
Das primäre Ziel von „Columbus Eye-Live Bilder von der ISS im Schulunterricht“ bzw. vom Nachfolgeprojekt KEPLER ISS ist es, basierend auf den Videos des NASA HDEV-Experiments Unterrichtsmaterialien zu entwickeln und das Videomaterial für die Bedürfnisse des Schulunterrichts aufzubereiten.
Hierzu betreiben die Arbeitsgruppe Fernerkundung der Universität Bonn und die Arbeitsgruppe Geomatik der Ruhr-Universität Bochum gemeinsam ein umfangreiches Webportal (www.columbuseye.uni-bonn.de), welches sowohl Highlight-Videos, als auch interaktive Lernmodule, Apps und Arbeitsblätter für die SuS und LuL zur Verfügung stellt (RIENOW et al. 2015a, 2016). Die interaktiven Unterrichtsmaterialien und auch die HDEV-Videos werden hauptsächlich in den Schulfächern Geographie, Physik, Biologie, Informatik und Mathematik verwendet.
Über ein Web-GIS besteht die Möglichkeit, die Highlight-Videos sowohl thematisch zu kategorisieren als auch räumlich zu verorten. Die Videos des NASA HDEV-Experiments und Astro-

[1] Ruhr-Universität Bochum, Arbeitsgruppe Geomatik, Universitätsstraße 104, D-44799 Bochum, E-Mail: johannes.schultz@rub.de

nautenbilder unserer Erde ermöglichen den SuS einen faszinierenden Blick aus dem All auf die natürlichen Prozesse, Phänomene und die auftretenden Wechselwirkungen in einem gekoppelten Mensch-Umweltsystem (VOß et al. 2010).
Im Folgenden wird zunächst kurz das HDEV-Experiment vorgestellt und auch auf die technischen Anforderungen für Empfang, Speicherung, Georeferenzierung und Atmosphärenkorrektur der Daten eingegangen.
Der Beitrag stellt, neben den Anwendungen für den Schulunterricht, auch die Potenziale und Anwendungsmöglichkeiten von Videodaten und Astronautenbildern in Bezug auf die Erdbeobachtung vor. In diesem Zusammenhang liegt der Schwerpunkt der Betrachtung auf der Erstellung von 3D-Modellen aus Erdbeobachtungsdaten von der ISS und deren Verwendung im Schulunterricht. Im Rahmen des Projektes werden verschiedene Lernmaterialien, z. B. Arbeitsblätter, Augmented Reality Apps (erweiterter Realität) oder computergestützte Lernmodule erstellt (ORTWEIN et al. 2017). Exemplarisch hierfür stellt der Beitrag neue Lerneinheiten, die auf Augmented Reality basieren, vor.
Neben diesen wissenschaftlichen und für die Schul-Curricula relevanten Anwendungen wird gezeigt, wie die Daten im Rahmen eines Schulwettbewerbs eingesetzt werden können. Die SuS sollen im Rahmen dieses Wettbewerbes lernen, wie sie interdisziplinäre Projektideen eigenständig auf der Schnittstelle zwischen Kunst, Naturwissenschaft und Ingenieurwesen entwickeln und umsetzen können.
Im Ausblick des Beitrages wird ein neues auf der ISS geplantes thermales Sensorsystem, welches mit einem sehr hochauflösenden Kamerasystem kombiniert ist, vorgestellt.

2 Das HDEV-Experiment

Im April 2014 wurden am Columbus-Modul der ISS vier Standard-Videokameras (commercial off-the shelf, COTS) mit Unterstützung des kanadischen Roboterarms installiert und in Betrieb genommen (RUNCO 2015). Die Kameras nehmen die Erde aus verschiedenen Blickwinkeln rund um die Uhr auf. Die Toshiba Kamera IK-HR1S schaut nadir, während die Hitachi GVHD30 nach vorne schaut. Zwei Kameras, Panasonic AGHMC150 und Sony FCBEH4300, schauen nach hinten (vgl. Abb. 1). Die räumliche Auflösung beträgt für die Nadir-Kamera ungefähr 500 m. Ziel des NASA HDEV-Experiments ist es, verschiedene Standardkameras zu vergleichen und zu überprüfen, ob kommerzielle Videokameras für zukünftige Langzeitmissionen zum Mond und Mars, insbesondere in Bezug auf die auftretende Strahlung, geeignet sind (RUNCO 2015, MURI et al., 2017). Die Kameras wurden in einem stickstoffgefüllten Gehäuse (siehe Abb. 1), welches sie auch vor den starken Temperaturschwankungen schützt, an der Außenseite des Columbus-Moduls der ISS angebracht.
Die ISS bewegt sich in einer durchschnittlichen Höhe von 400 km, jedoch ist die Flugbahn im Vergleich zu erdbeobachtenden Satellitensystemen, sowohl in Bezug auf die Flughöhe, als auch in Bezug auf die Lage der ISS, hoch variabel (STEFANOV et al. 2011). Bei Andockmanövern wird die ISS oft gedreht, die Blickrichtung der Kamera ändert sich somit auch. Darüber hinaus ist im Rahmen des HDEV-Experimentes kein Sternensensor vorhanden und da die Aufnahmezeit nicht mitgespeichert wird, ist eine automatische Georeferenzierung der Videos nicht möglich.

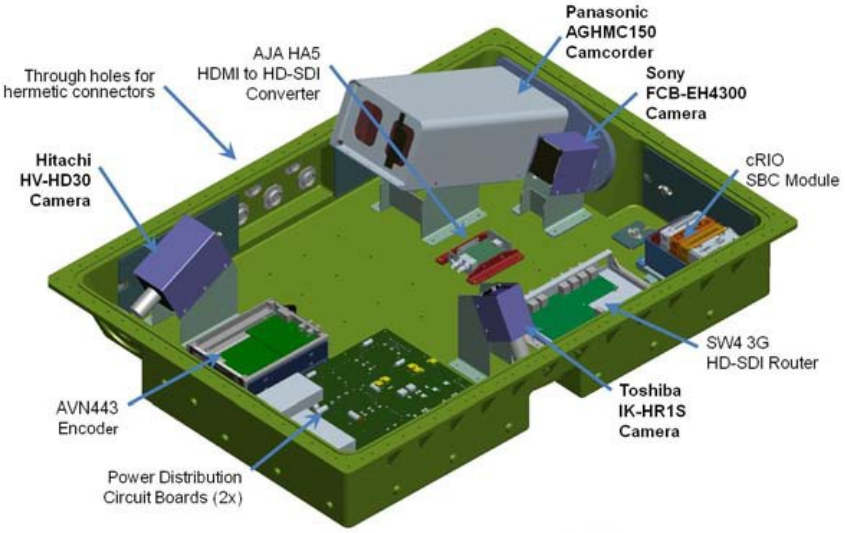


Abb. 1: Anordnung der Videokameras des NASA HDEV-Experiments. Quelle: (MURI et al. 2017)

Die Kameras sind nicht alle gleichzeitig aktiv, sondern ein Kamerazyklus legt fest, welche Kamera gerade ausgewählt wird. Dieser Kamerazyklus kann aber über TReK (Telescience Resource KiT) beliebig modifiziert werden. Die Signale der aktiven Kamera werden über die NASA Tracking und Data Relay Satelliten (TDRS) zum Bodensegment in White Sands, New Mexico und von dort aus an das Payload Operations Integration Center (POIC) am Marshall Spaceflight Center in Huntsville, Texas weitergeleitet. (RUNCO 2015)

Von dort aus werden sie direkt an den Columbus Eye-Server übertragen. KEPLER ISS ist der exklusive europäischer Partner von HDEV und ist auch für die Archivierung und Bearbeitung der Videos zuständig (RIENOW et al. 2016). Da die Daten über KU-Band übertragen werden ist ein kontinuierlicher Empfang nicht immer möglich. Dies ist darin begründet, dass auch andere Experimente das KU-Band verwenden. Durch die geringe Lichtempfindlichkeit der Kameras sind keine Aufnahmen in der Nacht möglich.

Neben dem Videoarchiv, welches derzeit ungefähr 40 Terabyte umfasst, werden auch Astronautenbilder und erdbeobachtende Sensoren der ISS, wie z. B. Hyperspectral Imager for the Coastal Ocean (HREP-HICO) (GAO, &. LI, 2010) verwendet, um Unterrichtseinheiten für den Schulunterricht zu entwickeln**.**

3 Lernen mit Erdbeobachtungsdaten von der ISS

Alle Unterrichtsmaterialien der Projekte Columbus Eye und Kepler ISS werden in einem Online-Lernportal der breiten Öffentlichkeit zur Verfügung gestellt (www.columbuseye.uni-bonn.de) und orientieren sich an den Schul-Curricula der Länder (RIENOW 2015b).

Im Portal steht neben den Unterrichtsmaterialien auch ein Livestream von der ISS zur Verfügung. Die Nutzer haben die Möglichkeit, in einem Web-GIS die aktuelle Position der ISS auf der Weltkarte zu verfolgen. Die Seite stellt auch Hintergrundinformationen zu ISS und zu den Missionen von Alexander Gerst „Blue Dot und „horizons“ zur Verfügung. Das auf der Seite befindliche frei zugängliche Videoarchiv umfasst mehr als 200 Videos. Diese Highlights zeigen spek-

takuläre Aufnahmen von der ISS, z. B. Algenblüten im Ozean, Waldbrände, Städte oder Wetterphänomene wie Hurrikans, Sandstürme oder Schwerewellen.
Durch die atmosphärische Streuung (Mie- und Rayleigh Streuung) zeigen die Highlight-Videos nicht die natürlichen Farben. Ferner sind die Kameras fest eingestellt, sodass sie nicht auf veränderte Belichtungsverhältnisse reagieren können. Dies macht es erforderlich, das Rohvideomaterial entsprechend aufzubereiten. In MATLAB werden hierzu entsprechende Algorithmen zur Anpassung von Helligkeit, Kontrast und Farbe verwendet (SCHULTZ et al. 2017). Für alle Highlight-Videos gibt es auch einen Hintergrundtext und über ein Web-GIS besteht die Möglichkeit, die Videos räumlich und thematisch zu filtern.
Die auf dem Lernportal präsentierten Videos und Hintergrundinformationen sind verbunden mit zahlreichen Lernmaterialien, die ein breites Spektrum von computergestützten Lernmodulen, analogen Arbeitsblättern bis hin zu Smartphone basierten Apps abdecken (ORTWEIN et al. 2017). Aus einfachen analogen Arbeitsblättern werden durch Augmented Reality virtuelle Lernumgebungen. Das didaktische Konzept aller Lernmaterialien basiert auf dem moderaten Konstruktivismus und auf den Erkenntnissen, die mit dem vom DLR geförderten Projekt „Fernerkundung in Schulen“ (FIS) (www.fis.uni-bonn.de) gemacht worden sind (VOß et al. 2011).
Im Folgenden wird auf das Arbeitsblatt Stereoskopie bzw. auf die Erstellung von 3D-Modellen eingegangen. Im Anschluss daran wird kurz die Augmented Reality-App und der Schülerwettbewerb vorgestellt.

3.1 3D-Modelle aus ISS- Erdbeobachtungs-Daten

Die im Rahmen des Projektes Columbus Eye entwickelte Lerneinheit „Stereoskopie“ verwendet sowohl Astronautenbilder der Erde als auch Videomaterial vom NASA HDEV-Experiment für die Erstellung von stereoskopischen Aufnahmen. Darüber hinaus kamen auch 3D-Videos aus dem Innenraum der ISS in der Einheit zur Anwendung. Das übergeordnete Ziel ist es, den SuS die physikalischen Grundlagen der Stereoskopie zu vermitteln und ihnen das Prinzip des räumlichen Sehens zu erläutern. Die Einheit zeigt, wie man aus zweidimensionalen Bildern dreidimensionale Raumeindrücke erzeugen kann.
Die 3D-Modelle wurden mit Agisoft PhotoScan© erstellt. Hierbei hat sich herausgestellt, dass insbesondere bei Landschaften mit hoher Reliefenergie bzw. mit sehr großen Höhenunterschieden erhebliche Fehler in den Modellen entstehen können (SCHULTZ et al. 2017). Die Minimierung von Artefakten im 3D-Modell ist insbesondere bedeutend für die Ableitung von 3D-Modellen aus HDEV-Daten. Hierzu wurden exemplarisch hochauflösende Astronautenbilder vom 3.776 m hohen Fuji verwendet. Für den Fuji stehen sehr viele Aufnahmen im „Gateway To Astronaut Photography of Earth“ (Earth Science and Remote Sensing Unit, NASA Johnson Space Center) zur Verfügung. Allein am 8.2.2016 wurden 32 Bilder vom Fuji von der ISS aufgenommen, nur 7 davon wurden für die Erstellung des 3D-Modells verwendet (siehe Abb. 2).
Insbesondere von Vorteil war der sehr geringe zeitliche Abstand zwischen den Bildern, die leicht unterschiedlichen Aufnahmewinkel, die hohe Auflösung (4.928×3.280 Pixel) und die, durch die Verwendung eines 1.150 mm Objektives, kleine Ground Sampling Distance (GSD) von ≈ 4m. Die großen Höhenunterschiede von über 2.600 m, reliefbedingte Schatten, die Bewölkung am Fuß des Fuji, sowie die fehlenden Informationen in Bezug auf Aufnahmewinkel, Flughöhe und Position,

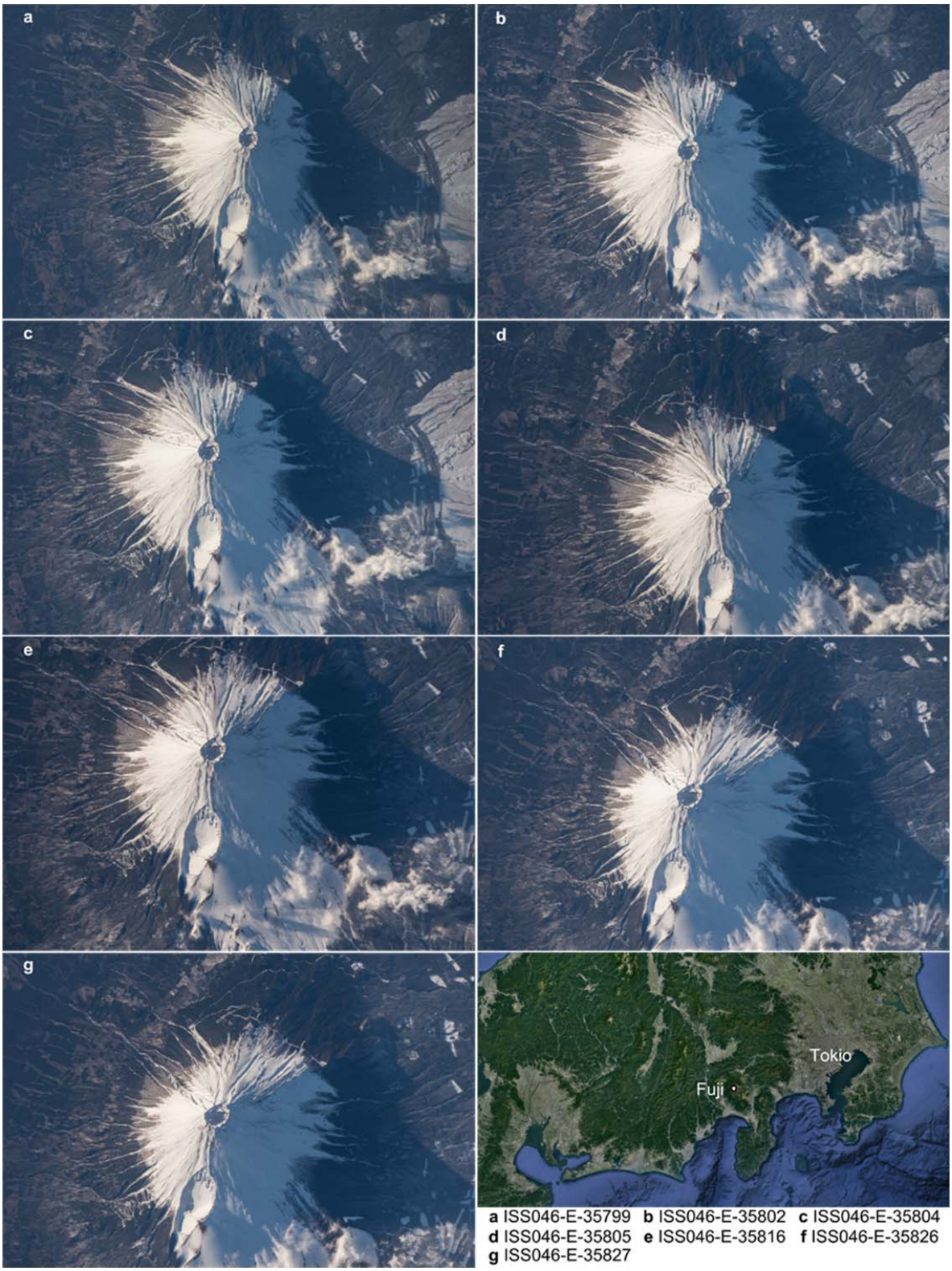


Abb. 2: Der westlich von Tokio gelegene Vulkan Fuji von der ISS aus betrachtet. ISS Bild ID siehe Abbildung. Quelle: Earth Science and Remote Sensing Unit, NASA Johnson Space Center

machen es jedoch schwer, ein georeferenziertes 3D-Modell zu erstellen. Von den 32 Bildern war ein Minimum von 7 Bildern notwendig um ein „visuell ansprechendes“ 3D-Modell zu erzeugen. Im Fuji-Modell traten jedoch relativ viele Fehler auf; diese konnten durch die Verwendung von 5 manuell gesetzten Markierungspunkten deutlich reduziert werden. In Abbildung 3 ist eines der sieben Originalbilder mit blauen Referenzpunkten zu sehen. Diese werden basierend auf den manuell ausgewählten Markierungspunkten (grüne Fähnchen) automatisch vom Programm gesetzt (AGISOFT 2018). Die manuelle Erstellung von Punkten war in diesem Beispiel von geringem Aufwand, da die Gesamtzahl nur 35 (7×5) beträgt. Das Programm errechnet basierend auf der Überlappung der Bilder und der unterschiedlichen Aufnahmewinkel die genaue Position der Kamera. Aus diesen Daten werden eine 3D-Punktwolke und ein Mesh berechnet. Besonders Bildbereiche mit Wolken führen oft zu Fehlern (Abb. 4 Punktwolke).

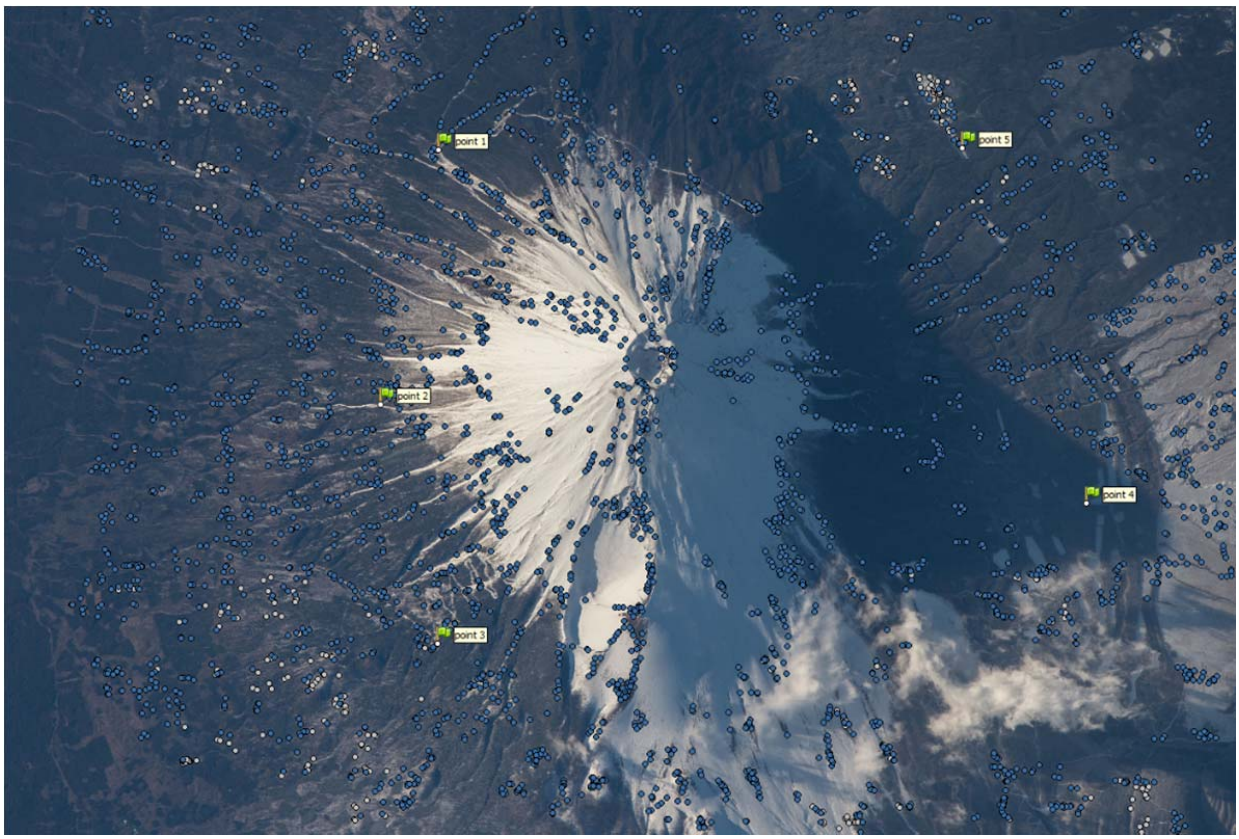


Abb. 3: Astronautenbild vom Fuji mit Referenzpunkten

Auf dieses Mesh werden dann die einzelnen Bilder projiziert und daraus die Textur erzeugt. Die Ergebnisse können als einfaches, frei drehbares 3D-Model (Abb. 5) dargestellt werden; ergänzend können auch Anaglyphenbilder (MICHEL 2013), die einen dreidimensionalen Seheindruck vermitteln, erzeugt werden. Das vom Fuji erstellte 3D-Modell kann auch als animiertes PDF aus Agisoft PhotoScan© exportiert werden.

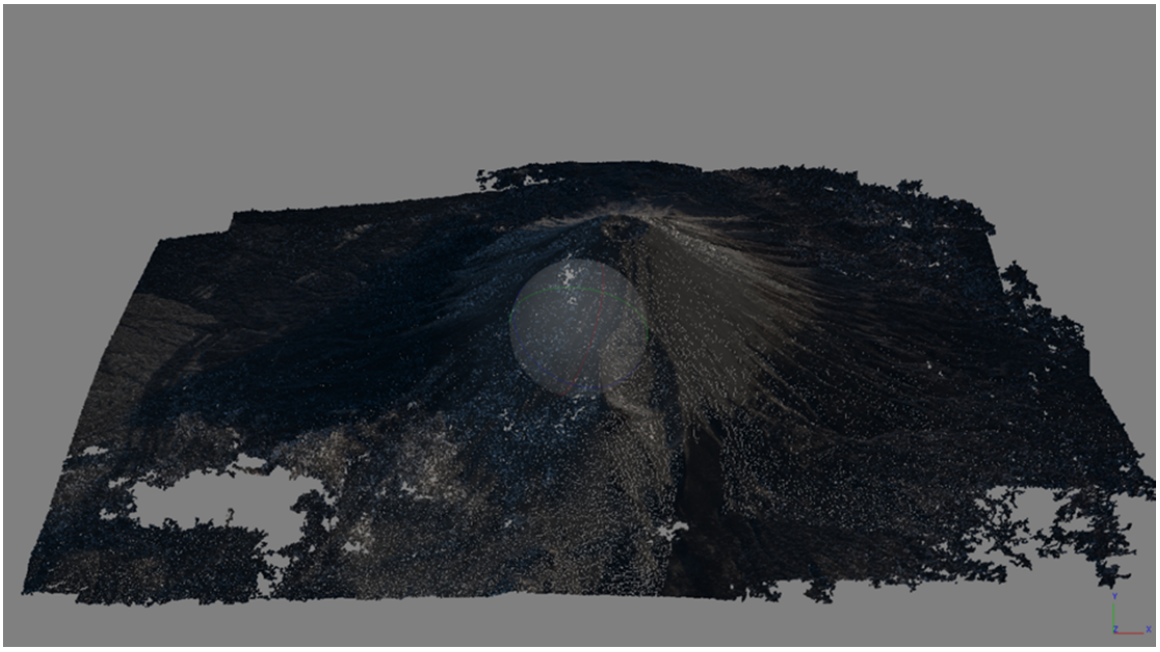


Abb. 4: 3-D Punktwolke

Neben den Astronautenbildern können auch aus HDEV-Videos 3D-Modelle abgeleitet werden (SCHULTZ et al. 2017). Dazu ist es aber zunächst erforderlich, die Videos aufzubereiten.

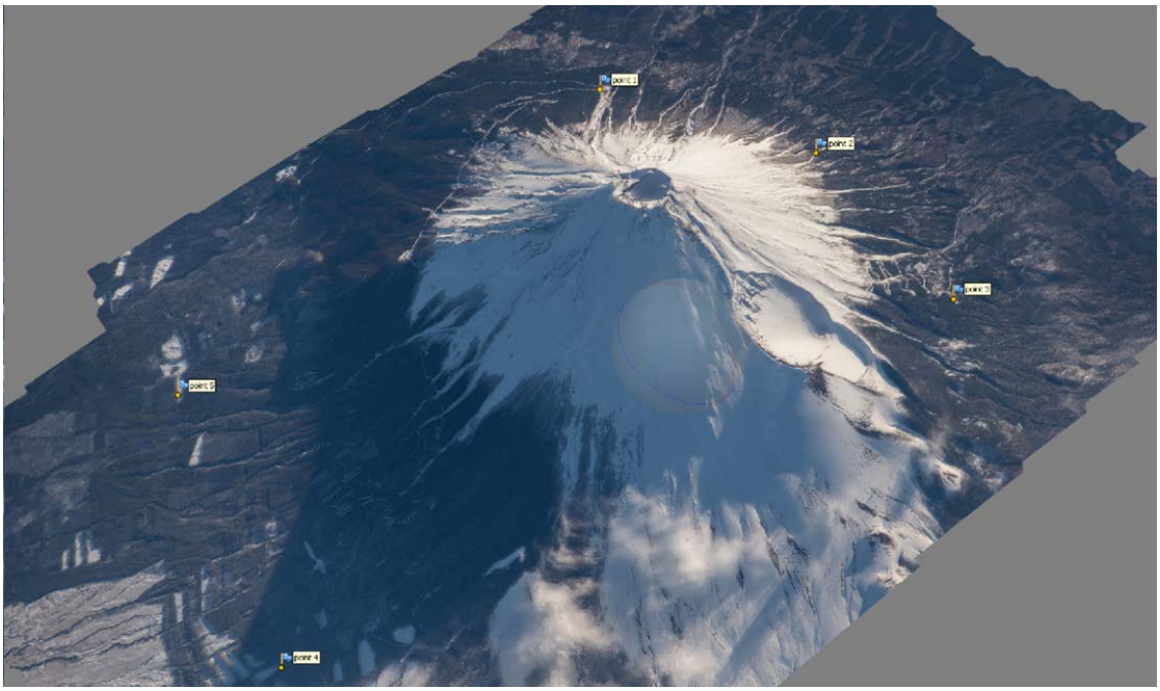


Abb. 5: 3D-Modell vom Fuji, abgeleitet aus Astronautenbildern von der ISS

Insbesondere müssen atmosphärische Einflüsse und der Kontrast verbessert werden. Hierfür werden entsprechende Algorithmen in MATLAB© verwendet (SCHULTZ et al. 2017). Im Anschluss daran werden die Videos in MATLAB in Einzelbilder zerlegt und in Agisoft Photo-

Scan© eingeladen. Die Kameras nehmen die Erde mit einer Framerate von 60 Bilder pro Sekunde auf.
Trotz der hohen Geschwindigkeit der ISS von ungefähr 7.600m/s ist es vollkommen ausreichend, nur jedes zwanzigste Bild für die Erstellung des 3D-Models zu verwenden, da die Überlappung dann immer noch weit über 95 % beträgt.
Bei der Generierung von 3D-Modellen aus HDEV-Daten haben sich vier Aspekte als besonders schwierig erwiesen:

- die geringe Auflösung der HDEV-Videos, die selbst im Nadir nur bei ca. 500 m liegt,
- Wolken, die oft zum Auftreten von Artefakten in 3D-Modellen führen,
- sich stark ändernde Beleuchtungsbedingungen, wenn sehr große Gebiete abgedeckt werden,
- statische Bildelemente wie Sonnensegel, die immer aus demselben Kamerawinkel betrachtet werden und aus denen somit kein stereoskopisches Bild berechnet werden kann.

Statische Bildelemente können den dreidimensionalen Seheindruck der gesamten Szene negativ beeinflussen. Idealerweise sollten deswegen solche Bereiche aus den einzelnen Bildern herausgeschnitten werden.
Der nächste Schritt ist es, gewonnene Erkenntnisse aus dem 3D-Fuji-Modell (Abb. 5) auf die HDEV-Videos anzuwenden. Bisher wurden aus HDEV-Videos nur 3D-Modelle ohne die Verwendung von manuell gesetzten Markierungspunkten abgeleitet. Für die Visualisierung im Schulunterricht ist die Auflösung von HDEV vollkommen ausreichend. Auf der ISS sind auch schon höher auflösende Kameras und Sensoren im Einsatz und in der Planung, somit erlangt die ISS auch eine immer größer werdende Bedeutung für die Erdbeobachtung (STEFANOV & EVANS 2014, SCHULTZ et al. 2017).

3.2 Augmented Reality Apps

E-Learning ist, sowohl an den Hochschulen, als auch an Schulen, kein neues Konzept, jedoch können interaktive Lernmaterialien den SuS helfen, autodidaktische Kompetenzen zu entwickeln, und ihre Fähigkeiten zur Selbstorganisation stärken z. B. durch computergestützte Lerneinheiten oder durch die selbstständige Internetrecherche (VOß et al. 2011). Relativ neu hingegen ist jedoch der Einsatz von Augmented Reality (AR, erweiterte Realität), mit der analoge Lernmaterialien z. B. durch Smartphone Apps um die virtuelle Welt erweitert werden können. Smartphones bzw. m-learning (KOROCU & ALKAN 2011) Applikationen haben den großen Vorteil gegenüber Computern, dass sie in den Klassenräumen nahezu ubiquitär vorhanden sind und die SuS sehr leicht zu begeistern sind, neue Apps auszuprobieren (NMC 2017, WEISER 1991). AR basiert auf der Aufnahme von Echtzeit-Kamerabildern, die durch eine Applikation erkannt und mit hinterlegten Zusatzinformationen auf dem Display überlagert werden (HÖLLERER & FEINER 2004, DUNLEAVY et al. 2009).
Im Rahmen von Columbus Eye und KEPLER-ISS wurden bisher zwei Arbeitsblätter, die sich auch Augmented Reality (HÖLLERER & FEINER 2004) zu Nutze machen namens „Im Auge des Sturms“ und „Aralkum – vom See zu Wüste“, entwickelt. Die SuS nehmen mit ihrem Smartphone die im Arbeitsblatt platzierten Marker, z. B. ein Satellitenbild vom Aralsee, auf. In ihrem Display wird der Marker durch eine interaktive Zeitreihe von Satellitenbildern (hier Landsat und Sentinel) ersetzt, mit der sie das Schrumpfen des Aralsees anschaulich nachvollziehen können.

Durch die Überlagerung und Bereitstellung von Zusatzinformationen können sie im Arbeitsblatt die verschiedenen Uferlinien einzeichnen und einfache Flächenberechnungen durchführen (Abb. 6). In der App „Im Auge des Sturms“ können die SuS in der Boden- und Höhenwetterkarte die Zugbahn und Zuggeschwindigkeit des Zyklons Maysak nachvollziehen und eigene Prognosen und Berechnungen durchführen.

Abb.6: Aralkum-App

Die Apps wurden mit dem Android Developer© erweitert mit Vuforia© und können über das Columbus Eye-Portal und den Google Play Store© installiert werden. Eine Version für das iPhone ist in Vorbereitung.

3.3 Schülerwettbewerb Self-eSTEAM

Mit dem Schülerwettbewerb Self-eSTEAM ist die Herangehensweise für die Stärkung der MINT Kompetenz bei den SuS eine andere. Hier werden nicht auf Erdbeobachtungsdaten von der ISS basierende Unterrichtsmaterialien für die MINT Fächer entwickelt, sondern die SuS sollen eigene interdisziplinäre Projektideen an der Schnittstelle zwischen Kunst, Naturwissenschaft und Ingenieurwesen entwickeln und umsetzen.
Die Wortschöpfung Self-eSTEAM kombiniert den englischen Ausdruck self-esteem (dt. Selbstwertgefühl) über das e des e-learnings mit der englischen Abkürzung STEAM. Sie entspricht in etwa dem deutschen ‚MINT’ plus Kunst und repräsentiert die naturwissenschaftlichen Fächerschwerpunkte science, technology, engineering, und mathematics in Kombination mit arts. Ziel des Wettbewerbs ist die künstlerische Auseinandersetzung der SuS mit den Naturwissenschaften unter Integration der ISS-Videos und sonstigen Erdbeobachtungsdaten von der ISS. Hierbei können sowohl Kollagen, Filme oder Bilder angefertigt werden. Es besteht auch die Möglichkeit, etwas zu programmieren, z. B. mit einem Raspberry Pie, oder auch politische Fragestellungen mit den Daten zu adressieren. Das Motto „Naturwissenschaft trifft Kunst und Gesellschaft“ kann frei umgesetzt werden. Der Wettbewerb richtet sich an die SuS der Sekundarstufe I, teilnahme-

berechtigt sind ganze Schulklassen, AGs oder einzelne Teams unter der Leitung eines Lehrkörpers. Da der Wettbewerb Alexander Gersts Raumflug im Jahr 2018 begleiten soll, können die eigenständig erstellten Beiträge innerhalb des Zeitraums vom 01.03. bis zum 31.10.2018 eingereicht werden. Im Nachgang werden die Beiträge durch eine Jury begutachtet und prämiert. Die ersten drei Plätze werden zu einem Astronautentreffen eingeladen.

Abb. 7: Teaser des Wettbewerbs Self-eSTEAM

4 Ausblick

Die oben dargelegte Vorgehensweise kann auch auf die vom japanischen Experimentmodul „KIBO" aufgenommenen 4K-Videos angewendet werden. Die hier verwendete Kamera ist extrem lichtempfindlich, sodass sie selbst in der Nacht Aufnahmen von der Erde machen kann. Da die Kameraeinstellungen hier, im Gegensatz zu den HDEV-Kameras, frei wählbar sind, können auch sehr gute Aufnahmen unter schwierigen Beleuchtungsverhältnissen aufgenommen werden.
Ein noch leistungsfähigeres Kamerasystem soll auf der ISS im Rahmen des in der Vorbereitung befindlichen DIEGO-Projektes installiert werden. „DIEGO – Dynamische Infrarot-Erdbeobachtung im ISS-Orbit" ist eine gemeinsame Initiative der Arbeitsgruppe Geomatik der Ruhr-Universität Bochum und der Remote Sensing Research Group (RSRG) der Universität Bonn. DIEGO wird die einzigartigen Vorteile des ISS-Orbits für die Erdbeobachtung nutzen und erstmals räumlich sehr hochauflösende (<60m) multispektrale Daten im sichtbaren Wellenlängenbereich und nahen Infrarot (VNIR) sowie im mittleren (MIR) und thermalen Infrarotbereich (TIR) von der Erdoberfläche aufnehmen. DIEGO wird die erforderliche räumlich höher auflösende multispektrale Infrarot-Sensortechnologie bereitstellen, die zur Ableitung von essentiellen Klimavariablen wie z. B. „Fire disturbance" notwendig ist. DIEGO verfügt sowohl über einen Sternensensor zur genauen Georeferenzierung als auch über ein freies schwenkbares hochauflösendes Kamerasystem, welches mit den 11 Bändern im VNIR, MIR und TIR koregistriert ist. Hier besteht somit auch die Möglichkeit, Thermaldaten mit Videomaterial zu verschneiden und sehr hochaufgelöste multispektrale 3D-Modelle zu generieren.

Zudem soll das DIEGO Projekt, neben den Beitrag zur Grundlagenforschung und zur angewandten Wissenschaft, auch das Wissen über die Schlüsseltechnologien Raumfahrt und Erdbeobachtung im Schulunterricht verbreiten und vertiefen, und neue Wege des naturwissenschaftlichen Wissenstransfers in der digitalen Bildung und des erfahrbaren Lernens beschreiten.

5 Literaturverzeichnis

AGISOFT PHOTOSCAN (2018): User Manual Professional Edition Version 1.4 http://www.agisoft.com/pdf/photoscan-pro_1_4_en.pdf, letzter Zugriff 19.01.2018.

DUNLEAVY, M., DEDE, C., & MITCHELL, R., 2009: Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning. Journal of Science Education and Technology, **18**, 7-22.

GAO, B.-C. LI, R.-R. 2010: Spectral calibrations of HICO data using atmospheric bands and radiance adjustment based on HICO and MODIS data comparisons. Geoscience and Remote Sensing Symposium (IGARSS).

HÖLLERER T, & FEINER S.2004: Mobile augmented reality. In: Karimi H, Hammad A (eds.): Telegeoinformatics: location-based computing and services. Taylor & Francis Books, London.

KORUCU, A.T. & ALKAN, A., 2011: Differences between m-learning (mobile learning) and e-learning, basic terminology and usage of m-learning in education. Procedia Social and Behavioral Sciences, **15**, 1925-1930.

MICHEL, B., 2013: Digital Stereoscopy: Scene to Screen 3D Production Workflows, Stereoscopy News, Stereoscopy News, Belgien.

NEW MEDIA CONSORTIUM (NMC) (2017): NMC Horizon Report: 2017 Higher Education Edition. http://cdn.nmc.org/media/2017-nmc-horizon-report-he-EN.pdf, letzter Zugriff 22.1.2018.

ORTWEIN, A. KREBS, B., MEHRING, J., SCHULTZ J. RIENOW, A. 2017: Neue Dimensionen der Erdbeobachtung im Schulunterricht: Raumfahrt, 3-D und Augmented Reality. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 26, T. Kersten (Hrsg.), 332-341.

RIENOW, A., GRAW, V., HEINEMANN, S., SCHULTZ, J., SELG, F. & MENZ, G., 2016: Mathematikunterricht aus dem All – Interdisziplinäre Lernwerkzeuge für den Einsatz von Erdbeobachtung im Schulunterricht. Dreiländertagung der DGPF, der OVG und der SGPF, Bern, Switzerland, 2016, 7.-9. June, 428-435.

RIENOW, A., GRAW, V., MENZ, G., SCHULTZ, J., SELG, F. & WEPPLER, J., 2015a: Experiencing Space by Exploring the Earth – Easy-to-use Image Processing Tools in School Lessons. IAC-15-E.1.2.2, 66th International Astronautical Congress, Jerusalem, Israel, 2015, 12-16 October.

RIENOW, A., HODAM, H.; SELG, F. & MENZ, G., 2015b: Columbus Eye: Interactive Earth Observation from the ISS in Class Rooms. GI-Forum. Journal for Geographic Information Science, Wichmann, Berlin, 349-353.

RUNCO, S., 2015: International Space Station – High Definition Earth Viewing (HDEV), http://www.nasa.gov/mission_pages/station/research/experiments/917.html, letzter Zugriff 19.01.2018.

RUPNIK, E., PIERROT DESEILLIGNY, M., DELORME, A., & KLINGER, Y. (2016). Refined satellite image orientation in the free open-source photogrammetric tools APERO/MICMAC. ISPRS Annals Photogramm Remote Sens Spatial Information Sciences, **III**-1, 83-90. doi:10.5194/isprsannals-III-1-83-2016.

SCHULTZ, J., ORTWEIN, A. & RIENOW, A. 2017: Technical Note: Using ISS Videos in Earth Observation - Implementations for Science and Education. In: European Journal of Remote Sensing.

Stefanov, W. & Evans, C.A., 2014: The international space station: A unique platform for remote sensing of natural disasters. JSC Biennal Research Report, 108-110.

Stefanov, W., Evans, C.A., & Dasgupta, R. 2011: NASA Earth observations resources. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20110019797.pdf, letzter Zugriff 19.01.2018.

VOß, K., GOETZKE, R., HODAM, H. & RIENOW, A., 2011: Remote Sensing, New Media and Scientific Literacy - A New Integrated Learning Portal for Schools Using Satellite Images. Learning with GI 2011 - Implementing Digital Earth in Education. Berlin, 172-180.

VOß, K., HODAM, H. & GOETZKE, R., 2010: Feuerspuren im Satellitenbild – Mit Fernerkundung die Bewertungskompetenz stärken. Learning with Geoinformation IV -Lernen mit Geoinformationen IV, Jekel, T., Koller, A., Donert, K. & Vogler, R. (Hrsg.), Wichmann, Berlin, 171-181.

VUFORIA DEVELOPER'S GUIDE, 2016: https://library.vuforia.com/, letzter Zugriff am 22.01.2018.

WEISER, M. 1991: The Computer for the 21st Century (1991). Scientific American, **265**(3), 94-104.

Geschichten erzählen mit verlinkten Visualisierungen – ein Review des Alters-Atlas Konzepts

SUSANNE BLEISCH[1], ADRIANA ZANDA[1] & SAFAK KORKUT[2]

Zusammenfassung: Der Alters-Atlas bereitet eine Vielzahl von Themen und Informationen rund um das Alter und Altern auf. Dafür wurde das neuartige Story-Network-Konzept von miteinander verlinkten und kurz erläuterten Visualisierungen entwickelt. Entlang der Verlinkungen im Netzwerk verlaufen sogenannte Geschichtenlinien. Die Nutzenden erkunden den Atlas frei oder navigieren entlang dieser Linien zu interessanten Atlas-Inhalten. Ziel dieses Reviews ist, das Konzept und seine Umsetzung im Alters-Atlas zu diskutieren und evaluieren. Zuerst werden das Konzept und die Inhalte des Alters-Atlas sowie die Möglichkeiten und Grenzen der Navigation und Interaktion im Atlas eingeführt. Dann erläutern wir die verwendete visuelle Sprache und wie die Visualisierungen aufgebaut und miteinander verknüpft sind. Die Evaluation der Rückmeldungen zeigt, dass die verlinkten Visualisierungen funktionieren und die Geschichtenlinien Gefallen finden. Allerdings ist das Kartenprinzip noch zu starr umgesetzt. Die Navigation wird unterschiedlich beurteilt, es bedarf noch zusätzlicher Erläuterungen.

1 Motivation

Atlanten haben eine lange Tradition in der Darstellung von Informationen mit Raumbezug. Sie sind häufig Sammlungen von gedruckten - oder heutzutage auch digitalen - Karten, die lokale, regionale, nationale oder internationale Themenbereiche abbilden und illustrieren. Sie dienen als Nachschlagewerke, Informationsquelle aber immer auch dem Entdecken und Verknüpfen von Informationen und Wissen.

Mit dem Alters-Atlas – einem zentralen Projekt der FHNW Strategischen Initiative Alternde Gesellschaft der Fachhochschule Nordwestschweiz (FHNW 2017) – wurde das Ziel verfolgt, den Altersstrukturwandel in seinen sozialräumlichen und lebenslagenspezifischen Dimensionen, z.B. der Altersgruppenverteilung oder demographischen Entwicklung, mittels einer „Landkarte der alternden Gesellschaft“ visuell darzustellen. Dabei war der Anspruch die erläuternden und verknüpfenden Dimensionen der Atlas-Tradition zu übernehmen, aber etwas weg zu kommen vom Image des dicken verstaubten Bibliothekbuchs – das manchen Atlanten scheinbar anhaftet. Somit wurde der Alters-Atlas als interaktive Plattform konzipiert, die via Internet potentiell einer Vielzahl von Nutzerinnen und Nutzern zur Verfügung steht (www.altersatlas.ch).

In diesem Review erläutern wir zuerst das dem Alters-Atlas zugrundeliegende Konzept und wie dieses umgesetzt wurde. Dabei vertiefen wir die Aspekte Inhalt und Navigation, die verwendete visuelle Sprache und wie Inhalte miteinander verknüpft werden. Mittels erster Rückmeldungen von Nutzenden evaluieren wir die Umsetzung des Alters-Atlas und erörtern seine Stärken und Schwächen. Abschließend geben wir einen Ausblick, wie der Alters-Atlas weiterentwickelt wird.

[1] FHNW Fachhochschule Nordwestschweiz, Institut Geomatik, CH-4132 Muttenz/Basel, Schweiz, E-Mail: [susanne.bleisch, adriana.zanda]@fhnw.ch
[2] FHNW Fachhochschule Nordwestschweiz, Institut für Wirtschaftsinformatik, CH-4052 Basel, Schweiz, E-Mail: safak.korkut@fhnw.ch

2 Konzept des Alters-Atlas

Der initiale Anspruch war, dass ein Atlas erstellt werden soll, der kein Atlas im traditionellen Sinne, beziehungsweise kein reines Kartenwerk ist. Um diesem Anspruch gerecht zu werden, wurde ein neuartiges Konzept entwickelt, das sogenannte „Story-Network Prinzip“ (schematisch aufgezeigt in Abb. 1), bei welchem kurz erläuterte Visualisierungen miteinander verlinkt werden und so ein Netzwerk von Informationen bilden. Dieses Konzept kann mit einer digitalen Kartei[3] verglichen werden. Dabei entsprechen die erläuterten und verlinkten Visualisierungen einzelnen „Karteikarten“. Die Karteikarten werden mittels Links entlang thematischer Gemeinsamkeiten, inhaltlicher Vertiefungen oder Ergänzungen miteinander zu einem Netzwerk verknüpft. Im Fall des Alters-Atlas wird bei diesen Karteikarten abgekürzt von „Karten“ gesprochen, welche immer aus einem Titel, einer grafischen Darstellung und einer Erläuterung bestehen. Dieses Grundkonzept erlaubt es Daten aus den unterschiedlichsten Quellen sowie auch qualitative Aussagen oder Bild- und Filmmaterial zu strukturieren und in das Alters-Atlas Netzwerk einzubinden.

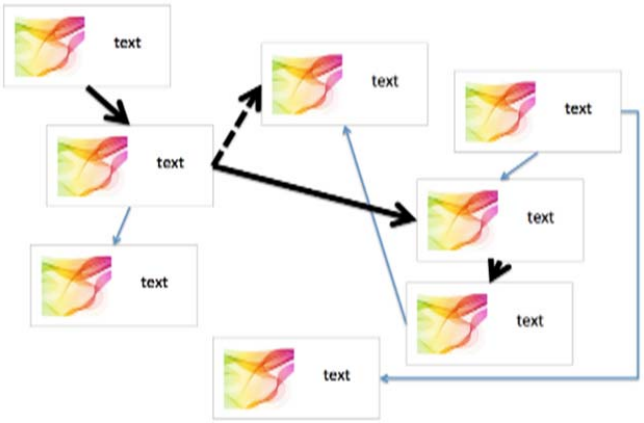


Abb. 1: Schematische Darstellung des Alters-Atlas „Story-Network Prinzips“ – die thematische Verknüpfung von Karteikarten mit Visualisierungen und erläuterndem Text zu einem Netzwerk. Um die Navigation zu erleichtern und gezieltes Entdecken zu ermöglichen, werden einzelne Verbindungen zu sogenannten „Story-Linien“ ausgestaltet (fetter dargestellt).

Durch das Hervorheben von gewissen Verknüpfungen können „Story-Linien“ (Geschichtenlinien) definiert werden, welche den Nutzenden Geschichten erzählen (in schematischer Abb. 1 fett hervorgehoben). Die Nutzenden, die nicht frei im Netzwerk auf Entdeckungsreise gehen möchten, können entlang dieser Linien navigieren und werden so zu interessanten Atlas-Inhalten sowie relevanten Kontextinformationen geleitet. Diese so erzählten „Geschichten“ sollen informieren, Interesse wecken und dazu animieren den Atlas und seine Inhalte weiter zu erkunden. Zusätzlich ermöglicht die flexible Struktur der zu Geschichten verknüpften Karten, das Atlas-Netzwerk, jederzeit eine Erweiterung und Ergänzung der Inhalte. Einzige Bedingung ist, dass jede Karte mit mindestens zwei weiteren Karten im Netzwerk verknüpft wird. Damit kann der Alters-Atlas zukünftig mit der Erhebung von neuen Datensätzen

[3] Eine Kartei ist eine „Sammlung von Karten gleicher Größe und gleichen Formats für einheitliche Aufzeichnungen, die nach bestimmten Ordnungsprinzipien wie Sachgruppen, Alphabet o. Ä. sortiert sind und in [Schub]kästen zur Einsichtnahme aufbewahrt werden.“ (DUDEN 2018)

einfach wachsen. Obsolete Inhalte oder Karten können entfernt werden, indem sichergestellt wird, dass alle Karten, die bisher auf diesen Inhalt verwiesen haben, weiterhin auf zwei unterschiedliche Karten verlinken.

3 Implementierung des Alters-Atlas

Der Alters-Atlas Version 1.0 wurde im Dezember 2017 auf www.altersatlas.ch veröffentlicht. Die folgenden Abschnitte erläutern kurz die technische Umsetzung des Story-Network Prinzips auf der Alters-Atlas Plattform bevor auf die Inhalte und die Möglichkeiten zur Navigation eingegangen wird. Zwei weitere Abschnitte beschäftigen sich mit der verwendeten visuellen Sprache der grafischen Darstellungen sowie der Art und Weise wie die Karten und Visualisierungen zu Geschichten verknüpft sind.

3.1 Technische Umsetzung

Serverseitig wurde das Story-Network Prinzip mit Hilfe der Graphendatenbank-Eigenschaften von OrientDB umgesetzt. Dabei sind die Atlas-Karten die Nodes (Knoten) und die Verlinkungen zwischen den Karten die Relations (Verbindungen). Client-seitig wurde eine auf HTML, JavaScript und CSS basierende Webseite entwickelt. Die grafischen Darstellungen verwenden die D3.js JavaScript Bibliothek. Für die Navigation und die Verlinkungen zwischen den Karten wird auf den zugrundeliegenden Alters-Altas Graphen in der Datenbank zugegriffen. Die verschiedenen Navigationsmöglichkeiten werden mit Hilfe weiterer JavaScript Bibliotheken auf der Alters-Atlas Webseite umgesetzt.

3.2 Inhalte und Navigation

Um einen Einblick in verschiedene Aspekte rund ums Thema Alter und Altern zu bieten, bereitet der Alters-Atlas ausgewählte Inhalte und Informationen aus einer Vielzahl von Quellen auf. Diese Quellen umfassen einerseits diverse weitere Projekte der FHNW Strategischen Initiative Alternde Gesellschaft (FHNW 2017) und werden andererseits durch statistische Daten vom Bundesamt für Statistik (BFS) der Schweiz ergänzt. Aufgrund der unterschiedlichen Herkunft und Granularitäten der Informationen aus verschiedenen Quellen, wurden alle Inhalte auf einzelne Aussagen heruntergebrochen, welche als einzelne Karten umgesetzt und verknüpft werden konnten (Beispiel in Abb. 2, links). Die Hauptelemente einer solchen Karte – Titel, grafische Darstellung und Erläuterung – stehen im Vordergrund und vermitteln die ausgewählten Aussagen und Inhalte immer gleich strukturiert. Ergänzt wird diese „Vorderseite der Karte“ durch Metadaten mit Informationen zu Quellen, AutorInnen und allfälligen weiteren Angaben zur Entstehung der Karte oder dem zugrundeliegenden Projekt. Die Metadaten zu jeder Karte können durch einen Mausklick auf „mehr Infos“ angezeigt werden. Dieser Aufbau unterstützt das im Alters-Atlas verfolgte „Drill Down-Prinzip“ (siehe Abschnitt 3.4), welches interessierten Nutzenden erlauben soll, sich vertiefte Informationen bzw. Hintergrundinformationen zu beschaffen, während die weniger an Details interessierten Nutzenden nicht mit Informationen überhäuft werden.

Im rechten oberen Bereich von Abb. 2 ist die Primär-Navigation des Alters-Atlas ersichtlich. Diese Navigation zeigt einen Ausschnitt des gesamten Atlas-Netzwerks rund um die aktuelle

Inhalts-Karte. Die Linien visualisieren die Verbindungen zu verknüpften Karten, wobei bei jeder Verzweigung der Titel einer Karte steht. Durch Anwählen eines Titels kann auf die jeweilige Karte navigiert werden. Die grüne Linie zuoberst zeigt die Verknüpfung der Karte innerhalb einer Geschichtenlinie an. Wer der grünen Linie entlang vorwärts navigiert, wird entlang einer Geschichte geführt. Die darunterliegenden Linien visualisieren die Verbindungen zu weiteren thematisch verknüpften Karten.

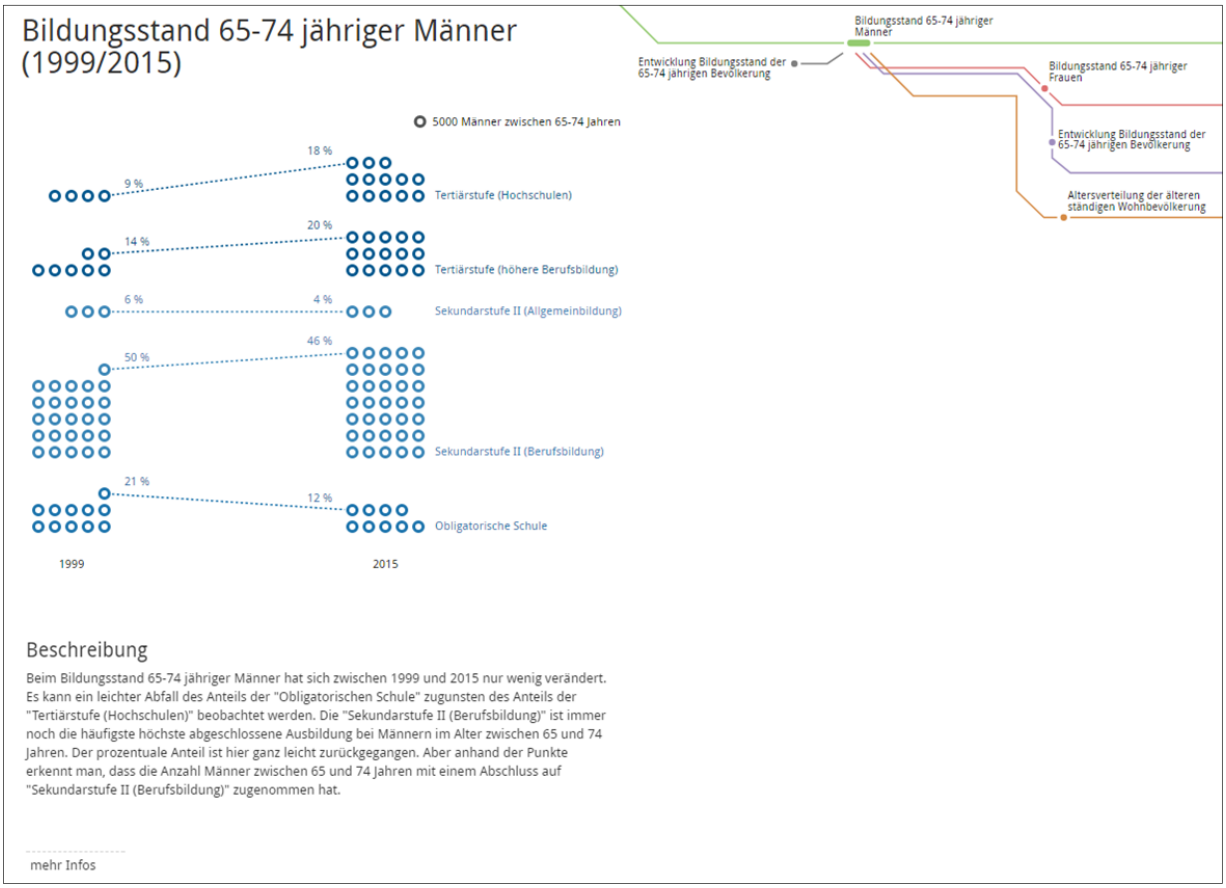


Abb. 2: Beispiel aus dem Alters-Atlas – Links: Karte zum Thema „Bildung“ mit Knopf ‚mehr Infos’ für die Einblendung der Metainformationen der Karte. Oben rechts: Primär-Navigation.

Um Sackgassen in der Netzwerk-Navigation zu vermeiden, muss jede Karte mit mindestens zwei weiteren Karten verlinkt sein. Zusätzlich kann sie auf fast beliebig viele weitere Karten verweisen. Die umgesetzte Navigation stellt bis zu fünf weitere Verbindungen zu verlinkten Karten dar. Eine Einschränkung der im Alters-Altas umgesetzten Navigation ist, dass nie der gesamte Inhalt, bzw. das ganze Netzwerk, betrachtet werden kann. Dies ist aber wenig erstrebenswert, da, abhängig von der Anzahl Karten, die Ansicht des ganzen Netzwerks und seiner Komplexität wenig zur Übersichtlichkeit oder Orientierung in den Inhalten des Alters-Atlas beitragen kann. Dafür bietet diese Art von Navigation den Nutzenden die Möglichkeit sich durch die Inhalte des Atlas führen zu lassen – entlang der Geschichtenlinien – oder sich zu jedem

Zeitpunkt selbst zu entscheiden, welche Karte als Nächstes interessant sein könnte – durch die Navigation zu weiteren verlinkten Karten.
Als Alternative oder Ergänzung zur Primär-Navigation oben rechts beinhaltet der Alters-Atlas noch eine Sekundär-Navigation (Abb. 3), welche unten an der Webseite dargestellt wird. Diese Navigation stellt alle Karten als kurze Striche, die aktuelle Karte als längeren Strich dar. Bereits besuchte Karten werden in dunkelorange hervorgehoben. Diese Navigation erlaubt beliebig – also unabhängig von den zugrundeliegenden Verknüpfungen – zwischen den Inhalten bzw. Karten zu wechseln. Für weitere Zugänge zu Alters-Atlas Inhalten können die Karten mittels Schlagwörter erreicht werden, ebenso möglich ist der Einstieg über eine Auflistung der umgesetzten Geschichtenlinien oder die gezielte Suche nach Karten mit eigenen Begriffen. Ebenfalls ist es möglich sich zufällig eine Karte anzeigen zu lassen.

Abb. 3: Sekundär-Navigation: Auflistung aller Karten als Striche; erlaubt einen beliebigen, von den Verlinkungen unabhängigen, Wechsel zwischen Inhaltskarten.

3.3 Visuelle Sprache

Ein Visualisierungs-Gesamtprojekt, wie der Alters-Atlas, hat zwingend eine visuelle Sprache nötig, die einheitlich verwendet wird und die den Wiedererkennungseffekt von zum Alters-Atlas gehörenden Visualisierungen ermöglicht. Die visuelle Sprache des Alters-Altas ist von vielen Quellen und Beispielen inspiriert, unter anderem von der Wiener Methode der Bildstatistik (NEURATH 1936). Diese besteht aus sogenannten Isotypen – kleine bildhafte Darstellungen –, die Informationen zusammenfassen und darstellen. Isotypen benötigen eine gewisse Mindestgröße, um lesbar zu sein, und verwenden neben der Bildbotschaft auch Farbe oder Größe, um weitere oder unterschiedliche Datendimensionen zu kommunizieren. Wenn wir die bildhaften Isotypen stark vereinfachen, erhalten wir geometrische Formen, wie Kreise, Quadrate oder Dreiecke. Da wir Menschen ein besonderes Faible für runde Dinge haben (FEW 2010), scheinen insbesondere Punkte eine interessante Darstellungsart. Und es überrascht nicht, dass diese Darstellungsart schon seit Jahrzehnten und für die verschiedensten Visualisierungen von Daten verwendet wird (bspw. Abb. 4, links).
Im Alters-Atlas werden Hierarchien von einfachen geometrischen Symbolen (Abb. 4, rechts unten) verwendet, um verschiedene Mengen quantitativer Daten darzustellen. Die unterschiedlichen geometrischen Formen werden für unterschiedliche Einheiten, z.B. Personen oder Prozente, und die verschiedenen Ausprägungen für eine Unterscheidung der Mengen verwendet. Dabei repräsentiert ein ausgefüllter Punkt z.B. 1000 Personen und jedes weniger ausgefüllte Symbol zeigt die Menge, wenn man eine Null streichen würde, also bspw. 1000-100-10-1. Außerdem werden die Symbole gruppiert (Abb. 4, rechts oben) um die Mengenabschätzung oder Zählung der Daten zu vereinfachen. Diese Prinzipien bilden die Basis für alle quantitativen grafischen Darstellungen.

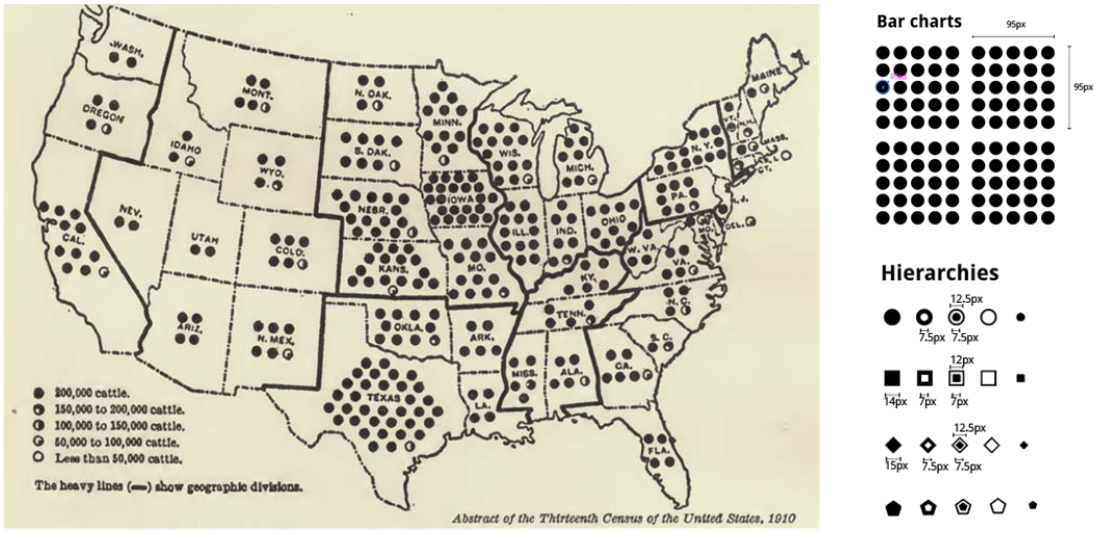


Abb. 4: Links - Kreisförmige, unterteilte Symbole um die Anzahl der Rindtiere in den US-Staaten darzustellen (BRINTON 1914: 215).
Rechts - Darstellungsstrukturen des Alters-Atlas; oben: Gruppierung von Punkten für eine einfachere Mengenabschätzung, unten: Hierarchien von einfachen geometrischen Symbolen.

Zusätzlich zu Form und Füllmenge wurde auch die Verwendung einer beschränkten Anzahl von Farben für den Alters-Atlas festgelegt. Dazu wurden jeweils sechs Primär- und Sekundärfarben ausgewählt (Abb. 5, exkl. grau). Für jede Farbe gibt es eine Farbreihe, wobei die Hauptfarbe die Mittlere ist, gefolgt von 2 helleren bzw. dunkleren Farbtönen, so dass qualitative und quantitative Daten differenziert visualisiert werden können.

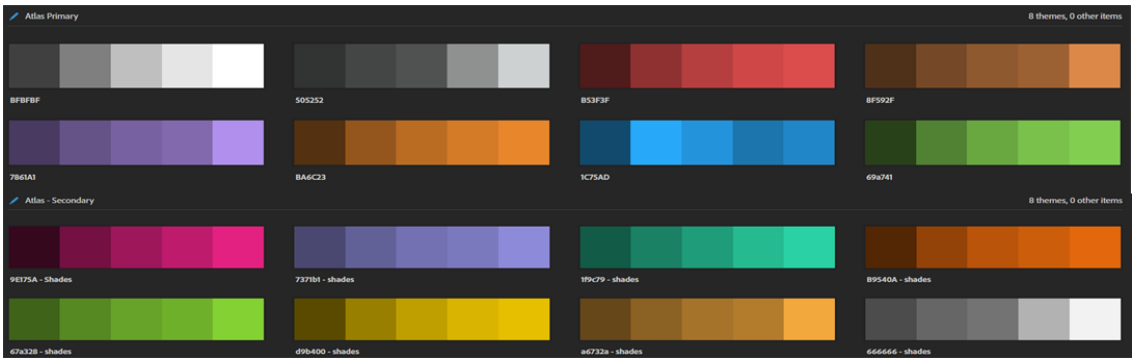


Abb. 5: Verwendete Farbreihen der Primär- und Sekundärfarben im Alters-Atlas.

Basierend auf den in diesem Abschnitt erläuterten Grundlagen der visuellen Sprache konnten verschiedene Diagramme für den Alters-Atlas erstellt werden. Eine kleine Auswahl an beispielhaften Darstellungen werden in Abb. 2 und Abb. 6 gezeigt. Für weitere Beispiele sei auf www.altersatlas.ch verwiesen.

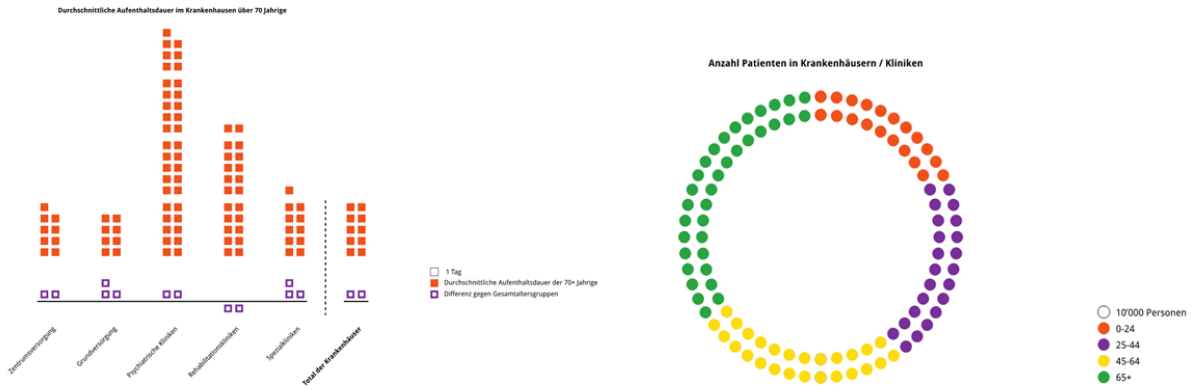


Abb. 6: Beispiele von Diagrammen im Alters-Atlas basierend auf der in Abschnitt 3.3 erläuterten Grundlagen der visuellen Sprache.

3.4 Verknüpfung der Inhaltskarten zu Geschichten

Entlang der Verlinkungen im Netzwerk der Visualisierungen verlaufen sogenannte Geschichtenlinien. Im Alters-Atlas wurden für die Umsetzung solcher Geschichten mehrere Ansätze verfolgt, bspw. das „Drill Down-Prinzip“, „Zoom-Out“ oder das Gegenüberstellen von Informationen („Contrast“, Abb. 7). Beim Drill Down-Prinzip enthält die erste Karte einer Geschichte eine ziemlich allgemeine Aussage zu einem Thema – z.B. Bildungsstand der Bevölkerung in der Schweiz – und die darauffolgenden Karten behandeln dann konkretere Aspekte des Themas mit immer höher aufgelöstem Detailgrad – z.B. der Vergleich des Bildungsstandes von Männern und Frauen im Alter zwischen 65 und 74 Jahren. Eine weitere Verknüpfung kann dann z.B. zu einer Karte verlinken, die diesen Informationen die Daten zum Bildungsstand von jüngeren Personen gegenüberstellt.

Diese Ansätze erlauben Nutzenden zwischen verschiedenen Inhalten auf verschiedenen Abstraktionsstufen zu wechseln und stellt sicher, dass eine Vielzahl von unterschiedlichen Nutzenden den Atlas verwenden können. Wer sich nicht vertieft für eine Thematik interessiert, springt z.B. von "Schlagzeile" zu "Schlagzeile", wen jedoch Details interessieren, kann z.B. von einer "Schlagzeile" aus in die Tiefe der Daten und Informationen navigieren. Dies ist innerhalb eines Themenfeldes aber auch quer zu Themenfeldern möglich.

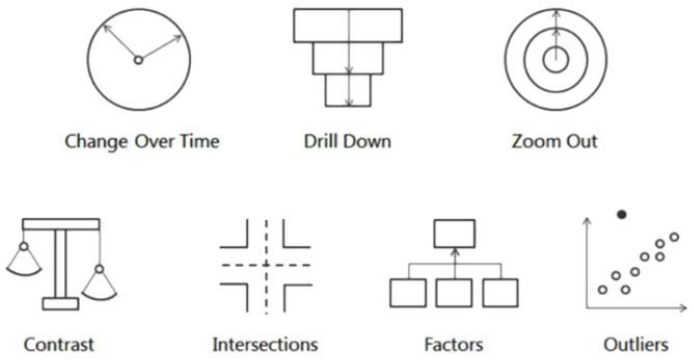


Abb. 7: Verschiedene Story-Typen, die mehr oder weniger häufig im Alters-Atlas für die Verbindung von Karten zu Geschichten oder Stories verwendet werden (JONES 2016).

4 Evaluation und Ausblick

Wie eingangs erwähnt, sollte der Alters-Atlas die verschiedensten Daten und Informationen zum Alter und Altern aufbereiten, visuell darstellen und verschiedenen Nutzenden zugänglich machen. Dabei sollten insbesondere die erläuternden und verknüpfenden Dimensionen traditioneller Atlas-Werke übernommen werden. Überlegungen zur Zugänglichkeit (z.B. sich nicht verloren fühlen in der Vielzahl von Informationen) sowie zur Attraktivität (z.B. was könnte Nutzende anregen, sich länger mit den Inhalten des Alters-Atlas zu beschäftigen) haben zur Entwicklung und Umsetzung des Story-Network Prinzips geführt. Geschichten zu erzählen ist ein jahrhundertealtes Konzept um Wissen zu bewahren und weiterzugeben, aber ebenso um zu unterhalten. Es gibt verschiedene Gründe warum die einen Geschichten funktionieren und andere weniger interessant scheinen. Drei Eigenschaften tragen dazu bei: 1) Geschichten haben einen definierten Start- und Endpunkt mit überlegten Inhalten, 2) sie verwenden durchgängig eine adressatengerechte, ansprechende (visuelle) Sprache und 3) sie haben definierte Charaktere bzw. beziehen sich auf Menschen und erlauben so einen persönlichen Bezug herzustellen. Mit den Geschichten im Alters-Atlas, bzw. dem Story-Network Prinzip und dessen Umsetzung in einer interaktiven Plattform, sollten diese Eigenschaften beachtet und genutzt werden. So wurde antizipiert, dass mittels der Geschichten die Navigation vereinfacht wird, dass die Geschichten informieren, dass sie Interesse wecken, und dass das Netzwerk, bzw. die verknüpften Karten, die Nutzenden dazu animieren, den Alters-Atlas und seine Inhalte weiter zu erkunden. Für eine Evaluation dieser dem Atlas zugrundeliegenden Ideen wurden erste Nutzende verschiedener Anspruchsgruppen gebeten eine kurze Rückmeldung zu machen. Der folgende Abschnitt fasst diese Rückmeldungen zusammen.

4.1 Rückmeldungen

Die ersten Rückmeldungen zum Alters-Atlas Version 1.0 von Nutzenden verschiedener Anspruchsgruppen waren mehrheitlich positiv. Die erwähnten Eigenschaften von Geschichten scheinen auch in der Umsetzung im Story-Network Prinzip recht gut zu funktionieren, wie aus den positiven Rückmeldungen zu Struktur, Gestaltung und Inhalt der Karten bzw. Geschichten herauszulesen ist. Insbesondere sei bei gewissen Geschichten das Drill Down-Prinzip sehr gut umgesetzt, wenn der Ablauf dem Prinzip „von allgemeinen Darstellungen eines Themas über Vertiefungen entlang verschiedener Dimensionen/Navigationswege bis hin zu Lösungen“ folgt.
Andere Rückmeldungen von Nutzenden umfassen Aussagen, wie „Die Geschichten sind informativ“ oder „Interessant gestaltete Homepage, die auf eine Erkundungsreise einlädt“. Wobei die Möglichkeit zur Erkundung von Inhalten von verschiedenen Nutzenden positiv erwähnt wurde. „Durch die Verknüpfung der einzelnen Karten kann der Leser oder die Leserin für ihn oder sie interessante Themen weiterverfolgen.“ Zusätzlich seien auch die Erläuterungen vor allem bei Diagrammen gut, da nicht alle Nutzenden geübt sind, mit Diagrammen umzugehen bzw. diese zu lesen. Dazu wird ergänzt, dass die Grafiken als schön empfunden wurden, jedoch wurden einige Darstellungen als zu klein und nicht leserlich bezeichnet, weil man nicht hineinzoomen kann.
Bei der Navigation sind sich die Nutzenden eher uneinig, eine Nutzerin fand sie „sehr intuitiv und einfach“. Ein anderer Nutzer hingegen stuft die Navigation als „nicht auf Anhieb

verständlich“ ein, weist aber im selben Satz auf die positive Unterstützung durch das Tutorial hin.

4.2 Stärken und Schwächen des Alters-Atlas-Konzept

Die Rückmeldungen von ersten Nutzenden aus verschiedenen Anspruchsgruppen hat ergeben, dass eine Stärke des Story-Network Konzept die verlinkten Visualisierungen sind, welche dem Nutzer bzw. der Nutzerin erlauben, den Atlas-Inhalt den eigenen Interessen entsprechend zu erkunden. Es gibt viele Möglichkeiten den Verbindungen zu folgen, ohne dass man sich in der Datenmenge verloren fühlt. Eine weitere Stärke ist die flexible Netzwerk-Struktur, welche es erlaubt, spannende Geschichten hervorzuheben. Insbesondere diese vordefinierten Geschichten finden Anklang bei den Nutzenden.

Bedingt durch das Prinzip von verknüpften Visualisierungen, wurden die Inhalte auf eine einheitliche Struktur von einzelnen „Karten“ heruntergebrochen. Dies hat Vorteile bezüglich der Übersichtlichkeit und Einheitlichkeit, lässt aber etwas Flexibilität in der Umsetzung der Inhalte vermissen. Gemäß Feedback kommen einige grafischen Darstellungen nicht ideal zur Geltung. Auch in der „Karteikarten“-Struktur wären individuellere Gestaltungen und grafische Darstellungen möglich. Es scheint jedoch, als hätte die „Fixierung“ auf Karten die Kreativität bei der Umsetzung etwas eingeschränkt.

Bezüglich Navigation sind sich die Feedbacks uneinig. Wir haben jedoch festgestellt, dass Rückmeldungen von Nutzenden, die schon bei anderen Gelegenheiten mit dem Story-Network Konzept vertraut gemacht wurden, die Navigation als einfacher oder intuitiver beurteilten, als Nutzende ohne dieses Wissen. Dabei fiel auch auf, dass dieses Konzept im Alters-Atlas selber nirgends erläutert wird, auch nicht im Tutorial.

4.3 Ausblick

Die Weiterentwicklung des Alters-Atlas wird sicher das bisher recht starr implementierte Konzept der „Karteikarten“ aufweichen und innerhalb der Karten mehr Interaktion zulassen. Neben Standard-Interaktionen wie zoomen, Daten via Mouse-Over anzeigen oder Daten-Selektion bzw. -Highlighting, wäre es allenfalls auch interessant direkt aus den Visualisierungen auf weiterführende Karten zu verlinken. Ein erster Schritt ist sicher auch die Implementierung von Karten, die zusätzlich zum bisherigen Tutorial, das Story-Network Prinzip erläutern und so hoffentlich den Zugang zu den Informationen im Alters-Atlas vereinfachen.

5 Danksagung

Die AutorInnen danken dem gesamten Alters-Atlas Entwicklungsteam für die Arbeit an der Umsetzung des Alters-Atlas Version 1.0 sowie der FHNW für die Finanzierung dieses Projekts im Rahmen der Strategischen Initiative Alternde Gesellschaft. Ebenso verdanken wir die Unterstützung vieler KollegInnen und PartnerInnen in den Projekten der FHNW Strategischen Initiative Alternde Gesellschaft sowie auch beim Bundesamt für Statistik der Schweiz. Herzlichen Dank auch den ersten Nutzern und Nutzerinnen, die den Alters-Atlas ausprobiert und erkundet haben und uns mit ihren Rückmeldungen diese erste Evaluation und Reflektion der Umsetzung und der Inhalte des Alters-Atlas ermöglicht haben.

6 Literaturverzeichnis

BRINTON, W.C., 1914: Graphic methods for presenting facts. New York: The Engineering Magazine Company.

DUDEN, 2018: Kartei. https://www.duden.de/rechtschreibung/Kartei, letzter Zugriff 18.01.18.

FEW, S., 2010: Our irresistible fascination with all things circular. Perceptual Edge Visual Business Intelligence Newsletter, 1-9.

FHNW, 2017: Alternde Gesellschaft. https://www.fhnw.ch/de/die-fhnw/strategische-initiativen/alternde-gesellschaft, letzter Zugriff 24.11.17.

JONES, B., 2016: Exploring 7 Kinds of Data Stories. BrightTALK. https://www.brighttalk.com/webcast/11375/191043/exploring-7-kinds-of-data-stories

NEURATH, O., 1936: International Picture Language: The first Rules of Isotype. (P. Kegan, Ed.). London: Trench, Trubner & Co.

Eignungsprüfung einer UAV-basierten Forstinventur als Ersatz zu traditionellen Feldverfahren in Verjüngungsbeständen

MARIUS RÖDER[1], EBERHARD GÜLCH[1] & HOOMAN LATIFI[2,3]

Zusammenfassung: Um Nachforschungen zur Regeneration in vom Borkenkäfer befallenen Beständen Mitteleuropas durchzuführen, ist das Monitoring der Waldstruktur von besonderer Bedeutung. Herkömmlich werden die Strukturparameter in kleinmaßstäbigen Gebieten durch zeit- und personalintensive Feldinventuren erhoben. In den letzten Jahren trat mit der Bildaufnahme mittels Unmanned Aerial Vehicle (UAV) eine neue Form der kleinmaßstäbigen forstlichen Fernerkundung in den Vordergrund. In dieser Arbeit wird geprüft, ob eine UAV-basierte Forstinventur als Ersatz zu traditionellen Feldverfahren in diesen Verjüngungsbeständen geeignet ist. Die Ergebnisse verdeutlichen, dass die UAV-Inventur in den kleinmaßstäbigen Untersuchungsgebieten der Verjüngungsbestände in Qualität und Quantität nur begrenzt mit der Feldinventur mithalten kann. Rein wirtschaftlich gesehen überwiegen allerdings die Vorteile gegenüber dem Feldverfahren.

1 Einleitung

Natürliche Störereignisse gehören zu den Hauptauslösern von Veränderungen in Waldökosystemen. Einer der Hauptgründe für solche Ereignisse in fichtendominierten Waldbeständen Mitteleuropas ist der Befall durch den Borkenkäfer (*Ips typographus* L.) (SVOBODA et al. 2012). Um Nachforschungen zur Regeneration dieser Gebiete durchzuführen, ist in den von Borkenkäfer getroffenen Beständen das Monitoring der Waldstruktur von besonderer Bedeutung. Herkömmlich werden die Strukturparameter (z.B. Baumhöhe (NAESSET 1997) und Kronendurchmesser (POPESCU 2003)) in kleinmaßstäbigen Gebieten durch sehr zeit- und personalintensive Feldinventuren erhoben (HYYPPÄ et al. 2000; MÄKELÄ et al. 2004).
Alternativ könnte die Waldstruktur mithilfe normalisierter Oberflächenmodelle (engl.: normalized Digital Surface Model (nDSM)) durch Fernerkundungsverfahren wie Airborne Laserscanning (ALS) (LEFSKY et al. 2002), photogrammetrische Luftbilder (BOHLIN et al. 2002) oder Satellitenbilder (ST. ONGE 2008) erfasst werden. Diese Methoden weisen in den kleinmaßstäbigen Flächen allerdings Nachteile in Bezug auf Qualität oder Wirtschaftlichkeit auf und sind eher für großflächige Inventuren von intakten Waldbeständen geeignet.
In den letzten Jahren trat mit der Bildaufnahme mittels UAV eine neue Form der kleinmaßstäbigen forstlichen Fernerkundung in den Vordergrund. Ihre geringen Kosten, die Flexibilität und ihre hohe räumliche Auflösung machen sie für solche Anwendungen sehr geeignet (TANG 2015).

[1] Hochschule für Technik Stuttgart, Laboratory for Image Understanding and Earth Observation Sensors (LIMES), Schellingstr. 24, 70174 Stuttgart, E-Mail: eberhard.guelch@hft-stuttgart.de
[2] 97074 Würzburg, Lehrstuhl für Fernerkundung, Universität Würzburg, Oswald-Külpe-Weg 86, E-Mail: hooman.latifi@uni-wuerzburg.de
[3] K. N. Toosi University of Technology, Faculty of Geodesy and Geomatics Engineering, Tehran 19697 64499, Iran

Allerdings sind die Erstellung von nDSM und die anschließende Inventur von Forstgebieten (speziell in z. T. sehr dichten vom Borkenkäfer befallenen Verjüngungsbeständen) mittels UAV noch unzureichend erforscht. PULITI et al. (2015) testen die Eignung von UAV für Inventuren in kleinen Waldgebieten. In DANDOIS et al. (2015) wurden Nachforschungen unternommen, um optimale Höhe, Überlappung und Wetterbedingungen für UAV-Befliegungen zur Ableitung von Forststrukturen zu ermitteln. TUOMINEN et al. (2015) untersuchten die Schätzung von Waldstrukturparametern aus UAV-Daten. Diese und weitere Literaturquellen beziehen sich allerdings allesamt auf erwachsene Waldbestände, die überwiegend aus intakten Bäumen bestehen. Untersuchungen zur Inventur mittels UAV in den vom Borkenkäfer befallenen Verjüngungsflächen sind nicht bekannt.
Die genannten Vorteile der UAVs, die aufgeführten Probleme der herkömmlichen Methoden und der unzureichende Forschungsstand führen dazu, durch weitere Untersuchungen die Eignung von UAV-nDSM für Forstinventuren in den vom Borkenkäfer befallenen Gebieten zu prüfen. Dazu sollen folgende konkrete Ziele erarbeitet werden: es soll untersucht werden, inwiefern sich die aus flugzeuggetragenen Luftbildern erstellten nDSMs und UAV-basierte nDSMs unterscheiden. Dadurch sollen die Vorteile der UAV-nDSMs gegenüber traditionellen Fernerkundungsverfahren in kleinmaßstäbigen Flächen manifestiert werden. Außerdem soll geklärt werden, ob sich die Inventurdaten einer UAV-Befliegung als Referenz anstelle der traditionell im Feld erhobenen Daten eignen. Um diesen Vergleich differenziert zu gestalten, soll ermittelt werden, wie gut sich die UAV-basierte Inventur für die jeweiligen Forsthabitattypen (lebende Fichten/Totholz) sowie die alleinstehende/gruppierte Individuen eignet.

2 Material und Methoden

2.1 Untersuchungsgebiet

Das Untersuchungsgebiet ist der Nationalpark Bayerischer Wald (NPBW). Hier fanden in ausgewählten Plots die Forschungen statt (siehe Abb. 1). Dabei handelt es sich um insgesamt neun Plots mit einer Größe von jeweils ca. 800 m², die sich im Gebiet um Rachel und Lusen in unmittelbarer Nähe zur tschechischen Grenze befinden (siehe Abb. 1). Die Stichproben liegen auf einer Höhe zwischen 1170 und 1340 m ü. NN. In den Jahren 1998 und 1999 starben dort die Fichtenbestände durch Borkenkäferbefall ab. Mittlerweile existieren dort vorwiegend kleine, junge Fichten und stehende bzw. liegende Totholzstämme der abgestorbenen Bäume (HILL et al. 2017). In den Plots lagen alle notwendigen Datengrundlagen (siehe Kapitel 2.2) vor.

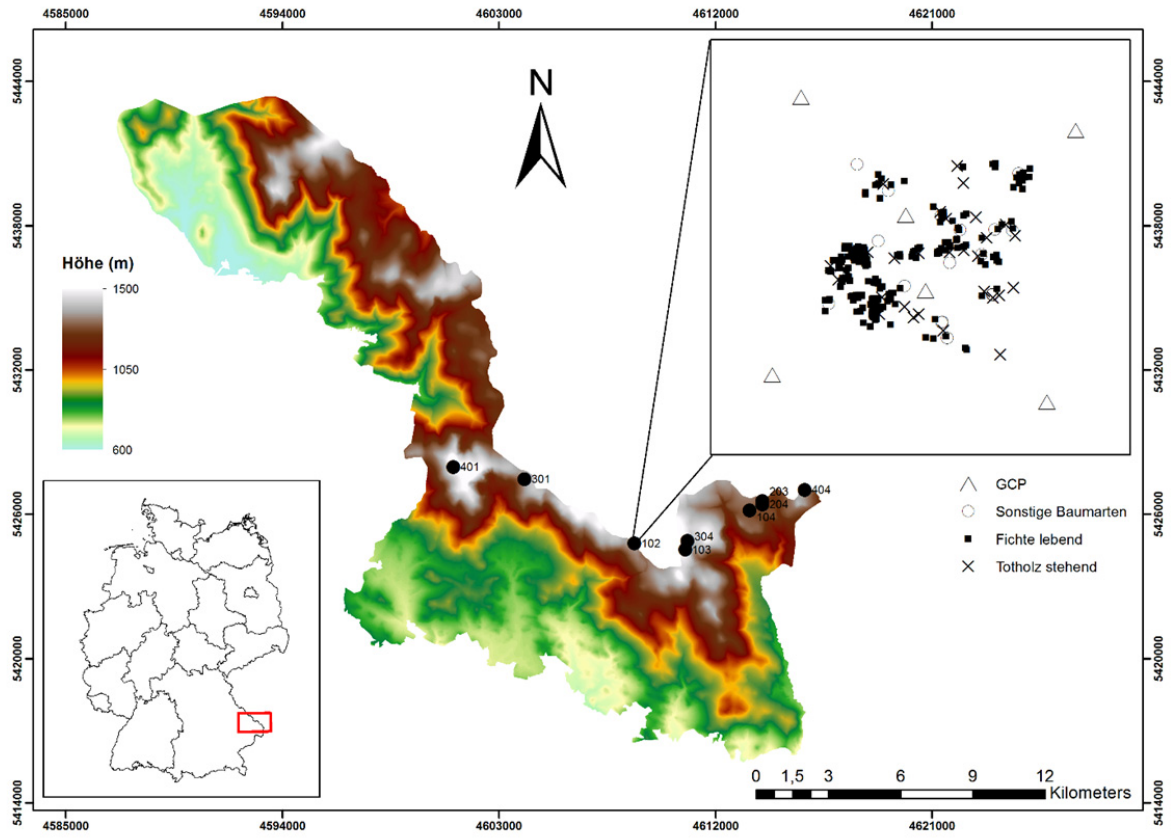


Abb. 1: Übersichtskarte NPBW mit Darstellung der Plots und Detailansicht eines Plots

2.2 Datengrundlagen

Die für den Vergleich herangezogenen photogrammetrischen Luftbilder wurden in einer Befliegung im Juni 2016 mit einer Längsüberlappung von 75 % und einer Querüberlappung von 60 % aufgenommen. Aufnahmesystem war die Z/I-Imaging DMC Kamera mit einer Brennweite von 120 mm. Die Befliegung wurde vom Ingenieurbüro ILV Wagner bei einer durchschnittlichen Flughöhe von 2918 m ü. NN. durchgeführt und durch Bündelblockausgleichung orientiert. Die Punktwolken wurden mit dem Semi-Global Matching Algorithmus (Software ERDAS Imagine 2014) selbst berechnet, anschließend zu DSM mit einer Gitterweite von ca. 30 cm interpoliert und von einem DTM subtrahiert, so dass für jeden Plot nDSM aus Luftbildern resultierten. Das verwendete DTM zur Ableitung der nDSM stammt aus einer ALS-Befliegung aus dem Jahr 2012 vorhanden (Details siehe (LATIFI et al. 2015)).

Zur cm-genauen Georeferenzierung der UAV-Aufnahmen wurden Ground Control Points (GCPs) verwendet. Pro Plot lagen zwei GCPs in der Mitte und vier an den Ecken der Plots vor (siehe Detailansicht Plot in Abb. 1 und Abb. 2). Sie wurden während der Feldinventur mit einer horizontalen Genauigkeit von ca. 10 cm per GPS-Verfahren aufgenommen. Die Messung jedes Punktes betrug mindestens 30 Minuten und wurde im Postprocessing ausgewertet. Es wurden die Daten der nächsten Referenzstation in Prachatice (Tschechien) verwendet, die ca. 40 km von allen Plots entfernt lag.

Die UAV-Daten wurden mit der Drohne DJI Phantom 3 Professional erfasst. Die Befliegungen wurden im Oktober 2016 in einer vierwöchigen Feldkampagne durchgeführt. Die Plots wurden mit einer durchschnittlichen Flughöhe von 40 m, quasi-Nadir Aufnahmen und einer 90-prozentigen Bildüberlappung beflogen. Die Anordnung der Bilder wurde mit der „Double Grid Mission Mode“ der Flugplanungssoftware Pix4D Capture festgelegt. Eine Befliegung eines Plots dauerte ca. 16 Minuten und ergab ca. 200 Bilder (JPEG) mit einer Gesamtgröße von ca. 1 GB pro Plot. Die Bilder wurden nach der Aufnahme auf Störeffekte wie Wind und Nebel kontrolliert und qualitativ schlechte Bilder gegebenenfalls nicht mit in die weitere Verarbeitung eingebunden. Die DSM wurden mit der Software Agisoft PhotoScan (Version 1.2.6) berechnet. Zu den Bearbeitungsschritten zählten u.a. die Herstellung der relativen Orientierung, die cm-genaue äußere Orientierung mittels GCPs sowie die Optimierung der Kameraparameter. Im Optimierungsschritt wurden Tie Points mit hohen Reprojektionsfehlern, hohen Rekonstruktionsunsicherheiten und geringen Projektionsgenauigkeiten entfernt und die Bündelblockausgleichung wiederholt. Ziel der Optimierung war, die besten Tie Points in den Bildern zu erhalten, um das geometrische Modell zu verbessern. Es wurde auch darauf geachtet, dass die Tie Points auf möglichst vielen Bildern vorkommen und auf Bildern und dem Terrain gut verteilt sind. Es resultierten Abweichungen < 10 cm für alle GCPs aller Plots. Eine detaillierte Beschreibung des Workflows und der Prozessierungsschritte findet sich in einem im Laufe der Arbeiten angefertigten Best Practice Tutorial[3]. Es wurden dichte Punktwolken berechnet (siehe Abb. 2), das UAV-DSM mit einer Gitterweite von 5 cm aus Agisoft PhotoScan exportiert und davon das ALS-DTM subtrahiert.

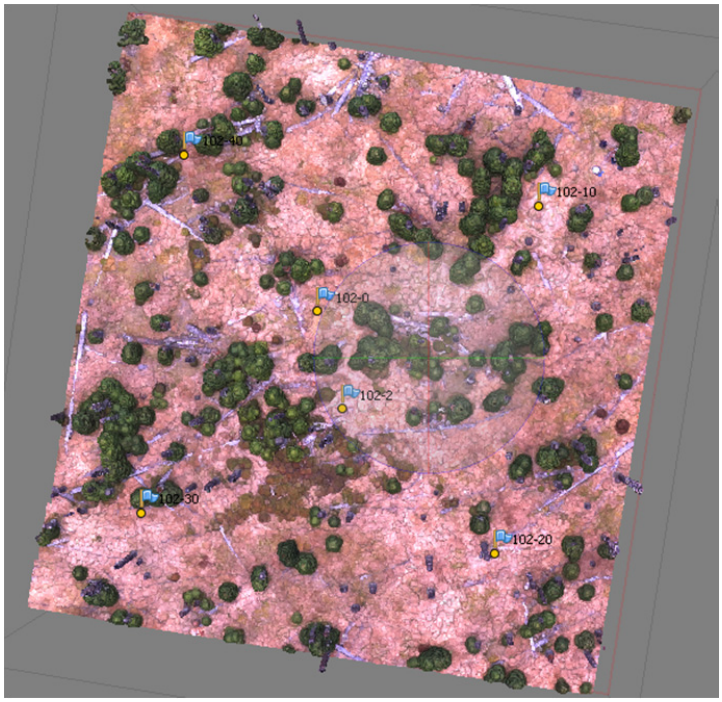


Abb. 2: Nadir-Ansicht einer Punktwolke und der GCPs des Plots 102 in Agisoft PhotoScan

[3]https://www.researchgate.net/publication/316473258_Best_practice_tutorial_Technical_handling_of_the_UAV_DJI_Phantom_3_Professional_and_processing_of_the_acquired_data

Die Referenzdaten aus einer Feldinventur wurden im Juli/August 2016 durch Forscher der Universität Prag durchgeführt. Dabei wurde die Realtime GIS-Software Fieldmap Ver. X5.0.5969 (www.fieldmap.cz), das Distanzmessgerät ForestPro Laser Rangefinder und das Winkelmessgerät MapStar Angle Encoder verwendet. Es wurden die Position (relativ zu den GCPs), die Höhe und der Durchmesser auf Brusthöhe der lebenden Fichten, der stehenden Totholzstämme und sonstigen Baumarten sowie der Kronendurchmesser in Nord-Süd und West-Ost Richtung gemessen (SVOBODA et al. 2016).

2.3 Methodik

Der Vergleich von Luftbild-nDSM und UAV-nDSM erfolgte rein visuell und wurde mithilfe der Gegenüberstellung von Profilen durchgeführt.
Als Kernpunkt der Arbeit wurde dann geprüft, ob sich eine mittels UAV durchgeführte Inventur als Ersatz zu traditionellen Feldverfahren eignet. Dazu wurden Referenz- und UAV-Daten in Hinblick auf Qualität, Quantität und Wirtschaftlichkeit miteinander verglichen.
Der **Qualitätsvergleich** wurde für die geometrischen Parameter Baumhöhe und Kronendurchmesser durchgeführt. Genauigkeitsmaße waren RMSE, R^2 und Bias. Dazu mussten zunächst die Bäume in den UAV-Daten erfasst werden. Zur Detektion wurden ein manueller und ein automatischer Ansatz verfolgt. Beim manuellen Ansatz wurden die Bäume in den farbcodierten Punktwolken in Agisoft PhotoScan markiert und denen der Feldinventur zugeordnet. Anschließend wurden Höhe und Kronendurchmesser von UAV- und Feldverfahren extrahiert und gegenübergestellt. Der Kronendurchmesser der UAV-nDSM wurde genau wie im Feldverfahren in Nord-Süd und West-Ost-Richtung in ArcGIS erfasst.
Das automatische Verfahren basiert auf der Erkennung von lokalen Maxima mit variierenden Fenstergrößen. Verwendet wurde ein nach HILL et al. (2017) modifizierter Algorithmus. Um ausschließlich Maxima zu erhalten, die zu einer Baumspitze gehören, wurde zusätzlich eine Baummaske erstellt. Die Maske basiert auf einer Random Forest Klassifikation der UAV-Orthomosaike.
Zum **Vergleich der Quantität** des manuellen Ansatzes wurden die Anzahl der in UAV- und Feldverfahren erfassten Bäume verglichen.
Zur qualitativen und quantitativen Auswertung des automatischen Verfahrens wurden die Ergebnisse aus Zeitgründen in lediglich drei ausgewählten Plots mit denen des manuellen Verfahrens verglichen.
Zur differenzierteren Auswertung wurden die Gegenüberstellungen beim Qualitäts- und Quantitätsvergleich auch nach lebenden Fichten, stehenden Totholzindividuen, alleinstehenden und gruppierten Individuen klassifiziert.
Der **Wirtschaftlichkeitsvergleich** basierte auf Zeit- und Kostenabschätzungen der Feldinventur und der UAV-Inventur.

3 Ergebnisse

3.1 Vergleich Luftbild- und UAV-nDSM

Abb. 3 zeigt exemplarisch die Profile von UAV-nDSM (schwarze Linie) und Luftbild-nDSM (rote Linie). Manuell wurden die Höhen der Feldinventurdaten (Referenzdaten) eingetragen.

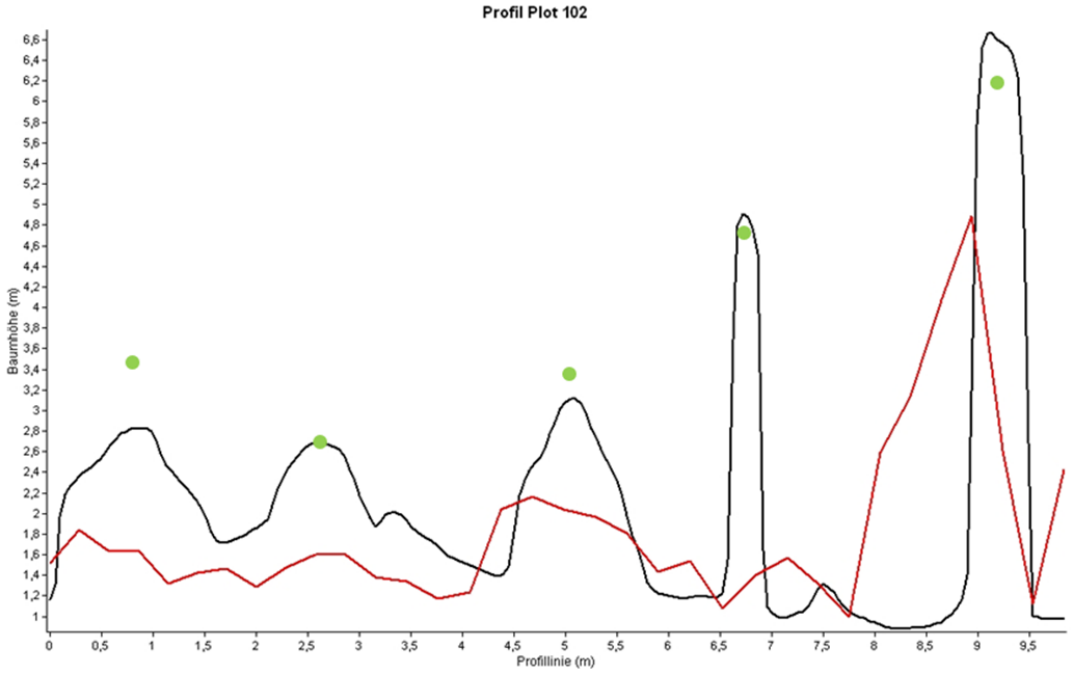


Abb. 3: Profilansicht des Plots 102 mit UAV-nDSM (schwarz), Luftbild-nDSM (rot) und Referenzdaten aus der Feldinventur (grün)

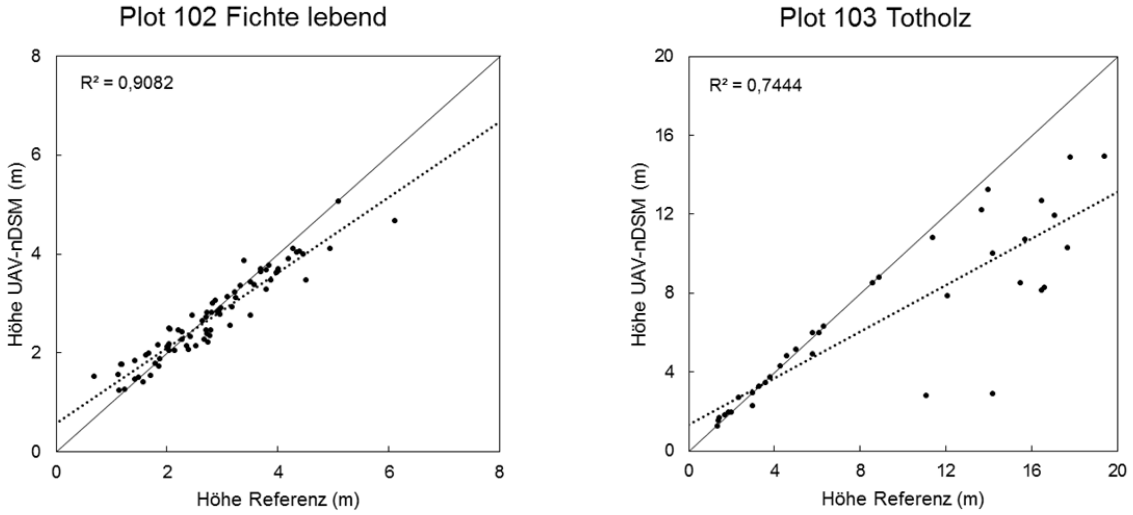
Im Profil des UAV-nDSM sind deutlich mehr Details zu erkennen als im Luftbild-nDSM. Kleine Fichten (links) sind deutlich erkennbar. Im Luftbild-nDSM wird die Höhe dieser Fichten deutlich unterschätzt. Auch im UAV-nDSM wird die Höhe unterschätzt, wenn auch nur geringfügig. Der stehende Totholzstamm bei ca. 7 m der Profillinie wird in den Luftbild-Daten überhaupt nicht modelliert. Der Totholzstamm bei ca. 9 m der Profillinie wird in den Luftbild-nDSM nur sehr grob durch einen einzelnen Punkt erfasst. In den UAV-nDSM werden die Totholzstämme genau rekonstruiert, wobei die Höhen beider Totholzstämme überschätzt werden.

3.2 Prüfung der UAV-Inventur als Ersatz zu traditionellen Referenzdaten

3.2.1 Qualität

Werden die qualitativen Ergebnisse des Parameters Baumhöhe (manueller Ansatz) betrachtet, so lässt sich feststellen, dass der RMSE im Mittel 1.56 m beträgt. R^2 ist im Mittel 0.74 und der Bias ist mit -0.73 m im Mittel negativ.

Der differenzierte Vergleich zeigt, dass lebende Fichten mit einem Durchschnitts-RMSE von 0.81 m deutlich geringere Residuen aufweisen als stehende Totholzstämme (2.34 m). R^2 ist mit einem Mittelwert von 0.89 für lebende Fichten über alle Plots größer als der Wert für Totholzindividuen (0.65). Auch der Bias ist für lebende Fichten (-0.62 m) betragsmäßig kleiner als der der Totholzstämme (-0.80 m). Generell sind die Residuen in Plots mit einem hohen Anteil an kleinen, lebenden Fichten geringer als in Plots mit einem hohen Anteil an großen, stehenden Totholzstämmen. Abb. 4 verdeutlicht diesen Unterschied visuell. In Plot 102 (a) existieren viele kleine Fichten mit moderater Baumhöhe, in Plot 103 (b) kommen große Totholzstämme häufig vor.



(a) (b)

Abb. 4: Diagramme zur Darstellung des Baumhöhenvergleichs von UAV- und Referenzdaten für lebende Fichten in Plot 102 (a) und Totholzstämme in Plot 103 (b)

Diese Unterschiede existieren bei der Differenzierung nach alleinstehenden/gruppiert stehenden Individuen nicht. Hier sind die Mittelwerte in der Regel ähnlich. Die Residuen des Parameters Kronendurchmesser sind beim manuellen Ansatz generell klein. Der RMSE beträgt im Mittel über alle Plots 0.12 m, $R^2 = 0.89$ und der Bias liegt mit 0.04 m im positiven Bereich.
Beim qualitativen Vergleich des automatischen und manuellen Ansatzes bzgl. des Parameters Baumhöhe zeigen sich hohe Korrelationen (R^2 im Mittel 0.99). Die Residuen sind in einem Plot mit hohem Totholzanteil höher als in solchen mit vielen lebenden Fichten. Der Bias ist im Mittel immer positiv, die Baumhöhen werden im automatischen Ansatz also durchschnittlich höher gemessen als im manuellen Verfahren.

3.2.2 Quantität

Beim manuellen Verfahren werden im Mittel nur 38 % der in der Feldinventur erfassten Bäume durch die UAV-Inventur erfasst. Lebende Fichten (37 %) werden dabei in geringerem Maße detektiert als stehendes Totholz (61 %). In Gruppen stehende Bäume weißen mit 30 % eine geringere Detektionsrate als alleinstehende Individuen (56 %). Eine hohe Korrelation von lebenden Fichten und gruppiert stehenden Individuen sowie stehendem Totholz und alleinstehenden Individuen ist zu erkennen. Es fällt außerdem auf, dass im Bereich von 0-1 m wesentlich mehr Bäume in den Referenzdaten detektiert wurden als in den UAV-Daten (siehe Abb. 5).

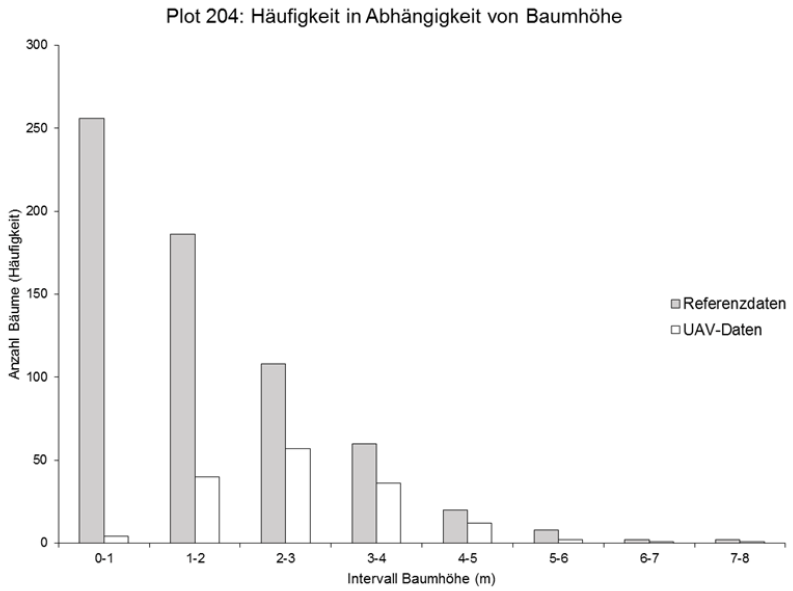


Abb. 5: Anzahl der Bäume (Häufigkeit) in Abhängigkeit von der Baumhöhe in Plot 204

Werden die quantitativen Ergebnisse des automatischen Verfahrens mit denen des manuellen Verfahrens verglichen, so fällt auf, dass im Mittel nur 63 % aller im manuellen Verfahren erfassten Bäume auch durch das automatische Verfahren erkannt werden.

3.2.3 Wirtschaftlichkeit

Der Zeitaufwand für die Feldinventur wurde abgeschätzt. Grundlage waren die Anzahl an Arbeitstagen, die Anzahl der benötigten Mitarbeiter und die daraus resultierende Anzahl an Arbeitsstunden insgesamt. Dabei wurde von einem Tagespensum von 9 Stunden Arbeitszeit pro Arbeiter pro Arbeitstag ausgegangen. Insgesamt ergibt sich ein Zeitaufwand von 842 Arbeitsstunden für die Feldinventur. Hinzu kommen die Kosten für Soft- und Hardware, die insgesamt 14.500 € betragen. Auch für die UAV-Inventur wurde die aufgebrachte Zeit ermittelt. Dabei wurde vom manuellen Ansatz für die Erfassung der Bäume in den UAV Daten ausgegangen. Für die Planung, Durchführung und Ausführung dieser Inventur wurden insgesamt 191 Arbeitsstunden aufgebracht. Die Kosten für Soft- und Hardware betragen 10.700 €. Bei beiden Zeitabschätzungen wurde davon ausgegangen, dass die Arbeiten durch erfahrene Mitarbeiter ausgeführt werden. Die Ergebnisse sind in Tab. 1 zusammengefasst.

Tab. 1: Vergleich Zeitaufwand und Equipmentkosten Feld- und UAV-inventur

	Feldinventur	UAV-Inventur
Zeitaufwand in Stunden	842	191
Kosten Equipment	14500 €	10700 €

4 Diskussion

4.1 Vergleich Luftbild- und UAV-nDSM

Der Profilvergleich hat bestätigt, dass die UAV-nDSM wesentlich besser für die Inventur in den kleinmaßstäbigen Gebieten geeignet sind als die Luftbild-nDSM. Durch die höhere räumliche Auflösung sind mehr Details in den nDSM der UAV-Aufnahmen zu erkennen, was eine zuverlässigere Ableitung von Strukturparametern ermöglicht. Als Gründe können die niedrigere Flughöhe und dadurch bessere Auflösung und die höhere Bildüberlappung genannt werden.

4.2 Prüfung der UAV-Inventur als Ersatz zu traditionellen Referenzdaten

4.2.1 Qualität

Für den Vergleich des Strukturparameters Baumhöhe (manueller Ansatz) lässt sich beurteilen, dass die Qualität weniger davon abhängt, ob die Individuen alleinstehend oder gruppiert sind. Vielmehr hat die Kombination von Forsthabitattyp und Baumhöhe einen entscheidenden Einfluss auf die Qualität der in den UAV-Daten gemessenen Baumhöhe. Gute Ergebnisse wurden für niedrige Baumhöhen der lebenden Fichten und Totholzstämme erzielt. Generell wurden die Baumhöhen der lebenden Fichten unterschätzt. Grund dafür ist, dass in den normalisierten Oberflächenmodellen die Baumspitze nicht exakt modelliert ist. Schlechte Resultate ergeben sich für hohe Totholzindividuen. Grund dafür ist wohl, dass in diesen Plots eine zu niedrige Flughöhe des UAV gewählt wurde. Dadurch weisen hohe Bäume, die sich näher an den Aufnahmepunkten der Drohne befinden, höhere Disparitäten bzw. Parallaxen zwischen den Bildern auf. Durch hohe Disparitäten kommt es beim Dense Image Matching unter Umständen zu Problemen bei der Lokalisierung homologer Punkte.

Unter Ausschluss der qualitativ schlechten Höhenmessungen der Totholzstämme sind die Ergebnisse (Mittelwerte Fichte lebend: RMSE = 0.81 m, R^2 = 0.89) den verwandten Untersuchungen von LISEIN et al. (2013), PULITI et al. (2015), ZARCO-TEJADAA et al. (2014) und DANDOIS et al. (2013) sehr ähnlich bzw. teilweise besser.

Für den Parameter Kronendurchmesser ergaben sich insgesamt sehr gute Resultate. Die meisten in den UAV-Daten gemessenen Werte korrelierten hoch mit denen der Referenzdaten. Grund dafür ist, dass in den hochaufgelösten UAV-nDSM Start- und Endpunkt der Messung sehr gut erkennbar sind.

Die Untersuchung eines Ansatzes zur automatischen Baumerkennung hat gezeigt, dass dadurch der manuellen Baumerfassung qualitativ ähnliche Ergebnisse erzielt werden können. Die stets positiven Bias-Werte deuten darauf hin, dass durch die automatische Baumerkennung die Baumhöhe größer geschätzt wird als durch das manuelle Verfahren. Gleichzeitig beweist diese Tatsache, dass durch die manuelle Baumerkennung nicht immer der höchste Punkt der Individuen erfasst wurde. Dies unterstreicht wiederum die Subjektivität des manuellen Verfahrens. Der positive Bias bei der Messung der lebenden Fichten zeigt, dass in diesem Fall die automatische Baumerkennung die Baumhöhe besser schätzt als die manuelle Baumerkennung.

4.2.2 Quantität

Die Ergebnisse des Quantitätsvergleichs (manueller Ansatz) zeigen, dass mit durchschnittlich 38 % nur ein geringer Teil der Feldinventurdaten durch die UAV-Inventur erfasst werden kann. Der Unterschied zwischen gruppierten und alleinstehenden sowie lebenden Fichten und Totholz ist jeweils sehr groß. Wie bereits erwähnt, korrelieren gruppiert stehende Bäume mit lebenden Fichten sehr hoch. In den Baumgruppen befinden sich sehr viele kleine Fichten unterhalb (engl.: understory, dt.: Unterwuchs) der darüber stehenden Fichten (engl.: overstory, dt.: Oberwuchs). Da durch die photogrammetrischen Verfahren nur die Oberfläche modelliert wird, ist es nicht möglich den Unterwuchs in Baumgruppen zu erfassen. So ergibt sich im Mittel ein Anteil von 30 % der gruppiert stehenden Bäume. Alleinstehende Bäume korrelieren hingegen hoch mit Totholzstämmen. Unterwuchs-Individuen kommen hier so gut wie nicht vor. So können 56 % der alleinstehenden Bäume (meist Totholzstämme) in den UAV-Daten erfasst werden.
Allerdings liefert das automatische Verfahren quantitativ schlechtere Resultate als das manuelle. Im Mittel wurden 63 % der im manuellen Verfahren erfassten Individuen auch im automatischen Verfahren erkannt. Der Anteil an den Referenzdaten fällt demzufolge noch geringer aus als durch die manuelle Baumerkennung (Mittelwert 38 %). Zusammenfassend lässt sich feststellen, dass der zeitliche Vorteil, den die automatische Erkennung mit sich bringt, den quantitativen Nachteil des verwendeten Algorithmus in dieser kleinmaßstäbigen Inventur nicht wettmacht. Durch Weiterentwicklung und Anpassung des Algorithmus der automatischen Baumerkennung könnten sich der manuellen Baumerkennung gleichwertige Resultate ergeben.
Die qualitativen und quantitativen Ergebnisse dieser Arbeit lassen sich zusammenfassend in Abb. 6 übersichtlich darstellen. Hier wird zum einen veranschaulicht, dass die sehr hohen Totholzstämme auf der linken Seite der Grafik nicht vollständig modelliert werden. Die Distanz zwischen Kamera und Aufnahmeobjekt ist zu nah, wodurch es zu den großen Differenzen zwischen Referenz- und UAV-Daten kommt. Die Höhen von relativ niedrigen Totholzstämme links neben und in der Baumgruppe weisen keine Ausreißer auf, werden allerdings überschätzt. Innerhalb der Baumgruppe werden die Individuen im Unterwuchs nicht modelliert. Es ist klar ersichtlich, dass nur die Oberfläche der Baumgruppe und nicht deren vertikale Struktur durch das photogrammetrische Verfahren rekonstruiert wird. Dies führt zu den schlechten quantitativen Ergebnissen in Baumgruppen und bei lebenden Fichten. Die Baumspitzen der lebenden Fichten werden geglättet und somit die Höhen unterschätzt. Auf der rechten Seite ist symbolisch eine alleinstehende Fichte zu erkennen, deren Kronendurchmesser durch die UAV-nDSM gut ableitbar ist.

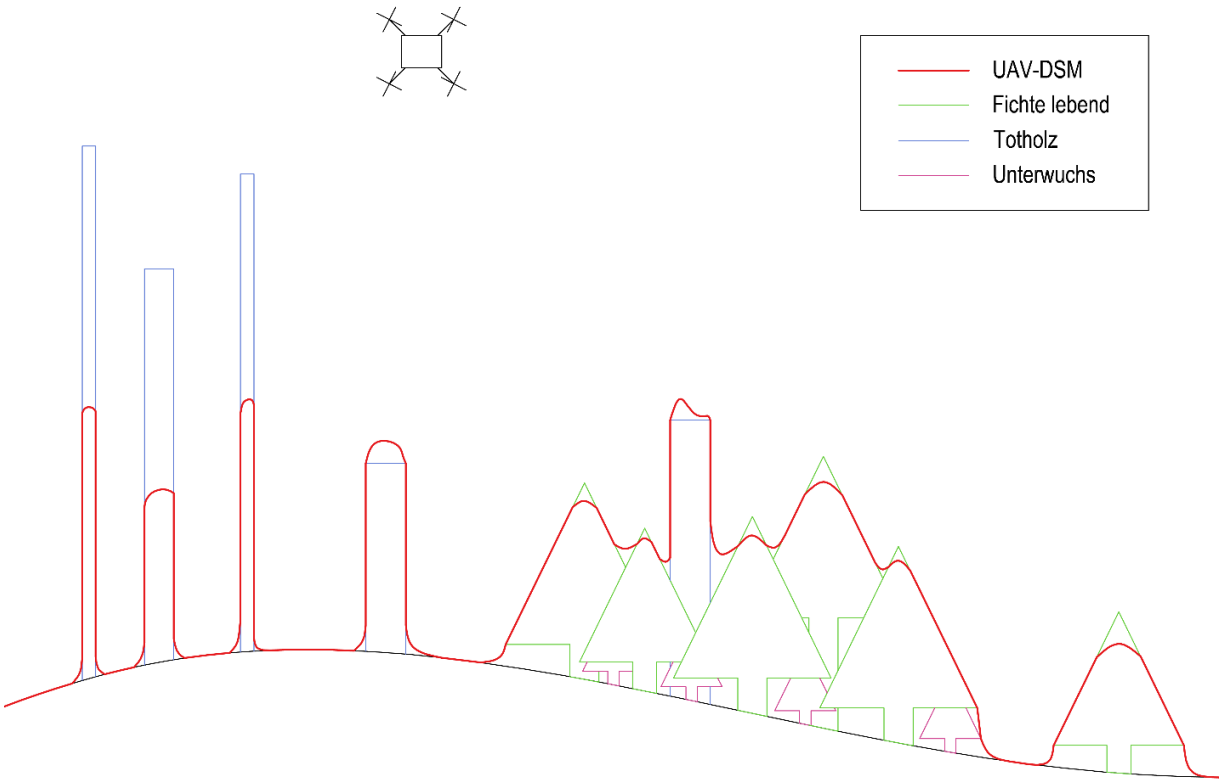


Abb. 6: Schematische Darstellung der Resultate - eigene Darstellung

4.2.3 Wirtschaftlichkeit

Die Untersuchungen zur Wirtschaftlichkeit haben veranschaulicht, dass durch die UAV-Inventur drei Viertel der Zeit gegenüber der Feldinventur eingespart werden kann. Außerdem sind die Kosten durch die UAV-Inventur im Rahmen dieser Arbeit um ein Drittel reduziert.

5 Fazit & Ausblick

Um die Eignung als Ersatz zu traditionellen, feldbasierten Verfahren bewerten zu können, müssen die Vergleichsergebnisse von Qualität, Quantität und Wirtschaftlichkeit gemeinsam betrachtet werden. Qualitative Nachteile ergaben sich in der Erfassung von sehr hohen Totholzstämmen. Positiv fallen hier die geringen Abweichungen bei der Erfassung von kleinen Fichten oder Totholzstämmen auf. In Plots mit einer großen Anzahl an Unterwuchs-Individuen ist die UAV-basierte Methode dem Feldverfahren quantitativ unterlegen. Durch das photogrammetrische Auswerteverfahren kann die vertikale Struktur der Bäume nicht erfasst werden, was insbesondere bei Baumgruppen der Fall ist. Im Gegensatz dazu resultieren in Plots mit wenig Unterwuchs und vielen alleinstehenden Individuen gute quantitative Ergebnisse. Zusammenfassend haben die Ergebnisse des Quantitätsvergleichs gezeigt, dass durch die UAV-Inventur nicht in dem Maße die notwendigen Informationen gesammelt werden können wie durch das Feldverfahren. Das automatische Verfahren erwies sich in qualitativer Hinsicht dem manuellen Verfahren mindestens gleichauf, zeigte jedoch Nachteile in Bezug auf die Quantität. Durch Weiterentwicklung der

Algorithmen der automatischen Baumerkennung könnten hier Verbesserungen erreicht werden. Wirtschaftlich gesehen ist die UAV-Inventur definitiv günstiger und mit einem wesentlich geringeren Zeitaufwand verbunden als die Feldinventur.
Eine mögliche Lösung dieser Probleme könnte auch darin bestehen, die UAV-Inventur mit den Methoden des Feldverfahrens zu kombinieren. Die GNSS-Messungen der GCPs nehmen wesentlich mehr Zeit in Anspruch als die UAV-Befliegungen. Die Arbeitskraft, die für die UAV-Aufnahmen zuständig ist, könnte parallel zu den GNSS-Messungen die notwendigen Informationen durch Feldverfahren erfassen, die durch die UAV-Inventur nicht ableitbar sind (z.B. Unterwuchs, sehr hohe Totholzstämme).
Alternativ sollte in Zukunft ein kombinierter Ansatz aus Photogrammetrie und Laserscanning getestet werden, wie es in (WALLACE et al. 2016) durchgeführt wurde. Ein zusätzlich an die Drohne montierter Laserscanner wäre in der Lage, durch First- und Last-Pulse Daten auch die vertikale Struktur zu erfassen. Durch die Last-Pulse Daten könnte ein DTM abgeleitet werden, was wiederum die Notwendigkeit eines externen DTM hinfällig macht. Hohe Totholzstämme würden durch die Laserimpulse auch mit relativ niedrigen Flughöhen erfasst werden. Ein zusätzlich in die Drohne integrierter, hochgenauer GNSS-Zweifrequenzempfänger würde die GCPs und damit die terrestrischen GNSS-Messungen obsolet machen. Dieser Ansatz würde auf der einen Seite Arbeitsstunden einsparen und Qualität sowie Quantität verbessern, auf der anderen Seite aber auch eine deutliche Kostensteigerung mit sich bringen. Wünschenswert wäre deshalb eine absteigende Preisentwicklung von teuren Produkten wie GNSS-Zweifrequenzempfänger oder Laserscanner durch weitere technische Entwicklungen.

6 Literaturverzeichnis

BOHLIN, J., WALLERMAN, J. & FRANSSON, J., 2012: Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high resolution DEM. Scandinavian Journal of Forest Research, **27**, 692-699.

DANDOIS, J. & ERLE, E., 2013: High spatial resolution threedimensional mapping of vegetation spectral dynamics using computer vision. Remote Sensing of Environment, **136**, 259-276.

DANDOIS, J., OLANO, M. & ERLE, E., 2015: Optimal Altitude, Overlap, and Weather Conditions for Computer Vision UAV Estimates of Forest Structure. Remote Sensing, **7**(10), 13895-13920.

HILL, S., LATIFI, H., HEURICH, M. & MÜLLER, J., 2017: Individual-tree- and stand-based development following natural disturbance in a heterogeneously structured forest: a LiDAR-based approach. Ecological Informatics, **38**, 12-25.

HYYPPÄ, J., HYYPPÄ, H., INKINEN, M., ENGDAHL, M., LINKO, S. & ZHU Y.-H., 2000: Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. Forest Ecology and Management, **128**, 109-120.

LATIFI, H., FASSNACHT, F., MÜLLER, J., THARANI, A., DECH, S. & HEURICH, M., 2015: Forest inventories by LiDAR data: A comparison of single tree segmentation and metric-based methods for inventories of a heterogeneous temperate forest. International Journal of Applied Earth Observation and Geoinformation, 162-174.

LEFSKY, M., COHEN, W., PARKER, G. & HARDING, D., 2002: Lidar Remote Sensing for Ecosystem Studies: Lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. BioScience, **52**, 19-30.

LISEIN, J., PIERROT-DECEILLIGNY, M., BONNET, S. & LEJEUNE, P., 2013: A Photogrammetric Workflow for the Creation of a Forest Canopy Height Model from Small Unmanned Aerial System Imagery. Forests, **4**, 922-944.

MÄKELÄ, H. & PEKKARINEN, A., 2004: Estimation of forest stand volumes by Landsat TM imagery and stand-level field-inventory data. Forest Ecology and Management, **196**(2-3), 245-255.

NAESSET, E., 1997: Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing, **52**(2), 49-56.

POPESCU, S. 2003: Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Can. J. Remote Sensing, **29**(5), 564-577.

PULITI, S., ORKA, H.O., GOBAKKEN, T. & NAESSET, E., 2015: Inventory of Small Forest Areas Using Unmanned Aerial System. Remote Sensing, **7**(8), 9632-9654

ST. ONGE, B., 2008: Mapping the height and aboveground biomass of a mixed forest using lidar and stereo Ikonos images. International Journal of Remote Sensing, **29**, 1277-1294.

SVOBODA, M., JANDA, P., NAGEL, T., FRAVER, S., REJZEK, J., & BACE, R., 2012: Disturbance history of an oldgrowth sub-alpine Picea abies stand in the Bohemian Forest, Czech Republic. J. Veg. Sci., **23**, 86-97.

SVOBODA, M. & WILD, J., 2016: Measuring of the permanent plots established by M. Bauer. Department of Forest Ecology, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague.

TUOMINEN, S., BALAZS, A., SAARI, H., PÖLÖNEN, I., SARKEALA, J. & VIITALA, R., 2015: Unmanned aerial system imagery and photogrammetric canopy height data in area-based estimation of forest variables. Silva Fennica, **49**(5), Article ID 1348.

WALLACE, L., LUCIEER, A., MALENOVSKY, Z., TURNER, D. & VOPENKA, P., 2016: Assessment of Forest Structure Using Two UAV Techniques: A Comparison of Airborne Laser Scanning and Structure from Motion (SfM) Point Clouds. Forests, **7**(3), 62.

TANG, L., 2015: Drone remote sensing for forestry research and practices. Journal of Forestry Research.

ZARCO-TEJADAA, P. J., DIAZ-VARELAA, R., ANGILERI, V. & LOUDJANI, P. 2014: Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods. European Journal of Agronomy, **50**, 89-99.

Bestimmung von Ortspositionen unter Nutzung von freien globalen Daten

THOMAS BRINKHOFF[1]

Zusammenfassung: Die Positionierung von Orten in webbasierten Karten ist oftmals unbefriedigend. Ursachen hierfür sind u.a. qualitativ unzureichende offene globale Datenbasen und vereinfachte Ansätze zur Ableitung der Position aus Flächendaten. Daraus leitet sich die Fragestellung ab, wie man unter Nutzung von freien, möglichst global vorliegenden Daten eine „gute" Ortsposition bestimmen kann. Ausgangspunkt sind einerseits Flächenbeschreibungen der Orte und andererseits Informationen über Siedlungsflächen, die aus OpenStreetMap-Daten oder dem „Global Urban Footprint" des DLR gewonnen werden können. Neben der Verschneidung und der geeigneten Bestimmung eines Ortszentrums ist dabei auch die Behandlung von Sonderfällen zu beachten, die sich aus den Eigenschaften der genutzten Daten ergeben.

1 Einleitung

Webbasierte Kartenanwendungen sind eine etablierte Technologie zur Visualisierung, Analyse und Bearbeitung von raumbezogenen Daten. Grundvoraussetzungen für solche Applikationen sind (a) geeignete Programmbibliotheken und (b) im Internet publizierbare Geodaten. In den letzten Jahren hat sich eine Reihe von JavaScript-basierten Application Programming Interfaces (APIs) etabliert, die clientseitig in Web Browsern genutzt werden können. Beispiele sind OpenLayers (http://openlayers.org/), Leaflet (http://leafletjs.com/), Google Maps JavaScript API (https://developers.google.com/maps/documentation/javascript/reference) oder ArcGIS API for JavaScript (https://developers.arcgis.com/javascript/). Solche Programmbibliotheken erlauben neben dem Einblenden von raster- oder SVG-basierten Hintergrunddaten insbesondere das Hinzufügen von anwendungsspezifischen, meist vektorbasierten Geodaten und die Programmierung spezifischer Visualisierungs-, Analyse- und Bearbeitungsfunktionalität (ROTH et al. 2014).
Die Hintergrundkarten werden oftmals über Webdienste in die Webanwendungen eingebunden. Diese Dienste sind entweder eng mit der API gekoppelt (z.B. bei Google Maps) oder werden über standardisierte Schnittstellen aufgerufen (z.B. WMS oder WMTS). Das wesentliche Alleinstellungsmerkmal einer webbasierten Kartenanwendung sind aber in der Regel die anwendungsspezifischen Geodaten und die darauf anwendbare Funktionalität. Von daher sind gerade für diese Daten die Qualitätsansprüche hinsichtlich Korrektheit und Visualisierungsqualität in der Regel hoch. Der zweite Aspekt kann beispielsweise die Beschriftung der Kartenobjekte betreffen, die durch die genannten Programmbibliotheken eher mangelhaft unterstützt wird, so dass ggf. eigene Lösungsansätze erforderlich sind (BRINKHOFF 2017). Der Fokus dieses Beitrags liegt auf dem erstgenannten Aspekt: Webbasierte Kartenanwendungen können Punkt-, Linien- und Flächengeometrien aufnehmen. Im Fall von komplexeren Geometrien kann die geeignete Generalisierung

[1] Jade Hochschule Wilhelmshaven/Oldenburg/Elsfleth, Institut für Angewandte Photogrammetrie und Geoinformatik, Ofener Str. 16/19, D-26121 Oldenburg, E-Mail: thomas.brinkhoff@jade-hs.de

eine wichtige Anforderung sein, um einerseits eine ansprechende Visualisierung und andererseits eine hinreichende Komprimierung der Daten zu erzielen. Bei Punktsymbolen stellen maßstabsabhängige Sichtbarkeit (BRINKHOFF 2017) und die Positionierung wichtige Kriterien dar. Punktsymbole repräsentieren oftmals flächenhafte Objekte der realen Welt. Dann muss festgelegt werden, wo der Punkt platziert wird. Im vorliegenden Beitrag wird dies für Orte wie Städte und Gemeinden betrachtet. Eine schlechte Platzierung ist für den Benutzer oftmals leicht erkennbar, sobald eine Satelliten- oder Luftbildkarte im Hintergrund verwendet wird und das Ortssymbol beispielsweise über unbesiedelten Gebiet oder über Wasserflächen eingeblendet wird. Die genannte Problematik ist sicherlich für einzelne Staaten, die qualitativ hochwertige amtliche Daten frei zur Verfügung stellen, leicht lösbar. Ein Blick in Google Earth zeigt aber, dass diese Aussage nicht im globalen Maßstab gilt: Bei vielen indischen Dörfern oder chinesischen Gemeinden (Townships) liegt das Ortssymbol oftmals über unbesiedelten Gebieten; Abbildung 1 illustriert die fehlende Korrespondenz.

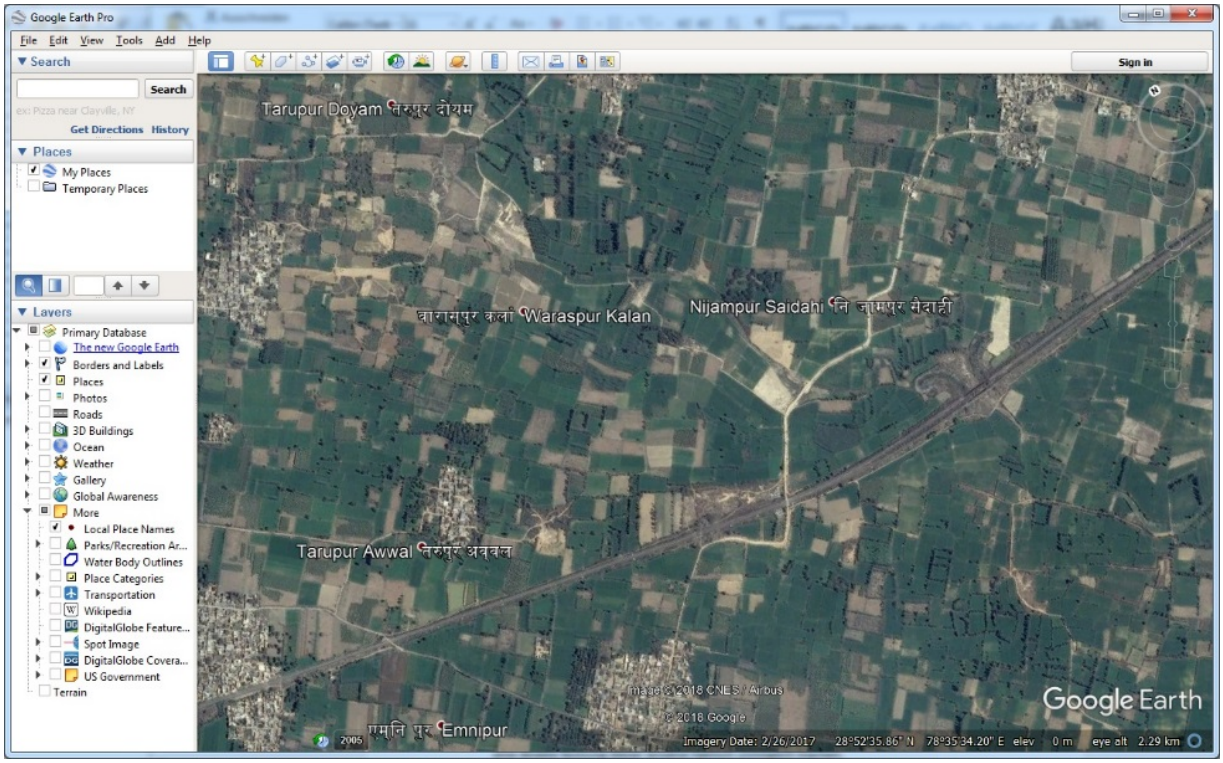


Abb. 1: Positionierung von Symbolen für indische Orte in Google Earth (© CNES/Airbus, Google)

Für die Nutzung von Geodaten in globalen Webanwendungen sind zudem rechtliche, organisatorische und finanzielle Nebenbedingungen zu beachten: Der kostenpflichtige Bezug hochwertiger Daten einzelner Staaten unter ggf. eingeschränkten Nutzungsbedingungen ist hierbei meist nicht zielführend. Daher werden nachfolgend nur Daten aus Quellen betrachtet, die unter den geltenden Lizenzbedingungen frei beziehbar und im Web publizierbar sind.

Damit kann die hier betrachtete Forschungsfrage wie folgt zusammengefasst werden: Wie kann über freie, global vorliegende Daten eine Ortsposition bestimmt werden, die möglichst gut den Ort repräsentiert? Dazu wird im nächsten Kapitel ein Überblick über freie punkt- und flächenhaf-

te Ortsdaten gegeben. Kapitel 3 behandelt Siedlungsflächen, die aus OpenStreetMap-Daten gewonnen wurden, bzw. den aus Radardaten abgeleiteten „Global Urban Footprint“ (GUF) des DLR. Das vierte Kapitel beschreibt die Bestimmung von Ortspositionen auf Basis von (durchaus ungenauen) Flächenbeschreibungen der Orte und den Siedlungsflächen. Neben der Verschneidung und der geeigneten Bestimmung eines Ortszentrums wird dabei auch auf die Behandlung von Sonderfällen eingegangen, die sich aus den besonderen Eigenschaften der genutzten Daten ergeben. Der Beitrag endet mit einem kurzen Fazit und Ausblick auf zukünftige Weiterentwicklungen.

2 Freie Geodaten für Ortspositionen

2.1 Punktdaten

In (BRINKHOFF 2015) werden vier wesentliche Quellen für freie globale Daten über Ortspositionen genannt: GeoNames, OpenStreetMap, Wikipedia und freie amtliche Daten. GeoNames (http://www.geonames.org/) stellt (Stand Januar 2018) für rund 11,5 Millionen geographische Objekte Punktkoordinaten in WGS84-Koordinaten zur Verfügung. Als Defizite sind fehlende nationale Schlüssel, unvollständige administrative Informationen und fehlende Qualitätsindikatoren für die Lageinformation zu nennen. Die Qualität von Ortspositionen ist nach Erfahrungen des Autoren oftmals schlecht.

OpenStreetMap (OSM; http://www.openstreetmap.org) enthält – neben vielen anderen Informationen – auch Ortspositionen. Dabei handelt es sich um Knotenelemente (node), die über das Tag „place“ gekennzeichnet sind (4,0 Millionen, Stand Januar 2018). Neben den WGS84-Koordinaten können solche Elemente Informationen über die administrative Ebene und den amtlichen Schlüssel besitzen. Quantitative Qualitätsindikatoren für die Lageinformation fehlen. Auch variiert die tatsächliche Belegung der genannten Attribute. Die Qualität von Ortspositionen variiert.

Wikipedia stellt für viele Artikel eine meist punktbezogene Georeferenzierung zur Verfügung, die über unterschiedliche Werkzeuge abrufbar ist. Praktikabler ist allerdings oftmals inzwischen die Nutzung von WikiData (https://www.wikidata.org/). Falls die entsprechende Information in WikiData hinterlegt ist, kann mit Hilfe einer SPARQL-Anfrage die Zuordnung der nationalen Schlüssel zu WikiData-Schlüssel bestimmt werden;

```
SELECT ?wd ?key WHERE { ?wd wdt:P1116 ?key . }
```

führt dies exemplarisch für griechische Verwaltungseinheiten durch. Mit Hilfe des WikiData-Schlüssels können dann die Koordinaten abgerufen werden. Ursprüngliche Herkunft und Qualität lassen sich allerdings nicht nachvollziehen. In (BRINKHOFF 2015) wird ein Werkzeug vorgestellt, das die Abweichungen zu amtlichen Daten klassifiziert. Zudem treten (in geringen Maße) Fehler bei der Zuordnung der amtlichen Schlüssel auf.

Amtliche Ortspositionen sind für viele Länder, aber nicht global frei verfügbar. Das Qualitätsversprechen beruht auf der Quelle der Daten, aber in der Regel nicht auf quantitativen Indikatoren. Zudem kann es aufgrund einer (ggf. vor Freigabe) reduzierten Koordinatenauflösung zu Lageabweichungen kommen. Amtliche Schlüsselangaben sind in solchen Datensätzen meist enthalten, können aber ggf. veraltet sein.

2.2 Flächenhafte Daten

Ortsymbole repräsentieren flächenhafte Objekte der realen Welt. Dies können administrative Einheiten (z.B. Gemeinden in Deutschland), statistische Gebiete (z.B. Census-Designated Places in den USA) oder aus Bebauungsinformationen abgeleitete Flächen sein (z.B. Built-up Areas in England und Wales).

OpenStreetMap und amtliche Stellen sind wichtige Quellen für flächenhafte Ortsdaten. Die Aussagen aus dem vorigen Abschnitt gelten in analoger Weise für beide Quellen. Hier betrifft die Lageungenauigkeit den Grenzverlauf. Abbildung 2 zeigt den Grenzverlauf zwischen Bremen und Oyten/Niedersachsen im BKG-Datensatz „Verwaltungsgebiete 1:250 000 - Stand 01.01.2017“ (http://www.geodatenzentrum.de/geodaten/). Die in der Dokumentation angegebene Genauigkeit von 100m wird leider nicht eingehalten, so dass bevölkerungsreiche Gebiete von Bremen-Tenever vermeintlich in Niedersachsen liegen.

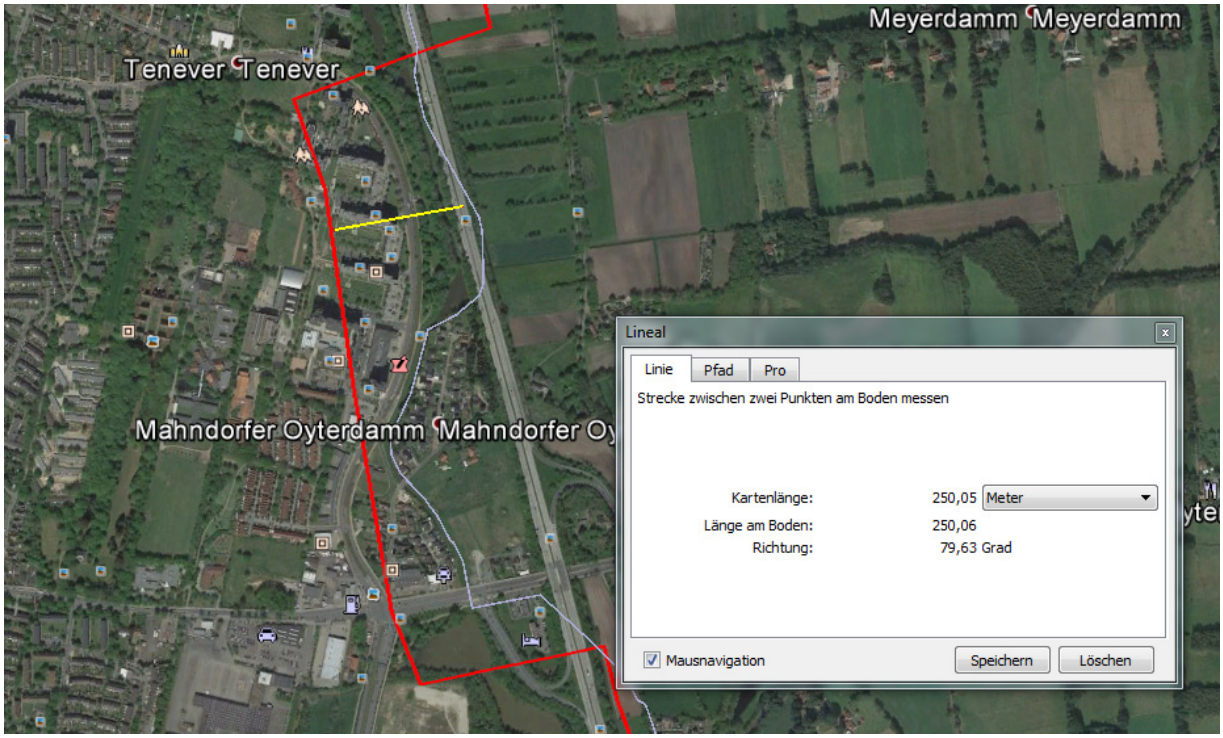


Abb. 2: Abweichung des Grenzverlaufs zwischen Bremen und Oyten/Niedersachsen: rot der Verlauf gemäß den freien BKG-Daten (© GeoBasis-DE / BKG 2018), in hellblau der tatsächliche Verlauf (© Google, GeoBasis-DE / BKG 2009)

Eine weitere Quelle für globale administrative Daten stellt die GADM-Datenbank dar (http://www.gadm.org/). Aufgrund oftmals mangelhafter Lagegenauigkeit sollte von einer Nutzung abgesehen werden.

3 Bestimmung von Siedlungsflächen

Die Ausführungen in Kapitel 2 machen deutlich, dass freie Geodaten für Ortspositionen oftmals ohne eine weitere Verarbeitung nicht geeignet sind, qualitativ hochwertig Ortssymbole zu platzieren. In Kapitel 4 wird dargestellt werden, wie dies mit Hilfe von Siedlungsflächen erreicht werden kann. Daher sollen hier zuvor zwei Ansätze vorgestellt werden, wie man solche Siedlungsflächen im globalen Maßstab erhalten kann.

3.1 OSM-Daten

In (BRINKHOFF 2016) wurde ein Verfahren vorgestellt, das erlaubt, aus OpenStreetMap-Daten Siedlungsflächen abzuleiten. Der Ansatz ist mehrstufig: Im ersten Schritt werden die Landnutzungsflächen („landuse") extrahiert, die die Werte „residential", „industrial", „commercial", „retail" oder „garages" aufweisen. Für Staaten wie Deutschland erhält man damit schon sehr gute Informationen über Siedlungsflächen, während bei anderen Ländern kaum Informationen vorliegen. Flächen, deren Ränder eine geringe Dichte an Stützpunkten aufweisen, werden eliminiert, da solche Objekte oftmals fehlerhaft oder nur sehr ungenau sind. Flächen mit geringen Abstand werden miteinander verschmolzen, so dass Schneisen, die beispielsweise durch Straßen in den Daten vorliegen, eliminiert werden (vgl. Abbildung 3).

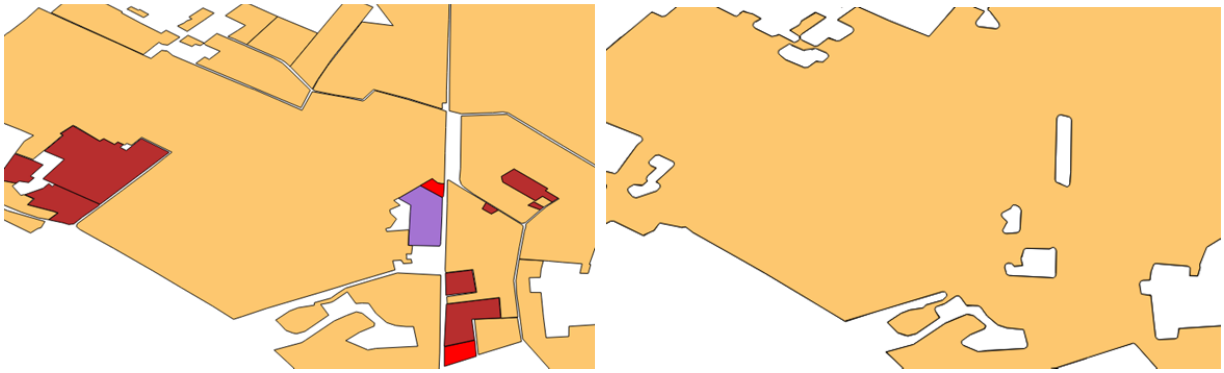


Abb. 3: Verschmelzen von Landnutzungsflächen (links: Flächen vor dem Verschmelzen; Färbung gemäß dem Attributwert von „landuse"; rechts: verschmolzene Flächen)

Im zweiten Schritt werden Gebäudedaten („building") extrahiert, die nicht bereits vollständig von den im ersten Schritt erzeugten Flächen überdeckt sind. Benachbarte Gebäude werden über eine Pufferbildung zusammengeführt, kleine Löcher entfernt und das Ergebnis generalisiert (vgl. Abbildung 4, links).

Im dritten Schritt werden die Straßen („highway") verarbeitet, die die Attributwerte „residential", „living_street" oder „unclassified" aufweisen und nach Differenzbildung mit den Ergebnissen aus den ersten beiden Schritten verbleiben. Benachbarte Straßen werden über eine Pufferbildung zusammengeführt, kleine Löcher entfernt und das Ergebnis generalisiert (vgl. Abbildung 4, rechts).

Abb. 4: Selektion von Gebäuden bzw. Straßen (links/rechts; jeweils grün) und deren Verschmelzen nach Pufferbildung (braune Flächen)

Abschließend werden die Ergebnisse zusammengeführt, miteinander verschmolzen, dabei entstandene kleine Löcher entfernt und das Ergebnis – je nach Anwendungszweck – ggf. weiter generalisiert.
Eine vorläufige Evaluation hat gezeigt, dass die resultierenden Siedlungsflächen eine recht gute Qualität aufweisen, allerdings Industrieflächen oftmals fehlen. Letzteres ist allerdings für die hier verfolgte Aufgabe kein Manko, da ein Ortssymbol in der Regel nicht über Industriegebieten positioniert werden sollte.

3.2 Global Urban Footprint

Ein alternativer Ansatz zur Ableitung globaler Siedlungsflächen ist die Nutzung von Fernerkundungsdaten. Beispiele hierfür sind die Nutzung von Daten des Landsat 8 OLI (Operational Land Imager) (BHATTI & TRIPATHI 2014), von radiometrischen Daten von ASTER (Advanced Spaceborne Thermal Emission and Reflection) (MIYAZAKI et al. 2014) und von Bildern des Landsat TM/ETM+ (ZHANG et al. 2014).
Der in diesem Zusammenhang aktuell relevanteste Datensatz ist der „Global Urban Footprint“ (GUF), der aus Radardaten abgeleitet worden ist (ESCH et al. 2012; ESCH et al. 2013). Dazu wurden insgesamt 180.000 Einzelaufnahmen aus den Jahren 2010 bis 2013 der beiden Radarsatelliten TerraSAR-X und TanDEM-X verarbeitet und analysiert. Die Erdoberfläche wird in besiedelte und unbesiedelte Bereiche untergliedert. Die resultierenden Datensätze wurden Anfang 2017 von dem DLR in zwei Auflösungen (12m und 84m) für die wissenschaftliche bzw. nichtkommerzielle Nutzung endgeldfrei zur Verfügung gestellt (http://www.dlr.de/guf/). Abbildung 5 zeigt die Siedlungsflächen mit der Auflösung 12m im Bereich der Innenstadt von Bremen.

3.3 Vergleich

Ein systematischer Vergleich der beiden vorgestellten Datensätze steht noch aus. Erste Untersuchungen zeigen, dass in entwickelten Staaten beide Ansätze ähnliche Siedlungsflächen erkennen. Abbildung 6 illustriert dies für die Stadt Oldenburg mit Umgebung.

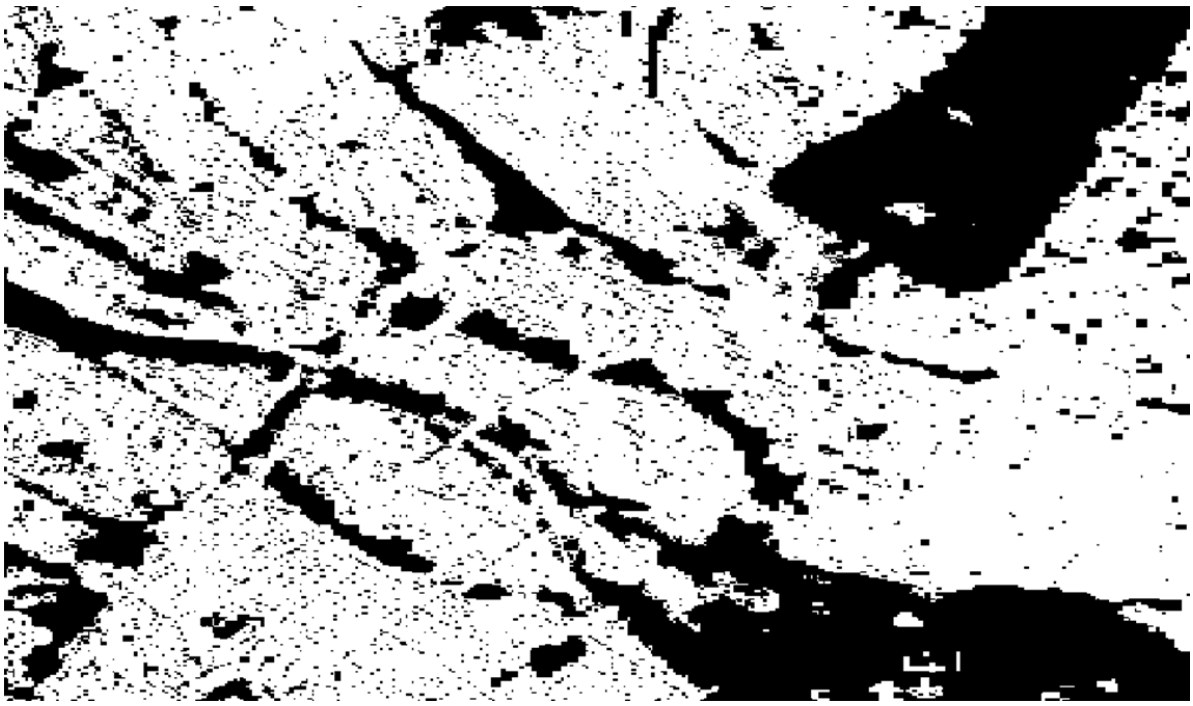


Abb. 5: Ausschnitt aus dem Global Urban Footprint (GUF) im Bereich Bremen-Innenstadt in Auflösung 12m: weiße Pixel entsprechen Siedlungsflächen (© DLR 2016)

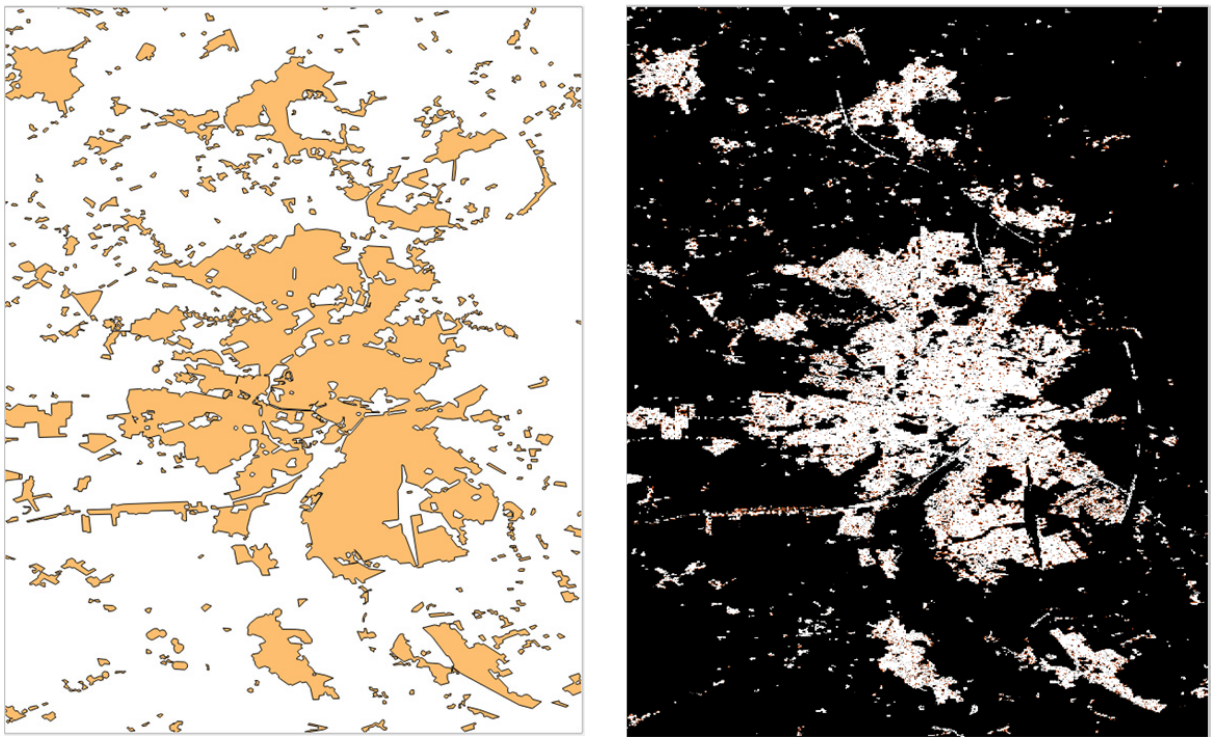


Abb. 6: Vergleich der Siedlungsflächen aus OSM-Daten (links) und GUF-Daten (rechts) im Bereich Stadt Oldenburg und Umgebung

In Entwicklungsländern werden im GUF deutlich mehr Siedlungsflächen erkannt. Abbildung 7 macht dies für den Großraum Lagos deutlich.

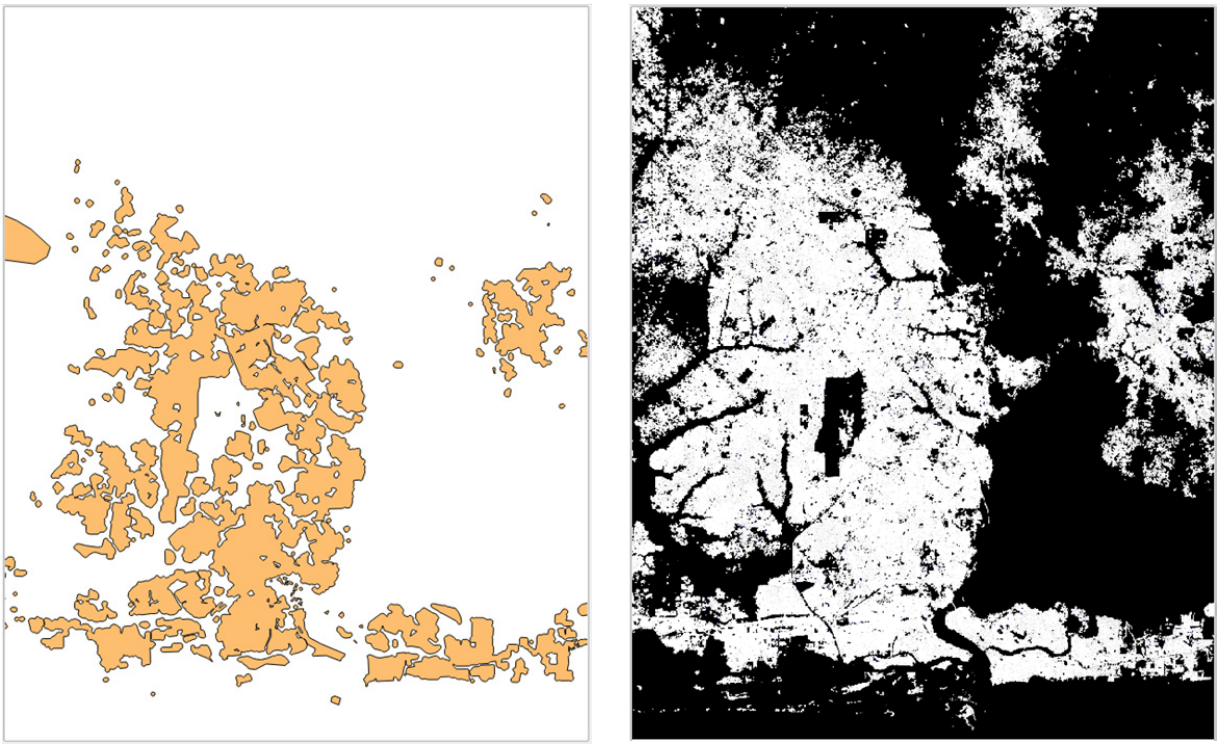


Abb. 7: Vergleich der Siedlungsflächen aus OSM-Daten (links) und GUF-Daten (rechts) im Großraum Lagos

4 Bestimmung von Ortspositionen

Ausgangspunkt für die Bestimmung der Ortspositionen sind Ortsflächen (vgl. Abs. 2.2) und Siedlungsflächen (vgl. Abs. 3). Auf deren Basis kann leicht ein Algorithmus skizziert werden, der daraus die Ortspositionen ableitet:

1. Die Ortsflächen werden mit den Siedlungsflächen verschnitten. Hierbei können drei Fälle eintreten: Für eine Ortsfläche entsteht eine Geometry Collection, die (a) kein Polygon, (b) genau ein Polygon oder (c) mehrere Polygone enthält. Im Fall (c) wird das flächenmäßig größte Polygon als Teilergebnis bestimmt.
2. In das Polygon aus Schritt 1, das ggf. Löcher besitzt, wird geeignet zentriert die Ortsposition gelegt.

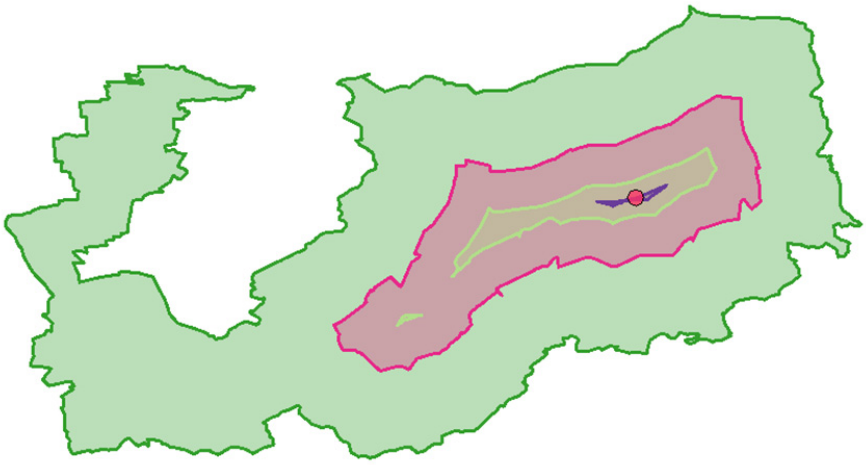
Algorithmische Lösungen für Schritt 2 werden im folgenden Abschnitt behandelt. In der geschilderten Form führt der Algorithmus für reale Daten oftmals zu offenkundig ungeeigneten Ergebnissen, die auf die speziellen Eigenschaften der Eingabedatensätze zurückzuführen sind. Abschnitt 4.2 erörtert dies näher. Erste Ergebnisse werden im dritten Unterabschnitt vorgestellt.

4.1 Positionierung

Falls ein Datensatz mit Ortspositionen zur Verfügung steht, dessen Positionen eine variierende Lagequalität aufweisen (was bei OSM-Daten mitunter der Fall ist), kann folgender Vorverarbeitungsschritt ausgeführt werden: Falls eine Ortsposition aus diesem Datensatz in dem im Schritt 1 berechneten Polygon liegt, ist diese Position das Ergebnis für Schritt 2.

Falls es einen solchen Datensatz nicht gibt oder der Vorverarbeitungsschritt zu einem negativen Ergebnis führt, ist es nun erforderlich, für ein Polygon mit Löchern eine geeignete Position als Zentrum zu bestimmen. Die Berechnung des Polygonschwerpunkts oder des Mittelpunkts des minimal umgebenden Rechtecks sind zwar mit linearen Aufwand auszuführen, führen aber offenkundig zu schlechten Ergebnissen. Für die Bestimmung der Ortsposition existieren in der Literatur verschiedene Ansätze, die exakt oder approximativ das „visuelle Zentrum“ eines Polygons bestimmen. Beispiele sind Verfahren, die das Zentrum unter Nutzung von Voronoi-Diagrammen (ASANO & TOUSSAINT 1986), Polygon-Skeletten (POLLACK et al. 1989) oder Quadtrees (AGAFONKIN 2016) berechnen. GARCIA-CASTELLANOS & LOMBARDO (2007) entwickelten ein approximatives Verfahren, das ausgehend von einer Küstenlinie (= Polygon) iterativ jeweils verkleinerte Flächen berechnet. Dieser Ansatz lässt sich mit einer GIS-Programmbibliothek leicht über eine inverse Pufferberechnung (d.h. die Berechnung einer Pufferzone mit negativer Distanz) realisieren. Falls nach einem Iterationsschritt keine Fläche verbleibt, muss die absolute Distanz verringert und der Schritt erneut ausgeführt werden. Falls das Ergebnis eines Iterationsschritts ein Multipolygon ist, wird nur das flächenmäßig größte Polygon im nächsten Iterationsschritt verwendet. Falls der Flächeninhalt des resultierenden Polygons kleiner als ein vorgegebener Wert ist (z.B. 1% relativ zum Ausgangspolygon), endet das Verfahren und gibt den Polygonschwerpunkt als Ergebnis zurück. Neben der einfachen Implementierbarkeit kann eine für den hier vorgesehenen Anwendungsfall hinreichende Performanz und Qualität des Resultats konstatiert werden. Abbildung 8 zeigt eine Beispielsberechnung in drei Iterationsschritten.

Abb. 8: Berechnung des visuellen Zentrums über inverse Pufferbildung



4.2 Erforderliche Anpassungen

Die Abbildungen 6 und 7 zeigen, dass beide potenziellen Datensätze für Siedlungsflächen eine große Anzahl kleinerer Löcher und Einbuchtungen aufweisen. Diese gehen durch die Verschneidung nicht verloren, so dass sie auch im Ausgangspolygon für die Bestimmung der Ortsposition enthalten sind und – wie in Abbildung 9 (links) zeigt – die Berechnung beeinflussen. Daher sollten kleinere Löcher (Abb. 9, Mitte) und ggf. auch Einbuchtungen (Abb. 9, rechts) vorher entfernt werden.

Abb. 9: Die Berechnung des visuellen Zentrums mit den ursprünglichen Siedlungsflächen (links), nach Entfernung von kleineren Löchern (Mitte) und nach Entfernung von kleineren Löchern und Einbuchtungen (rechts) am Beispiel Oldenburg

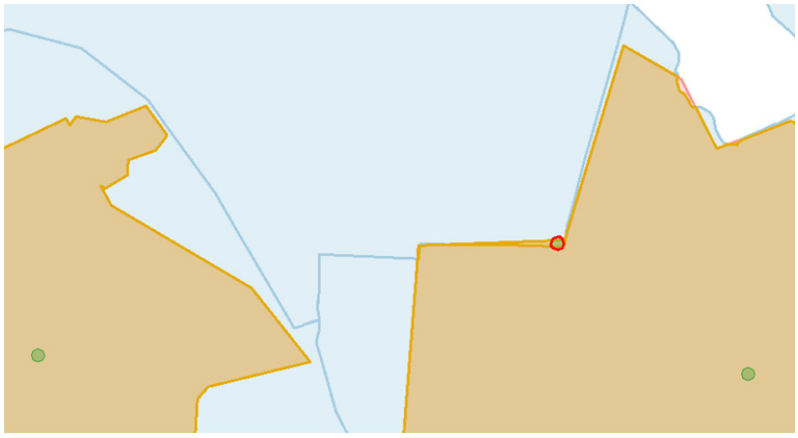


Abb. 10: Der rote Punkt würde ohne geeignete Vorverarbeitung als Ortsposition für die obere hellblaue Fläche bestimmt werden.

Eine weitere Problematik resultiert daraus, dass sowohl die Siedlungsflächen als auch frei verfügbare administrative Grenzen Lageabweichungen von der tatsächlichen Situation aufweisen. Dadurch kann das Resultat insbesondere bei Orten, deren Siedlungsflächen in den Siedlungsdaten fehlen, und die an einen Ort grenzen, der eine an die Ortsgrenzen reichende Siedlungsfläche

aufweist, verfälscht werden. Abbildung 10 zeigt dies an einem Beispiel. Daher sollten die Ortsflächen vor der Bestimmung der Ortspositionen mindestens mit einem Abstand verkleinert werden, der die Ungenauigkeiten der beteiligten Daten berücksichtigt.

4.3 Ergebnisse

In diesem Abschnitt soll das Ergebnis zunächst anhand der Großgemeinden (Municipios) in Brasilien erörtert werden. Ausgangsdaten sind Flächengeometrien des Instituto Brasileiro de Geografia e Estatistica für die 5.570 brasilianischen Municipios. Für 1.106 Gemeinden (=20%) konnte eine und für 3.609 Gemeinden (=65%) konnten mehrere Siedlungsflächen aus den OSM-Daten zugeordnet werden. Die Berechnungszeit für die Positionierung aller Orte lag bei etwa 40 Sekunden. Abbildung 11 zeigt, dass der vorgestellte Ansatz bei großflächigen Gemeinden mit geringen Siedlungsanteilen zu ansprechenden Ergebnissen führt.

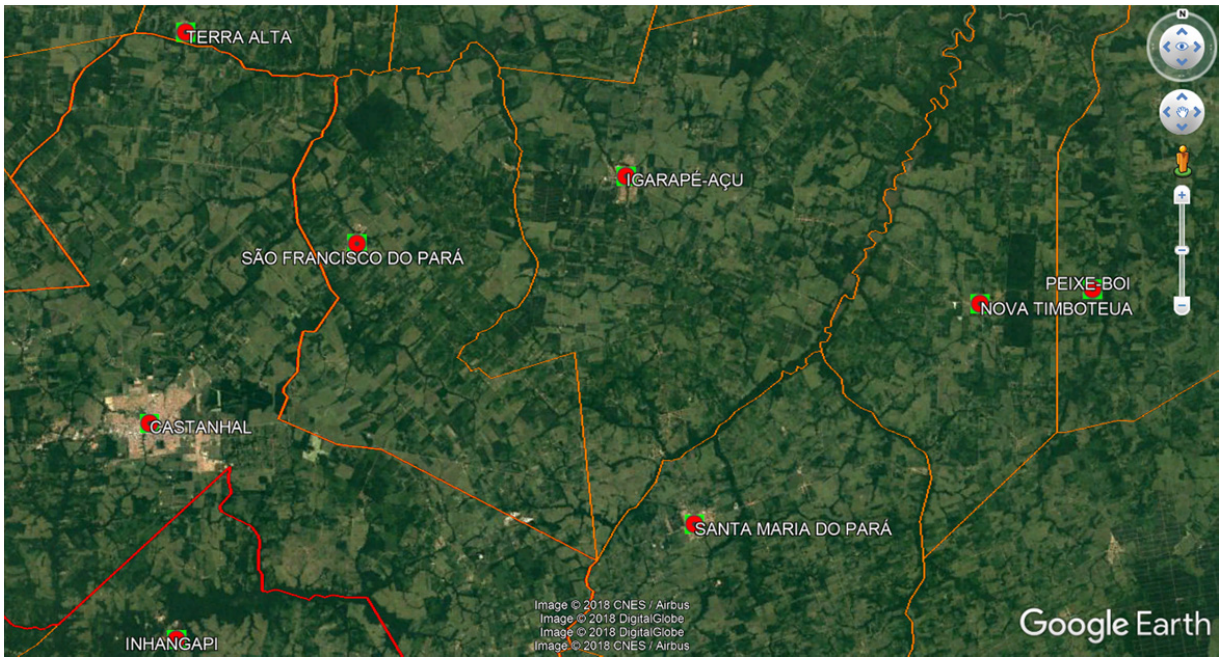


Abb. 11: Ortspositionen für brasilianische Großgemeinden

Für kleinräumigere israelische Ortschaften (Localities) war die Bestimmungsrate deutlich höher: Hier konnten für 1.094 Orte 1.088 Ortspositionen (=99,5%) ermittelt werden. Erste Untersuchungen für indische Dörfer zeigen, dass die Siedlungsflächen aus den OSM-Daten zu unvollständig sind, so dass hierfür der Global Urban Footprint genutzt werden sollte.

5 Fazit & Ausblick

In dem Beitrag wurde ein Ansatz vorgestellt, der die Positionierung von Orten in webbasierten Karten verbessern soll. Ausgangspunkt sind einerseits (durchaus ungenaue) Flächenbeschreibungen der Orte und andererseits Informationen über Siedlungsflächen, die aus OpenStreetMap-Daten oder dem Global Urban Footprint des DLR gewonnen werden können. Neben der Ver-

schneidung und der geeigneten Bestimmung eines Ortszentrums ist dabei auch die Behandlung von Sonderfällen zu beachten, die sich aus den Eigenschaften der Siedlungsflächen und der Ungenauigkeit der Ausgangsdaten ergeben. In ersten Untersuchungen konnte demonstriert werden, dass der Ansatz für einen hohen Anteil von Orten automatisiert eine geeignete Ortsposition bestimmen kann.
Einer der nächsten Schritte ist die Durchführung eines systematischen Vergleichs zwischen den OSM- und GUF-Siedlungsdaten. Zudem stehen noch ausführlichere Untersuchungen hinsichtlich der Nutzung des Global Urban Footprints aus.

6 Literaturverzeichnis

AGAFONKIN, V., 2016: A new algorithm for finding a visual center of a polygon. https://blog.mapbox.com/a-new-algorithm-for-finding-a-visual-center-of-a-polygon-7c77e6492fbc

ASANO, T. & TOUSSAINT, G.T., 1986: Computing the geodesic center of a simple polygon. Perspectives in Computing: Discrete Algorithms and Complexity, Proceedings of Japan-US Joint Seminar, 65-79.

BHATTI, S.S. & TRIPATHI, N.K., 2014: Built-up area extraction using Landsat 8 OLI imagery. GIScience & Remote Sensing, **51**(4), 445-467.

BRINKHOFF, T., 2015: Offene Geodaten: Lage von Orten im Vergleich. Tagungsband FOSSGIS 2015, Münster, 2015, 18-24.

BRINKHOFF, T., 2016: Open Street Map Data as Source for Built-up and Urban Areas on Global Scale, XXIII Congress of the International Society for Photogrammetry and Remote Sensing 2016, Prague, Czech Republic. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B4, 2016, 557-564; doi:10.5194/isprs-archives-XLI-B4-557-2016

BRINKHOFF, T., 2017: Supporting Dynamic Labeling in Web Map Applications. Proceedings 20th AGILE Conference on Geographic Information Science, Wageningen, the Netherlands.

ESCH, T.; TAUBENBÖCK, H.; ROTH, A.; HELDENS, W.; FELBIER, A.; THIEL, M.; SCHMIDT, M.; MÜLLER, A. & DECH, S., 2012: TanDEM-X mission-new perspectives for the inventory and monitoring of global settlement patterns. Journal of Applied Remote Sensing, **6**(1), 061702; doi:10.1117/1.JRS.6.061702

ESCH, T.; MARCONCINI; M., FELBIER, A.; ROTH, A.; HELDENS, W.; HUBER, M.; SCHWINGER, M.; TAUBENBÖCK, H.; MÜLLER, A. & DECH, S., 2013: Urban Footprint Processor – Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission. IEEE Geoscience and Remote Sensing Letters, **10**(6), 2013, 1617-1621; doi:10.1109/LGRS.2013.2272953

GARCIA-CASTELLANOS, D. & LOMBARDO U., 2007: Poles of Inaccessibility: A Calculation Algorithm for the Remotest Places on Earth. Scottish Geographical Journal, **123**(3), 227-233.

MIYAZAKI, H., SHAO, X., IWAO, K. & SHIBASAKI, R., 2014: Development of a Global Built-Up Area Map Using ASTER Satellite Images and Existing GIS Data. In: Weng, Q. (ed.), Global Urban Monitoring and Assessment through Earth Observation, CRC Press, 121-142.

POLLACK, R., SHARIR, M. & ROTE, G., 1989: Computing the Geodesic Center of a Simple Polygon. Discrete & Computational Geometry, **4**, 611-626

ROTH, R.E., DONOHUE, R.G., SACK, C.M., WALLACE, T.R. & BUCKINGHAM, T., 2015. A process for keeping pace with evolving web mapping technologies. Cartographic Perspectives, **78**, 25-52; doi:10.14714/CP78.1273

ZHANG, J., LI, P. & WANG, J., 2014. Urban Built-Up Area Extraction from Landsat TM/ETM+ Images Using Spectral Information and Multivariate Texture. Remote Sensing, **6**(8), 7339-7359.

Generalisierung mittels Deep Learning (CNN) am Beispiel der Straßenextraktion aus GPS-Trajektorien

FRANK THIEMANN[1], SERCAN CAKIR[1] & FLORIAN POLITZ[1]

Zusammenfassung: Convolutional Neural Networks (CNN) sind in der Lage, wesentliche Merkmale der Eingabedaten zu erlernen. Mit dieser Arbeit wollten wir die Eignung von CNN zur Generalisierung am Beispiel der Extraktion von Straßenachsen aus GPS-Trajektorien testen. Als Eingangsdaten wurden aus den Trajektorien Dichtekarten abgeleitet. Als Referenzdaten haben wir das zugrundeliegende Straßennetz gerastert. In drei Experimenten wurden CNN auf simulierten Kfz- und realen Fahrrad-Trajektorien trainiert und dann auf die jeweiligen anderen Datensätze übertragen. Das Paper beschreibt die beobachteten Probleme.

1 Einleitung

GPS-Trajektorien werden an vielen Stellen erzeugt. Navigationssysteme, Smartphones, Sportuhren, etc. sind mit einfachen GPS-Empfängern ausgestattet, die die Position auf wenige Meter genau bestimmen können. Diese Positionsdaten werden als Zeitreihen mit wenigen Sekunden Auflösung aufgezeichnet und können auf diversen Internetplattformen gesammelt und geteilt werden. Frei verfügbare Trajektorien gibt es zum Beispiel im Openstreetmap-Projekt (OSM, openstreetmap.org). Sport- und Fitnessplattformen sammeln ebenfalls Trajektorien und bieten vielfältige Auswertungen an. Sehr beliebt sind dabei Heatmaps (Dichtekarten) aus den gesammelten Trajektorien. Unter Anderem ermöglichen Heatmaps häufig benutzte und damit vermeintlich beliebte Routen zu entdecken. Des Weiteren können sie auch Wegeverläufe sichtbar machen, die ansonsten nicht in Luftbildern und Karten erkennbar bzw. verzeichnet sind. So stellt die Fitnesstracking-Plattform Strava.com Heatmaps für OSM bereit, die als Grundlage zur Wegeerfassung dienen können.
In unserer Arbeit wollten wir testen, inwieweit sich die Extraktion des Straßennetzes aus Heatmaps mittels Deep Learning automatisieren lässt. Dabei fassen wir das Extraktionsproblem als Generalisierung auf, wobei die Straßenachsen die vereinfachte Repräsentation der Trajektorien bzw. der daraus abgeleiteten Heatmap darstellt.

2 Verwandte Arbeiten

Deep Convolutional Neural Networks (tiefe faltende neuronale Netze, CNN) werden seit Jahren erfolgreich für die semantische Klassifikation und Objektdetektion in Bildern eingesetzt (KRIZHEVSKY et al. 2012; GIRSHICK et al. 2013). In Form von Autoencodern (GOODFELLOW et al. 2016) können solche Netze die wesentlichen Merkmale der Eingangsdaten erlernen und diese zur Dimensionsreduktion bzw. zum Komprimieren der Daten nutzen. Dabei werden unwesentliche

[1] Leibniz Universität Hannover, Institut für Kartographie und Geoinformatik, Appelstraße 9 a, D-30167 Hannover, E-Mail: [Frank.Thiemann, Politz]@ikg.uni-hannover.de

Merkmale unterdrückt. Diese Eigenschaft von Autoencodern legt nahe, dass sie auch zur Generalisierung im kartographischen Sinne verwendet werden können.
SIMO-SERRA et al. (2016) haben bereits gezeigt, dass CNN zum Vereinfachen von groben Skizzen verwendet werden können. Die Autoren nutzen dabei ein CNN mit Autoencoder-Architektur. Anstatt, wie bei Autoencodern üblich, die Originalbilder sowohl als Eingangs- als auch als Referenzdaten zu nutzen, verwenden sie als Referenzdaten die vereinfachten Versionen der Eingangsbilder.

3 Methodik

Ziel unserer Arbeit ist es, die Eignung von CNN zur Generalisierung zu testen. Dazu adaptieren wir den Ansatz SIMO-SERRA et al. (2016) auf die Extraktion der Straßenachsen aus Heatmaps. Als Eingangsdaten dienen Dichtekarten von sowohl simulierten als auch realen GPS-Trajektorien. Als Referenzdaten werden die zugrundeliegenden Fahrbahn- bzw. Fahrwegachsen genutzt.
Für das Training des Netzes werden weitgehend fehlerfreie Referenzdaten benötigt. Eine direkte Verwendung von bestehenden Straßendatensätzen ist nur bedingt möglich, da das Straßennetz nicht überall ausreichend oft befahren wird, um eine Mindestanzahl an Trajektorien je Straße zu gewährleisten. Somit müssen zunächst nicht befahrene Straßenteile aus den Referenzdaten gelöscht werden. Auf Grund der geometrischen Ungenauigkeit, sowohl der Trajektorien und der Straßendatensätze, ist eine automatische Zuordnung der befahrenen Straßen nicht immer und nur mit großem Aufwand möglich.
Um die Korrektheit der Trainingsdaten zu gewährleisten, haben wir zunächst Trajektorien auf dem aus Openstreetmap extrahierten Straßennetz von Hannover simuliert. Des Weiteren haben wir versucht, Referenzdaten zu realen Fahrrad-Trajektorien zu generieren, was uns wegen der beschriebenen Probleme jedoch nicht in der benötigten Genauigkeit gelang.
Mit beiden Datensätzen haben wir ein CNN trainiert und dieses dann auf dem jeweils anderen Datensatz getestet. In einem weiteren Test haben wir die gelernten Modelle auf einen Datensatz aus realen Kfz-Trajektorien angewendet und die Ergebnisse visuell beurteilt.

3.1 Eingangs- und Referenzdatensätze

Als Quelle für die Referenzdaten haben wir das Straßennetz aus Openstreetmap verwendet. Aus den Heatmaps von Fahrrad-Trajektorien von Strava (heatmaps.strava.com) haben wir Dichtekarten generiert. Des Weiteren haben wir zeitlich hochaufgelöste Kfz-Trajektorien von der Stadt Chicago verwendet (AHMED et al., 2015).
Die Erstellung eines Referenzdatensatzes gestaltet sich für größere Straßendatensätze äußerst aufwendig. Um einen fehlerfreien Referenzdatensatz zu erhalten, muss das Straßennetz passend zu den Trajektorien vektorisiert werden. Selbst bei der Verwendung von existierenden Straßendatensätzen wie z. B. ATKIS oder OSM, ist eine erhebliche manuelle Nachbearbeitung notwendig. Straßen und Straßenteile, für die keine Trajektorien vorliegen, müssen entfernt werden, da sie nicht prädiziert werden können. Ebenso müssen fehlende Straßen oder Fahrspuren ergänzt und Geometrien ggf. korrigiert werden. Insbesondere für die Fahrrad-Trajektorien gestaltete sich die Zuordnung zum Wegenetz schwierig, da alternative Spuren wie z. B. Radweg und rechter

Fahrbahnrand oft sehr eng nebeneinanderliegen. Zudem ist die Vollständigkeit der Radwege in OSM deutlich geringer als die der Fahrbahnachsen.
Um die Vollständigkeit und geometrische Korrektheit der Referenzdaten zu gewährleisten, haben wir Trajektorien auf einem ausgewählten Straßendatensatz simuliert. Dazu wurde zunächst das für Kraftfahrzeuge nutzbare Straßennetz von Hannover extrahiert. Im Folgenden wurden zufällig Knoten im Straßennetz selektiert und ein Routing zwischen den Knoten durchgeführt. Die daraus entstandene Route wurde dann in regelmäßigen Abständen abgetastet. Auf die so erhaltenen Punktkoordinaten wurde anschließend ein zufälliges Rauschen (σ = 0,6 m) und eine zufällige, systematische Verschiebung je Trajektorie (σ = 5 m) addiert.
Zur Verwendung eines CNN müssen die Trajektorien und das Referenzstraßennetz in Bild- bzw. Rasterdaten überführt werden. Aus den Trajektorien wurde dazu mit einem Gaußkern die Dichte berechnet. Die Straßenachsen wurden direkt gerastert (siehe Abb. 1). Als Pixelgröße wurden 3 Meter gewählt, da dies etwa der Genauigkeit der einzelnen Trajektorien entspricht. Bei der gewählten Auflösung sind parallele Fahrbahnachsen im Raster noch unterscheidbar. Zum Rastern eines Radwegenetzes ist diese Auflösung jedoch nicht mehr ausreichend. Hier beginnen einige Achsen, wie z. B. die Achsen von straßenparallelen Radwegen mit der Fahrbahnachse, zu verschmelzen (siehe Abb. 2).

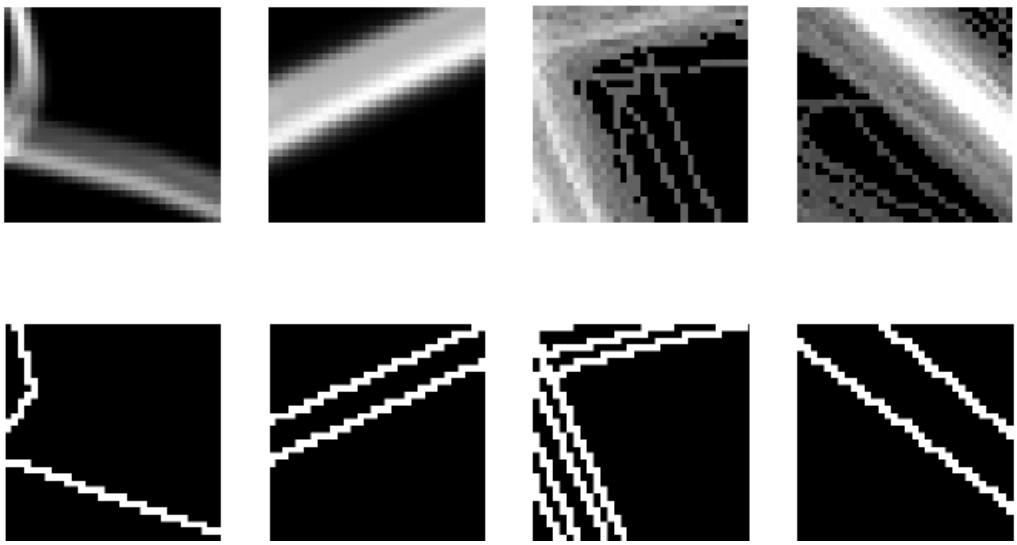


Abb. 1: Simulierte Dichtekarten (oben) und dazugehörige Referenzdaten (unten)

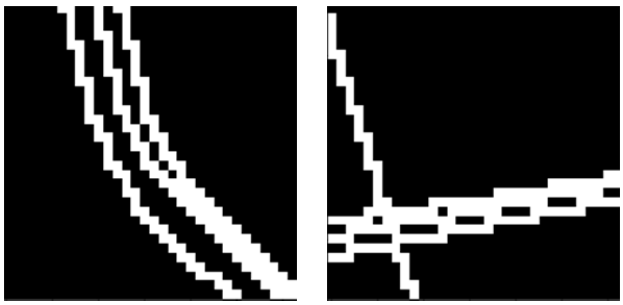


Abb. 2: Radwege- und Fahrbahnachsen verschmelzen beim Rastern mit 3 m Auflösung

3.2 Convolutional Neural Network

Das für unsere Arbeit genutzte Convolutional Neural Network wurde von dem Modell von SIMO-SERRA et al. (2016) durch Reduktion der Parameter abgeleitet. Das vollständige CNN ist in Abb. 3 dargestellt. Die Größe der Eingangsbilder ist dabei frei wählbar, da das Modell auf voll vernetzte Layer verzichtet. Wir haben quadratische Bildpatches mit einer Größe von 32 x 32 Pixel verwendet. Der erste Teil des Netzwerkes fungiert als Encoder und soll die Eingabebilder komprimieren, indem er signifikante Merkmale extrahiert. Statt einer Pooling-Funktion wird durch zwei Down-Convolutions mit einer Schrittweite von 2 das Eingabebild auf 1/16 seiner ursprünglichen Größe reduziert. Der Informationsverlust wird durch Erhöhung der Filterzahl ausgeglichen.

Der mittlere Teil des Netzwerkes extrahiert die grundlegende Information der Straßenachsen mit konstanter Bildgröße. Hier wird die Zahl der Filter bis zur Mitte des Netzes systematisch erhöht und anschließend wieder reduziert. Das CNN schließt mit einem dekodierenden Teil, der das nun generalisierte Bild wieder in die Ausgangsgröße zurückführt. Die Up-Convolutions sind als Up-Sampling gefolgt von einer Faltung mit einer 5 x 5 großen Filtermatrix implementiert.

Die Ergebnisse jeder Faltung werden mit einer nichtlinearen Funktion aktiviert. Dabei wird mit Ausnahme der letzten Schicht eine Rectified Linear Unit-Funktion (ReLU-Funktion) verwendet (NAIR & HINTON, 2010). Um die Ausgabe in das Intervall [0,1] zu normieren wird beim letzten Layer die Sigmoid-Funktion verwendet. Zur Vermeidung von Überanpassung werden die Parameter mit der L2-Norm regularisiert und es wird nach jedem Layer ein Dropout-Layer (SRIVASTAVA et al., 2014) mit einer Dropout-Rate von 10 % eingefügt.

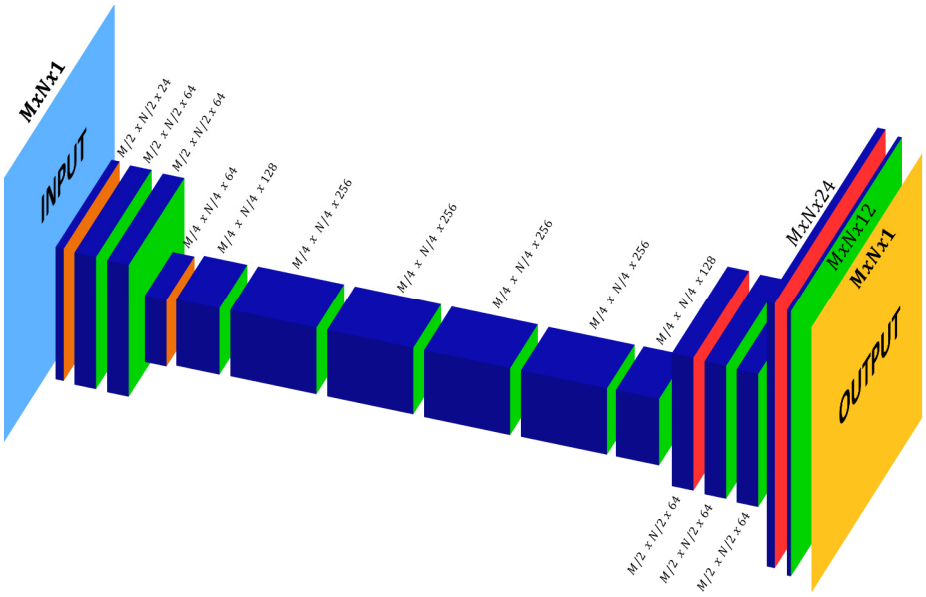


Abb. 3: Verwendetes CNN (Zeilen x Spalten x Kanäle)

Das Modell wird mit den Heatmaps als Eingabe und den gerasterten Straßenachsen als Referenzdaten trainiert. Als Kostenfunktion wurde die binäre Kreuz-Entropie verwendet. Diese ist besonderes für binäre Klassifikationen geeignet. Als weitere Metrik wurde die Kovarianz von Ausgabe- und Referenzbild berechnet. Um auch geringe Lageabweichungen der prädizierten

Straßen zu erlauben, wurden nur die inneren 30 × 30 Pixel als gleitendes Fenster über das 32 × 32 Pixel Referenzbild geschoben und die maximale Kovarianz verwendet.
Um Probleme mit der Einstellung der Lernrate zu umgehen, wurde die Optimierung mit dem Adadelta-Verfahren (ZEILER 2012) durchgeführt.

Tab. 1: CNN-Struktur. Die Farben wurden analog zu Abb. 3 gewählt.

Art der Faltung	Größe (Zeilen x Spalten x Kanäle)	Schrittweite	Filtergröße
Eingabe	M x N x 1		-
down-conv-1	M/2 x N/2 x 24	2 x 2	5 x 5
flat-conv-1	M/2 x N/2 x 64	1 x 1	3 x 3
flat-conv-2	M/2 x N/2 x 64	1 x 1	3 x 3
down-conv-2	M/4 x N/4 x 64	2 x 2	3 x 3
flat-conv-3	M/4 x N/4 x 128	1 x 1	3 x 3
flat-conv-4	M/4 x N/4 x 256	1 x 1	3 x 3
flat-conv-5	M/4 x N/4 x 256	1 x 1	3 x 3
flat-conv-6	M/4 x N/4 x 256	1 x 1	3 x 3
flat-conv-7	M/4 x N/4 x 256	1 x 1	3 x 3
flat-conv-8	M/4 x N/4 x 128	1 x 1	3 x 3
up-conv-1	M/2 x N/2 x 64	2 x 2	3 x 3
flat-conv-9	M/2 x N/2 x 64	1 x 1	3 x 3
flat-conv-10	M/2 x N/2 x 64	1 x 1	3 x 3
up-conv-2	M x N x 24	2 x 2	3 x 3
flat-conv-11	M x N x 12	1 x 1	3 x 3
flat-conv-12	M x N x 1	1 x 1	3 x 3

3.3 Experimente und Ergebnisse

Damit sich Trainings- und Validierungsdaten nicht überlappen und möglichst ähnliche Straßennetzstrukturen abdecken, wurde das Gebiet in adjazente, nicht überlappende Kacheln von 32 x 32 Pixeln aufgeteilt. Kachelpaare, bei denen Eingabe- oder Referenzbild keine Straßenpixel enthielten, wurden vor dem Training aus den Datensätzen entfernt. Etwa 80 % der Daten wurden fürs Training und die restlichen 20 % zur Validierung verwendet.

3.3.1 Training mit simulierten Trajektorien

Für das erste Experiment wurde das Netz mit den simulierten Trajektorien trainiert. Wir haben 3772 Trainings- und 1000 Validierungsbilder verwendet. Das Training erfolgte über 250 Epochen. Die Kostenfunktion auf den Trainingsdaten fällt dabei stetig, d. h. die Optimierung konvergiert. Die Kosten im Validierungsdatensatz erreichen hingegen bereits nach ca. 15 Epochen ihr Minimum und beginnen dann wieder zu steigen. Die Kovarianzen steigen sowohl auf dem Trainings- als auf dem Validierungsdatensatz stetig, wobei der Wert für den Validierungsdaten-

satz deutlich unterhalb des Wertes des Trainingsdatensatzes verbleit. Die Werte für den Validierungsdatensatz lassen sich durch Variation der Dropout-Rate und der Gewichtung der L2-Regularisierung nicht wesentlich verbessern. Die Richtigkeit und die Vollständigkeit der prädizierten Straßenachsen auf dem gesamten Datensatz betragen rund 90 % (siehe Tab. 2). Der vermeintlich gute Wert ist jedoch auf eine Überanpassung an die Trainingsdaten zurückzuführen.

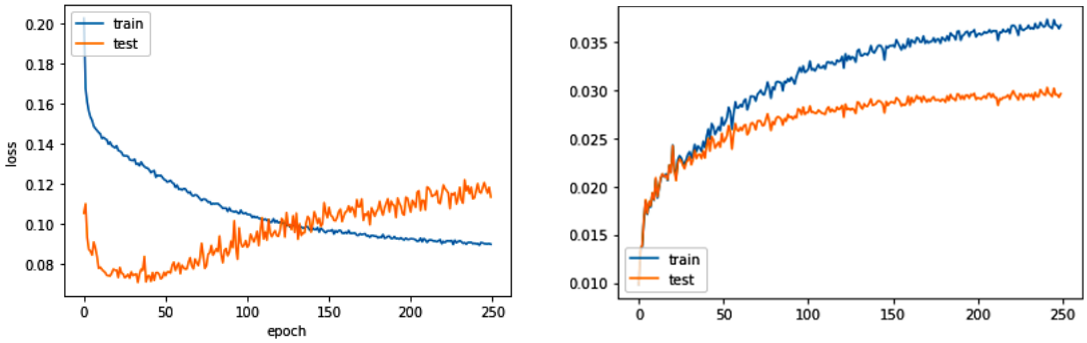


Abb. 4: Verlauf der Kostenfunktion (Kreuz-Entropie) (links) und der Kovarianz (rechts) über die 250 Trainingsepochen. Die absoluten Werte der Entropie von Trainings- und Testdaten lassen sich nicht vergleichen, da beim Training ein Dropout von 10 % verwendet wurde, welches beim Validieren nicht zur Anwendung kommt.

3.3.2 Training mit Fahrrad-Heatmap

Im zweiten Experiment haben wir als Eingabedaten die Heatmaps von Strava verwendet und als Referenzdaten das zugrundeliegende Radwegenetz aus OSM herangezogen. Es wurden die selben Hyperparameter wie in Kap. 3.3.1 verwendet. Auch hier konvergiert die Optimierung. Die Kovarianz von Prädiktion und Referenzdaten fällt jedoch geringer als beim Training mit den simulierten Trajektorien aus. Die Kovarianzen von Trainings- und Validierungsdaten unterscheiden sich weniger (siehe Abb. 5).

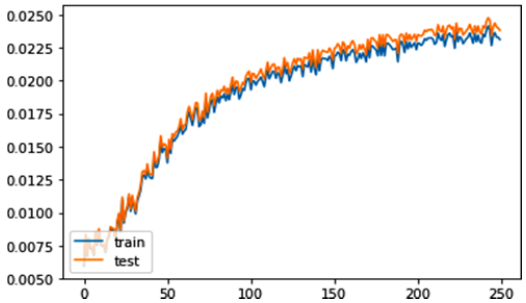


Abb. 5: Verlauf der Kovarianz über 250 Trainingsepochen beim Training mit der Fahrrad-Heatmap.

Die erzielte Richtigkeit und Vollständigkeit auf dem Strava-Datensatz (Trainings- und Testdaten zusammen) beträgt lediglich 38 % und 33 %. Dies kann auf die geringere Genauigkeit der Referenzdaten zurückzuführen sein, denn bereits beim Erstellen der Referenzdaten fiel auf, dass die OSM-Wegeachsen und die dazugehörigen Stellen in der Heatmap oft nicht gut zusammenpassen. Beim Testen des trainierten Netzwerkes auf den simulierten Daten wird lediglich eine Richtig-

keit von 17 % und eine Vollständigkeit von 10 % erreicht. Das auf Fahrradtrajektorien trainierte CNN lässt sich folglich nicht auf die simulierten Daten übertragen.

3.3.3 Training mit simulierten und realen Daten

Im dritten Experiment haben wir beide Datensätze, also die simulierten Trajektorien sowie die realen Strava-Heatmaps, gemeinsam als Trainingsdaten für das CNN verwendet. Es standen insgesamt rund 7000 Eingangsbilder für das Training zur Verfügung. Die erzielte Richtigkeit beträgt hier 53 %, und die Vollständigkeit 48 %. Die Kovarianz der Validierungsdaten ist in diesem Experiment nur geringfügig besser als beim zweiten. Auch hier bleibt der Wert für den Validierungsdatensatz deutlich kleiner als für den Trainingsdatensatz, was auf eine Überanpassung an den Trainingsdatensatz hindeutet.

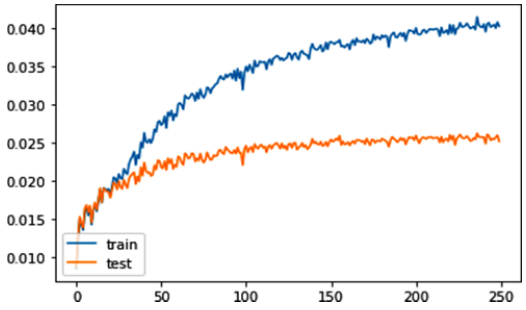


Abb. 6: Verlauf der Kovarianz über 250 Trainingsepochen mit kombiniertem Datensatz

3.3.4 Vergleich der Ergebnisse

Die trainierten Netzwerke wurden jeweils auf die einzelnen und den kombinierten Datensatz angewendet, um die Gesamtleistung zu messen. Die besten Ergebnisse wurden mit den simulierten Daten auf dem CNN mit den simulierten Trajektorien erzielt. Die Struktur der Trajektorien ist hier am einfachsten und die Genauigkeit der Referenzdaten am höchsten. Das so trainierte Netzwerk lässt sich jedoch nicht auf die realen Fahrradtrajektorien übertragen. Eine Ursache hierfür könnte sein, dass Fahrradfahrer ein anderes Spurwahlverhalten haben als Kraftfahrzeuge. Radfahrer fahren in der Regel immer am rechten Fahrbahnrand oder rechts neben der Fahrbahn auf dem Radweg. Auf einigen Plätzen können Radfahrer ihre Fahrspuren vollkommen frei wählen. Dieses Verhalten ist in den simulierten Trajektorien jedoch nicht abgebildet, wodurch das so trainierte CNN dieses nicht lernen konnte.
Das Training auf den realen Fahrrad-Heatmaps liefert mit 38 % Richtigkeit und 33 % Vollständigkeit nur schlechte Ergebnisse. Der Referenzstraßensatz passt hier vermutlich zu schlecht zur Heatmap. Das so trainierte Netzwerk konnte ebenfalls nicht auf den anderen Datensatz übertragen werden. Beim Training auf gemischten Trainingsdaten ergeben sich für die Fehlermaße geringfügig schlechtere Ergebnisse als beim Übertragen des CNN mit den simulierten Trajektorien. Visuell betrachtet wirkte das Ausgabebild jedoch plausibler.

Tab. 2: Vollständigkeit und Richtigkeit der Straßenpixel im Überblick

Validierung mit:	Training mit:	simuliert	Fahrrad	kombiniert
simulierten Trajektorien	Richtigkeit	0,90	0,18	0,54
	Vollständigkeit	0,89	0,37	0,63
realen Fahrrad-Heatmaps	Richtigkeit	0,17	0,38	0,28
	Vollständigkeit	0,10	0,33	0,22
kombinierten Datensätzen	Richtigkeit			0,53
	Vollständigkeit			0,48

Auf dem Datensatz von Chicago, welches auf Kfz-Trajektorien basiert, wurde zudem ein Test mit visueller Begutachtung durchgeführt (siehe Abb. 7). Das auf den simulierten Trajektorien trainierte CNN gibt nur ein unvollständiges Straßennetz als Ergebnis aus. Das lässt sich damit erklären, dass für das Training in unserem Referenzdatensatz Straßen, denen nur wenige Trajektorien zuzuordnen waren, gelöscht wurden. Nach dem Training mit realen Fahrrad-Heatmaps ist das Ausgabebild vollständiger. Es werden jedoch auf breiten Straßen mehrere Achsen prädiziert. Nach dem Training mit dem kombinierten Datensatz reduziert sich dieser Effekt. Es ergibt sich hier visuell das beste Ergebnis.

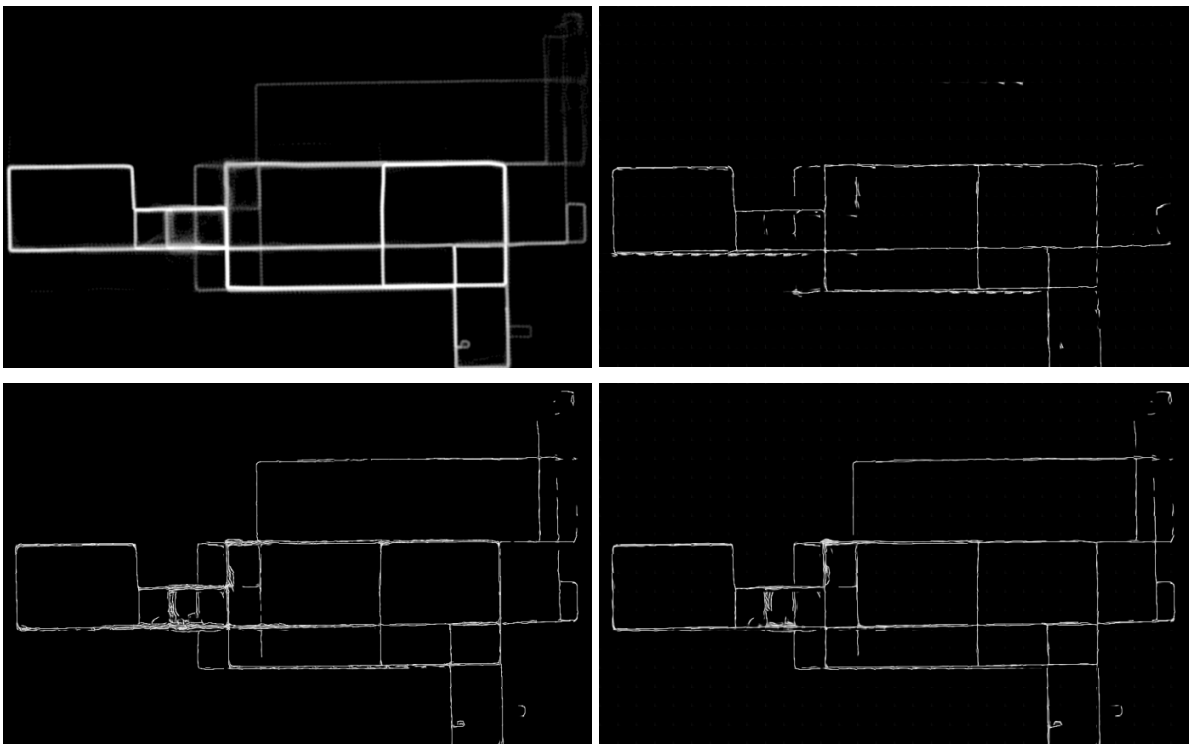


Abb. 7: Eingangsdaten: Heatmap von Chicago (a), Prädiktion (Wahrscheinlichkeit) der Straßenachsen: nach Training mit simulierten Trajektorien (b), mit realer Fahrrad-Heatmap (c) und mit kombiniertem Trainingsdatensätzen (d).

4 Fazit & Ausblick

Die mit dem Ansatz erzielten Ergebnisse sind bisher noch nicht zufriedenstellend. Das Erstellen der Referenzdaten stellte sich als sehr aufwendig heraus. Die Verwendung von bestehenden Straßendatensätzen als Referenz erfordert viel manuelle Nacharbeitung, um fehlende, nicht benutzte oder ungenaue Straßen zu korrigieren. Die Simulation von Trajektorien bedarf aufwendigerer Algorithmen, um ein realistisches Spurwahlverhalten auch auf breiteren Fahrbahnen nachzubilden.
Veränderungen am Modell haben auf die Ergebnisse nur wenig Auswirkung gezeigt. Die Kostenfunktion (Kreuz-Entropie) erreicht auf dem Validierungsdatensatz bereits nach ca. 15 Iterationen ihr Minimum und beginnt dann wieder zu steigen. Das Ausgabebild ist nach 15 Iterationen noch unscharf und wird erst mit größerer Epochenzahl schärfer. Mit zunehmender Schärfe liegt die Prädiktion aufgrund der Lageungenauigkeiten der Trajektorien und des Referenzdatensatz oft geringfügig neben der Referenz. Jegliche Lageabweichung ($\geq$ 1 Pixel) führt bei der Kostenfunktion jedoch zu einem starken Anstieg. Dem kann nur begegnet werden, wenn entweder die Referenzdaten optimal an die Trainingsdaten angepasst werden oder eine tolerantere Kostenfunktion genutzt wird, die kleine Lageabweichungen weniger bestraft. Die vorgestellte gleitende Kovarianz, ist dazu ein erster Ansatz, wenngleich sie nur auf kleinen Patches sinnvolle Ergebnisse liefern kann.
Der Überanpassung könnte künftig durch mehr Trainingsdaten begegnet werden. Eine stärkere Regularisierung und höhere Dropout-Raten führten lediglich zu unschärferen Ausgaben, was einer unsichereren Prädiktion gleichkommt.
Generell stellt der benutzte Ansatz ein interessantes Werkzeug für weitere Probleme zur Ableitung von Achsen bereit. So wollen wir künftig die Anwendbarkeit des Ansatzes zur Extraktion von Gewässer- und Wegeachsen aus digitalen Geländemodellen testen.

5 Literaturverzeichnis

AHMED, M., KARAGIORGOU, S., PFOSER, D. & WENK, C., 2015: A Comparison and Evaluation of Map Construction Algorithms. GeoInformatica, **19**(3), 601-632. Chicago Datensatz unter http://mapconstruction.org

GIRSHICK, R., DONAHUE, J., DARRELL, T. & MALIK, J., 2013: Rich features hierarchies for accurate object detection and semantic segmentation. Extended version of the published conference paper at CVPR 2014. arXiv: 1311.2524v5.

GOODFELLOW, I., BENGIO, Y. & COURVILLE, A., 2016: Deep Learning. MIT Press 2016, Kapitel 8. http://www.deeplearningbook.org

KRIZHEVSKY, A., SUTSKEVER, I. & HINTON, G., 2012: ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information (NIPS), **2012**(1), 1097-1105.

NAIR, V. & HINTON, G. E. 2010: Rectified Linear Units Improve Restricted Boltzmann Machines. In: ICML'10 Proceedings of the 27th International Conference on Machine Learning, 807-814.

SIMO-SERRA, E., IIZUKA, S., SASAKI, K. & ISHIKAWA, H., 2016: Learning to simplify: fully convolutional networks for rough sketch cleanup. In: ACM Trans. Graph, **35**(4), Artikel 121.

SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I. & SALAKHUTDINOV, R., 2014: Dropout: A Simple Way to Prevent Neural Networks from Overfitting. In: Journal of Machine Learning Research, **15**, 1929-1958.

ZEILER, M. D., 2012: ADADELTA: An Adaptive Learning Rate Method. arXiv: 1212.5701v1.

Simulation von Menschenmengen im urbanen Umfeld

AXEL MEINERT[1,2], STEFAN HAHMANN[1], STEPHAN KOPF[2] & PATRICK BRAUSEWETTER[1]

Zusammenfassung: Die Simulation von Fußgängern und Menschenmengen besitzt eine große Relevanz in der Planung von Veranstaltungen, der Simulation von Evakuierungen und der realitätsnahen Modellierung von Fußgängerströmen. Dieser Beitrag gibt zunächst einen Überblick zu für diese Aufgabenstellung geeigneten physikalisch-mathematischen Modellen, wobei das agentenbasierte Social-Force-Modell vertieft wird. Des Weiteren wird diskutiert, inwiefern OpenStreetMap-Daten in entsprechende Simulationen integriert werden können. Basierend auf diesen Erkenntnissen werden Anforderungen an ein Programm abgleitet, welches Fußgänger und Menschenmengen im urbanen Umfeld visualisiert. Eine prototypische Implementierung wird vorgestellt und deren Performanz in Bezug auf die maximale Anzahl der simulierbaren Menschen getestet.

1 Motivation

Die Simulation von Fußgängern und Menschenmengen besitzt eine große Relevanz in der Planung von Veranstaltungen und baulichen Anlagen, in der Entwicklung von Evakuierungsszenarien und in der Analyse menschlichen Verhaltens. Lange Zeit lag der Fokus der Simulationsmodellierung in der Entwicklung von Verkehrssimulationen. Eine in vielen Erdteilen wachsende Bevölkerung, die stärkere Urbanisierung und größer werdende Umweltprobleme führen dazu, dass alternative Fortbewegungsarten an Bedeutung gewinnen, weshalb auch die Analyse des Fußgängerverkehrs bzw. von Fußgängerströmen in den Vordergrund rückt.

Gerade für die Planung von Großveranstaltungen nimmt die Relevanz von Fußgängersimulationen zu. Nicht nur die wachsende Bevölkerung, sondern auch eine zunehmende Eventisierung und Kommerzialisierung innerhalb der Gesellschaft prägen diese Entwicklung. Damit einhergehende Unglücksfälle wie z.B. bei der Love Parade in Duisburg 2010 (HELBING & MUKERJI 2012) oder mehrere Vorfälle während des Hadsch in Mekka (JOHANSSON 2009) erhöhen das Gefahrenbewusstsein bei Veranstaltern und Sicherheitsbehörden, die dadurch intensiver Evakuierungsszenarien entwickeln, auch mit Unterstützung entsprechender Fußgängersimulations-Software. HELBING et al. (2002) nennen eine unvollständige Liste von Unglücksfällen, die große mediale Aufmerksamkeit erlangten. Ein Ziel dieses Beitrags ist daher die Unterstützung der Arbeit von Sicherheitsbehörden durch verbesserte Fußgängersimulationen.

Ein weiteres konkretes Ziel ist es, die Menschenmengen in Demonstrationszügen realistisch zu simulieren. Damit soll deren Darstellung in einer Simulationssoftware für das Training von Stabsarbeit verbessert werden. Wir leisten zudem einen Beitrag durch die Analyse des OpenStreetMap-Datensatzes hinsichtlich der für die Fußgängersimulation sinnvoll nutzbaren Objektarten.

[1] Fraunhofer IVI, Fraunhofer-Institut für Verkehrs- und Infrastruktursysteme, Dresden, Zeunerstraße 38, D-01069 Dresden, E-Mail: [axel.meinert, stefan.hahmann, patrick.brausewetter]@ivi.fraunhofer.de
[2] Hochschule für Technik und Wirtschaft Dresden, Friedrich-List-Platz 1, D-01069 Dresden, E-Mail: stephan.kopf@htw-dresden.de

2 Literaturüberblick

Um das Programm zur Fußgängersimulation aufzubauen haben wir die möglichen Modellierungsansätze und bereits bestehenden Fußgängersimulationen analysiert und entsprechende Erkenntnisse in den Kapiteln zum Forschungsstand zusammengefasst. Um in späteren Schritten eine Vereinfachung der Menschenmenge zu ermöglichen beschäftigt sich ein Kapitel mit Verfahren zur Ableitung von Umrissen. Wie bereits in der Einleitung dargestellt ist die Einbindung von OSM-Daten die zweite elementare Zielsetzung dieser Arbeit. Dazu entsprechend beschäftigt sich Kapitel 4 mit der Motivation zur Nutzung von OpenStreetMap und den bisher gewonnenen Erkenntnissen. Konkrete Anforderungen an das Programm und erste Ergebnisse werden in den nachfolgenden Kapiteln vorgestellt. Abschließend folgt ein zwischenzeitliches Resümee, mit Ausblick auf die weiteren Entwicklungsmöglichkeiten.

2.1 Forschungsstand – Modelltypen

Die Modellierung von Fußgängersimulationen kann mit unterschiedlichen Modellierungsansätzen umgesetzt werden. Gängige wissenschaftliche Unterteilungen differenzieren nach dem Aggregationsgrad der Modellierung (KRETZ 2007), dem Raumbezug, dem Verhalten der Personen, stochastischen oder deterministischen Verhalten, dem Zweck und der Verfügbarkeit. Der Aggregationsgrad gliedert sich in mikroskopische, mesoskopische und makroskopische Simulationen. Mikroskopische Modellierungsansätze beschreiben jeden Fußgänger individuell. Der mesoskopische Ansatz stellt die Fußgänger ebenso individuell dar, jedoch wird die Bewegung der Fußgänger einheitlich berechnet und entspricht somit keiner Individualisierung mehr. Sowohl die graphische Darstellung als auch die rechnerische Modellierung erfolgt beim makroskopischen Aggregationsgrad für alle Fußgänger als Kollektiv. Der Raumbezug unterscheidet zwischen Gitter- und Netzwerkstrukturen bzw. einem kontinuierlichen Raum. Das Verhalten der Personen kann auf künstlicher Intelligenz oder Regeln basieren, des Weiteren funktional oder implizit sein. Nach allgemeinen Fußgänger-Modellen oder spezifischen, wie z.B. einer Evakuierungssimulation, unterscheidet der Zweck der Modellierung. Die Verfügbarkeit differenziert nach Open Source, kommerziellen oder noch nicht implementierten Produkten.
Drei verschiedene Modelltypen werden im Folgenden untersucht: die fluiddynamische Modellierung, die Modellierung basierend auf Zellulären Automaten und die Kräfte-basierte Modellierung.
Die fluiddynamische bzw. gaskinetische Modellierung (KORMONOVÁ 2013) nimmt an, dass das Bewegungsverhalten, welches sich bei Flüssigkeiten oder Gasen beobachten lässt, denen von Menschen ähnelt. Darauf aufbauend werden physikalische Gleichungen, welche die Bewegungen der Partikel beschreiben, auf die Bewegung der Fußgänger adaptiert. Im Falle der fluiddynamischen Herangehensweise handelt es sich um Anpassungen basierend auf den Navier-Stokes-Gleichungen. Dahingegen nutzt die gaskinetische Modellierung den Boltzmann-Ansatz. Entsprechende Modelle gibt es bereits seit der Mitte des 20. Jahrhunderts. Die ersten populären Simulationen kamen von TOGAWA (1955), welcher die Evakuierung bei Bränden untersucht hat, und PREDTECHENSKII et al. (1978), der die Bewegung von Menschenmengen in Gebäuden modelliert hat. Neuere Modellierungsansätze sind u.a. durch HELBING (1990) und HUGHES (2000) beschrieben. Die Ansätze der fluiddynamischen Modellierung besitzen die Eigenschaften der makrosko-

pischen Aggregation. Im Fall einer Fußgängersimulation bedeutet das, dass Visualisierung und Berechnung immer die komplette Menschenmenge umfassen. Ein Nachteil dieser Modellierungsart ist, dass das menschliche Verhalten mit physikalisch exakten Gesetzen nur grob angenähert werden kann (SCHADSCHNEIDER et al. 2009).
Zelluläre Automaten (ZA) sind eine Modellierungsform, welche weit über das Gebiet der Verkehrssimulationen hinaus Verwendung findet. Erste ZA wurden ebenfalls bereits in der Mitte des 20. Jahrhundert entwickelt. Diese beschreiben im Allgemeinen einen Raum welcher durch ein Raster gebildet wird, dessen Zellen verschiedene Zustände annehmen können. Auf ZA basierende Simulationen arbeiten in der Regel mit diskreten Zeitschritten (TOFFOLI 1991). Jeder ZA muss vier Grundbedingungen erfüllen (BURSTEDDE et al. 2001; SCHADSCHNEIDER et al. 2009):

- Die geometrische Anordnung des ZA ist ein Raster. Die meisten ZA nutzen ein zweidimensionales Raster.
- Das Raster muss eine explizite geometrische Form besitzen, z.B. rechteckig oder hexagonal. Die geometrische Form muss regelmäßig sein, sodass Nachbarschaften entstehen, beispielsweise eine Vierer-, Sechser- oder Achter-Nachbarschaft.
- Die Anzahl der Zellzustände, welche das Raster bilden, muss festgelegt sein. Ein einfaches Beispiel ist die Belegung mit zwei Zuständen. Dabei hieße der erste Zustand, dass die Zelle besetzt ist und der zweite, dass sie frei ist.
- Der nächste Zustand einer Zelle wird ausschließlich aus dem Einfluss der direkten Nachbarzellen bestimmt, woraus folgt, dass der Zustand einer Zelle, welche sich nicht in der direkten Nachbarschaft einer anderen Zelle befindet, keinen Einfluss auf diese Zelle nehmen kann.

Die Fußgängermodellierung mit Hilfe von ZA ist dem mikroskopischen Aggregierungsgrad zuzuordnen. Im Unterschied zum fluiddynamischen Modell ist die Berechnung räumlich und zeitlich diskret, da sich die Fußgänger nur von Rasterzelle zu Rasterzelle bewegen können.
Kräfte-basierte Modelle beruhen auf der Berechnung von Differentialgleichungen und sind damit der kontinuierlichen Berechnungsmethodik zuzuschreiben. Vom Aggregationsgrad sind sie, wie die ZA, den mikroskopischen Modellen zuzuordnen. In der Modellierungsgruppe der Kräftebasierten Modelle gibt es verschiedene Modellierungsansätze. Ein weit verbreiteter Vertreter ist das Social-Force-Model (SFM) nach HELBING et al. (1995). Die Grundannahme des SFM ist, dass die Partikel, welche im Sinne des Modells Agenten genannt werden, sogenannten sozialen Kräften ausgesetzt sind. Diese Kräfte beeinflussen die Stärke und Richtung der Bewegung des Agenten. Die Größen dieser Kräfte umschreibt die folgende Formel, welche auf HELBING et al. (2005) beruht:

$$\vec{f}_{\propto}(t) = \vec{f}_{\propto}^{0}(\vec{v}_{\propto}) + \vec{f}_{\propto B}(\vec{r}_{\propto}) + \sum_{\beta(\neq\propto)} \vec{f}_{\propto\beta}\left(\vec{r}_{\propto}, \vec{v}_{\propto}, \vec{r}_{\beta}, \vec{v}_{\beta}\right) + \sum_{i} \vec{f}_{\propto i}(\vec{r}_{\propto}, \vec{r}_{i}, t) \quad (1)$$

Die Formel besteht aus vier Kraftkomponenten, welche auf die Agenten einwirken:

- Der erste Term beschreibt die Beschleunigungskraft welche auf den Agenten wirkt. Die Kraft resultiert daraus, dass sich der Agent mit einer gewünschten Geschwindigkeit in Richtung seines Zieles fortbewegen möchte.
- Der zweite Term beschreibt die Abstoßungskräfte zwischen Fußgängern und Hindernissen.

- Der dritte Term beschreibt abstoßende Kräfte zwischen den Agenten untereinander.
- Der vierte Term beschreibt anziehende Kräfte zwischen den Agenten untereinander (z.B. bei Gruppen) und von bestimmten Objekten.

Die Parameter dieses Modells sind bereits durch empirische Untersuchungen kalibriert worden, was dazu führt, dass fußgängerspezifische Prozesse heute durchaus realitätsnah beschrieben werden können. JOHANSSON et al. (2007) haben die Laufwege von Fußgängern mittels Video-Tracking in realen Umgebungen untersucht. Unter anderem basiert ihre Arbeit auf Videoaufnahmen des Haddsch. Weiterhin haben HOOGENDORN & DAAMEN (2007) und MOUSSAÏD et al. (2009) verschiedene Versuchsszenarien in Hallen nachgestellt, aufgenommen und im Fall von Moussaïd mit aufgenommenen Videos von belebten Straßen verglichen.
Aufgrund der vorliegenden durch empirische Untersuchungen kalibrierten Modelle haben wir uns als Modellgrundlage für das Social-Force-Modell nach Helbing entschieden. Das Modell hat gleich mehrere Vorteile, welche unseren Zielstellungen entsprechen. Dazu gehören insbesondere der mikroskopische Aggregationsgrad und die vergleichsweise leichte Implementierbarkeit.

2.2 Forschungsstand – Ableitung von Umrissen

Bei der Verarbeitung und Analyse von großen Datenmengen ist es vorteilhaft, diese zu strukturieren und zu vereinfachen. Gerade hinsichtlich der Visualisierung einer Fußgängersimulation kann es von praktischem Nutzen sein, die Darstellung so zu abstrahieren, dass anstelle einer Vielzahl von einzelnen Agenten nur der abgeleitete Umriss der entstehenden gesamten Menschenmenge betrachtet werden kann.
Im Bereich der Computergrafik gibt es zwei Methoden, die für die Ableitung von Umrissen von Punktwolken im zweidimensionalen Raum verwendet werden können: die konvexe Hülle und die nicht-konvexe Hülle. Eine konvexe Hülle definiert sich als kleinste konvexe Menge, in der die zu untersuchenden Punkte enthalten sind. Dabei bedeutet Konvexität, dass alle Verbindungsstrecken zweier beliebiger Punkte ebenfalls innerhalb der konvexen Hülle liegen (PREPARATA & SHAMOS 1985; LANG 2016).

Abb. 1: Screenshot einer Menge von Fußgängern mit einer Umrissableitung aufbauend auf einer nahezu konvexen Hülle (links) und einer nicht-konvexen Hülle (rechts).

Im Gegensatz dazu, kann die nicht-konvexe Hülle jede mögliche Form annehmen, wovon die konvexe Hülle ein Sonderfall ist. Aufgrund einer fehlenden strikten Definition ist die Ableitung nicht-konvexer Umrisslinien komplexer, so dass für eine Punktewolke üblicherweise eine Vielzahl an nicht konvexen Hüllen existiert. Im Bereich der Computergrafik gibt es dafür verschie-

dene Ansätze, wie beispielsweise die Methode der Alpha-Shapes (EDELSBRUNNER et al. 1983) oder Chi-Shapes (DUCKHAM et al. 2008). Die Methode der Chi-Shapes nutzt einen Algorithmus, welcher auch eine konvexe Hülle entstehen lassen kann. Viel wahrscheinlicher ist es jedoch, dass eine nicht-konvexe Hülle entsteht, wobei deren Bedingung ist, dass sie von der entsprechenden konvexen Hülle eingeschlossen wird. Dabei lässt sich hinzufügen, dass bei einer identischen Punktmenge in Abhängigkeit der verwendeten Parameter verschiedene Hüllen entstehen können. Der Parameter Chi legt fest, wie stark das entstehende Shape in Richtung Konvexität, oder ins Gegenteil, streben soll. Ein Beispiel für das Aussehen eines solchen Chi-Shapes im Vergleich zur konvexen Hülle ist in Abbildung 1 zu sehen. Es wird deutlich, dass die abgeleitete Umrisslinie besser der Form der Punktmenge entspricht, als die konvexe Hülle.

3 Softwareüberblick

Es existiert eine Vielzahl an Programmen für Fußgängersimulationen. So zählt KRETZ (2007) bereits 2007 um die 50 Programme bzw. Modellierungsansätze. Die nachfolgende Auswahl an Fußgängersimulationen stellt einen Ausschnitt daraus dar. Insbesondere werden frei verfügbare Programme und Bibliotheken betrachtet: die Pedestrian Crowd Simulation Library (PEDSim), Crowd-Sim und der Jülich Pedestrian Simulator (JuPedSim).

PEDSim ist eine in C++ programmierte Bibliothek (libpedsim). Die Bibliothek implementiert den mikroskopischen Modellierungsansatz des Social-Force-Modells nach Helbing. Ein weiterer Bestandteil ist eine graphische Benutzeroberfläche (Qt-Framework). Die Modifizierbarkeit und Wiederverwendung des Programms ist durch Verwendung der GNU General Public Licence (GPL) (graphische Benutzeroberfläche) und der GNU Lesser General Public License (LGPL) (eigentliche Bibliothek libpedsim) gegeben. PEDSim verfügt über eine ausführliche Dokumentation.

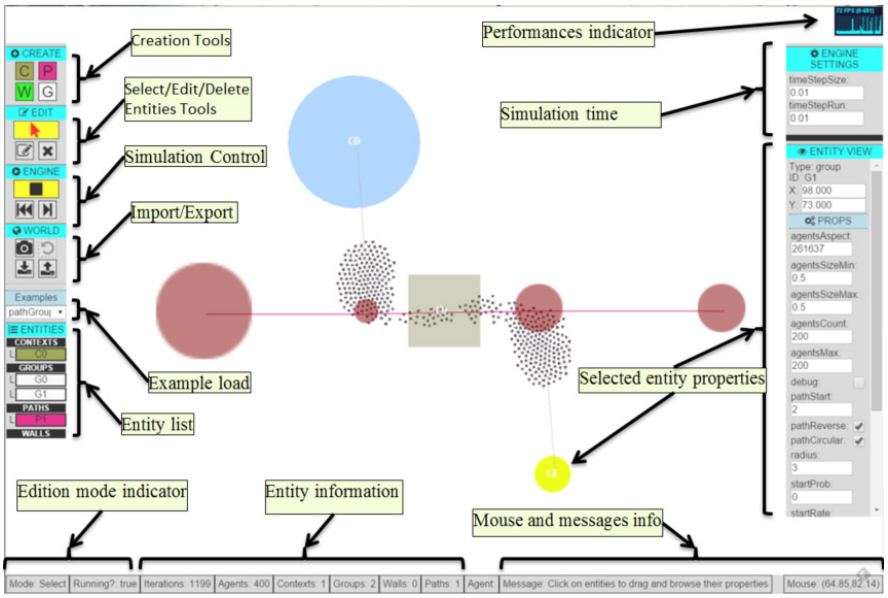


Abb. 2: Simulationsoberfläche der Fußgängersimulation Crowd-Sim, aufrufbar unter der Internetadresse https://rawgit.com/Jacq/crowd-sim/master/demo/index.html

Im Gegensatz zu PEDSim ist Crowd-Sim mehr als Simulationsumgebung denn als Bibliothek anzusehen. Der Modellierungsansatz ist wiederum derselbe, welcher auch in PEDSim genutzt wird, das SFM von Helbing. Crowd-Sim ist eine Browseranwendung und in JavaScript implementiert. Es steht unter der MIT-Lizenz. Crowd-Sim wird derzeit nicht weiterentwickelt. Die Darstellung der Parameter für die Fußgängersimulation ist intuitiv. Abbildung 2 zeigt die Anwendung mit den Möglichkeiten zum Setzen von Fußgängern, von Hindernissen, von Wegpfaden und der Simulationsparameter.
Sehr aktuell, und insbesondere für den wissenschaftlichen Betrieb gedacht, hat das Forschungszentrum Jülich mit JuPedSim eine Fußgängersimulation veröffentlicht (KEMLOH et al. 2016). Das Programm implementiert mehrere Modellierungsansätze. So stehen z.B. verschiedene Kräfte-basierte Modelle, wie beispielsweise das generalisierte Zentrifugalkraftmodell (CHRAIBI et al. 2010), und diverse Geschwindigkeits-basierte Modelle, wie u.a. das kollisionsfreie Geschwindigkeitsmodell (TORDEUX et al. 2015), zur Verfügung. JuPedSim ist in den Programmiersprachen C++ und Python geschrieben und wird unter der Lizenz LGPL bereitgestellt.

4 OpenStreetMap-Daten

Der OpenStreetMap (OSM) Datensatz kann als Datengrundlage für die Simulation von Menschenmengen in urbanen Umgebungen genutzt werden, da existierende Geoobjekte insbesondere in Großstädten sehr detailliert aufgenommenen sind und der Datensatz frei verfügbar ist. Die in den OSM-Daten vorhandenen semantischen Informationen in Form von Tags bzw. Attributen können genutzt werden, um für die Simulation nicht relevante Daten auszusortieren. Teilweise vorhandene Informationen z.B. zur Breite von Straßen und Wegen sind relevant für Fußgängersimulation im urbanen Kontext.
Für die Umsetzung der Fußgängersimulation wurden die vom Dienstleister Geofabrik aufbereiteten OSM-Daten genutzt. Diese gruppieren die OSM-Daten nach den Geometrietypen Punkt, Linie und Polygon. Dies entspricht einer Datenaufbereitung für die Verwendung in Geoinformationssystemen, welche bei den Primärdaten nicht vorhanden ist. Weiterhin sind in den Geofabrik-Daten einige Polygon-Objekte gleichzeitig als Punkt-Objekte abgebildet. Es ist zu beachten, dass dieser Datensatz nur eine bestimmte Auswahl der OSM Daten umfasst (RAMM 2017).
Aufbauend auf den Daten der Geofabrik konnte eine Tabelle entwickelt werden, welche als Vorlage für die Nutzung in einer Fußgängersimulation angesehen werden kann. Tabelle 1 zeigt eine Auswahl wichtiger Objektarten, sortiert nach dem Geometrietyp und den jeweiligen Layer-Namen.

Tab. 1: Auswahl an geeigneten OSM-Daten (basierend auf den Daten der Geofabrik) für die Integration als Hindernisse in eine Fußgängersimulation.

Geometrie	*Layer-Name*	*Feature*	*Beschreibung*
Punkt			
	gis.osm_natural		
		tree	Baum
	gis.osm_pois		
		post_box	Postkasten
		telephone	Telefonzelle
		recycling	Müllcontainer
		bench	Bank
		drinking_water	Trinkwasser
	gis.osm_traffic		
		traffic_signals	Ampeln
		stop	Stoppzeichen
		street_lamp	Straßenlaterne
		barrier_fence	starrer Zaun
Polygon			
	gis.osm_buildings		
		building	Gebäude
	gis.osm_landuse		
		scrub	Gebüsch
	gis.osm_water		
		water	Wasser (z.B. Teich)

Die Analyse der OSM-Daten zeigt, welche Probleme bei der Implementierung nach Tabelle 1 beachtet werden müssen:

- Einige Objekte im OSM-Datensatz sind nicht mit dem für eine Fußgängersimulation geeigneten Geometrietyp aufgenommen. Ein Beispiel hierfür sind Schranken, welche häufig als Punkt-Feature abgespeichert sind, für die Fußgängersimulation jedoch Linien sein müssten.
- Als Punkte gespeicherte Objekte, die in der Simulation als Hindernis betrachtet werden sollen, benötigen eine Art Standardgröße.
- Im OSM-Datensatz gibt es teilweise Redundanzen: so werden Objekte teilweise als Polygon und gleichzeitig als Punkt aufgenommen, etwa Point of Interests. Für eine bestmögliche Fußgängersimulation gilt es redundante Daten zu vermeiden.

5 Implementierung

5.1 Anforderungen

An die Entwicklung eines Programmes zur Fußgängersimulation kann eine Reihe an funktionalen Anforderungen gestellt werden. Die nachfolgende Liste zeigt einen Ausschnitt aus dem Anforderungskatalog, in welchem nur die elementarsten Anforderungen enthalten sind:

- Datenimport
    - Das Programm muss dem Nutzer erlauben Daten zu Fußgängern zu importieren.
    - Das Programm muss dem Nutzer erlauben Daten zu Hindernissen zu importieren.
    - Das Programm muss dem Nutzer erlauben Daten zu Wegpunkten zu importieren.
    - Das Programm muss dem Nutzer die Möglichkeit geben verschiedene Dateiformate zu importieren, z.B. OGC WKT und ESRI Shape.
- Information
    - Das Programm muss es dem Nutzer ermöglichen über verschiedene Informationen den Zustand der Simulation überblicken zu können.
- Simulationsparameter
    - Das Programm muss dem Nutzer ermöglichen verschiedene Simulationsparameter zu manipulieren.
- Darstellung
    - Im Programm muss an zentraler Stelle eine Karte untergebracht werden, in der die Simulation dargestellt wird.
    - Das Programm muss dem Nutzer die Möglichkeit geben die Sichtbarkeit von Objekten in der Karte zu steuern.

5.2 Umsetzung

Aufbauend auf dem Anforderungskatalog und den Forschungsständen der vorgestellten wissenschaftlichen Bereiche konnte ein Programm entwickelt werden, welches unsere grundlegenden Zielsetzungen erfüllt. Abbildung 3 zeigt einen Ausschnitt, bei welchem eine 1.000 Personen umfassende Menschenmenge durch die Prager Straße in Dresden läuft. Um die Hindernisse zu modellieren, wurden die OSM-Daten der Geofabrik genutzt.
Insgesamt gliedert sich das Programm in vier Bereiche. Der obere Bereich besteht aus den Schaltflächen für die Steuerung der Simulation (Start, Stopp, Pause) und den Schaltflächen für das Laden der Daten (Fußgänger, Hindernisse, Wegpunkte). An dieser Stelle konnten alle Anforderungen umgesetzt werden, insbesondere das Laden der Geometrien zu Hindernissen, Wegpunkten und Fußgängern. Diese Daten können sowohl als Shape-Datei sowie als CSV-Datei importiert werden. In letzterem Fall wird OGC Well Known Text (WKT) als Geometrieformat angenommen. Der rechte Bereich des Programms bietet eine Reihe von Möglichkeiten zum Ein- und Ausblenden von Elementen der Simulation. Im unteren, hier nicht dargestellten Abschnitt, können zusätzlich einige Simulationsparameter manipuliert werden, z.B. die Simulationsge-

schwindigkeit und die Durchschnittsgeschwindigkeiten der Fußgänger. Informationen bezüglich des Programms, insbesondere der Simulation, werden im untersten Teil der Nutzeroberfläche angezeigt, z.B. die Gesamtanzahl der simulierten Fußgänger und die durchschnittliche Dichte der Menschenmenge in Personen/m². Die Hauptfunktionalität des Programmes bildet die Kartenansicht bzw. Simulationsansicht im mittleren linken Bereich. Hier wird die eigentliche Simulation visualisiert.

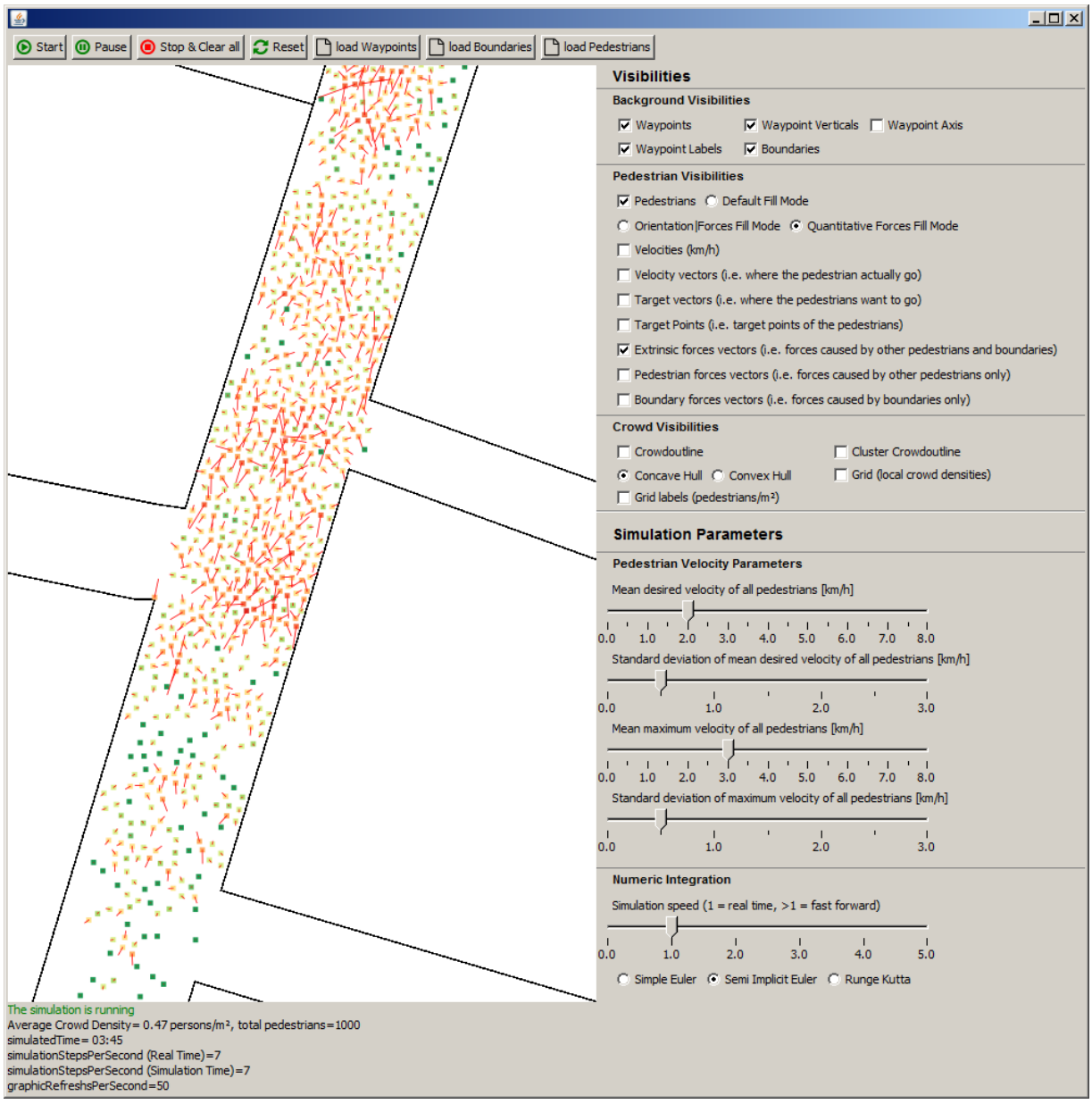


Abb.3: Ausschnitt aus der entwickelten Fußgängersimulation. Die Fußgänger laufen durch einen Teil der Prager Straße in Dresden, welche durch OSM-Daten visualisiert wurde. Die Farbdifferenzierung der Fußgänger quantifiziert die Stärke der Kräfte, die zu diesem Zeitpunkt auf sie wirken. Die roten Linien zeigen zusätzlich die Richtung der wirkenden Kräfte.

6 Fazit & Ausblick

Die entwickelte Fußgängersimulation wurde einem Performancetest unterzogen. Die Basis bildeten Shape-Dateien mit jeweils 300, 500, 600, 750, 1.000, 1.500 und 2.000 Fußgängern. Nach einer spezifischen Zeit, in der sich der Fußgängerstrom nach dem Start eingependelt hat, wurde die Performance, basierend auf den berechneten Simulationsschritten pro Sekunde, gemessen. Aus diesen Daten konnte eine Regressionskurve, mit folgender Formel, bestimmt werden:

$$y = \frac{12951}{x-4{,}35} \qquad (2)$$

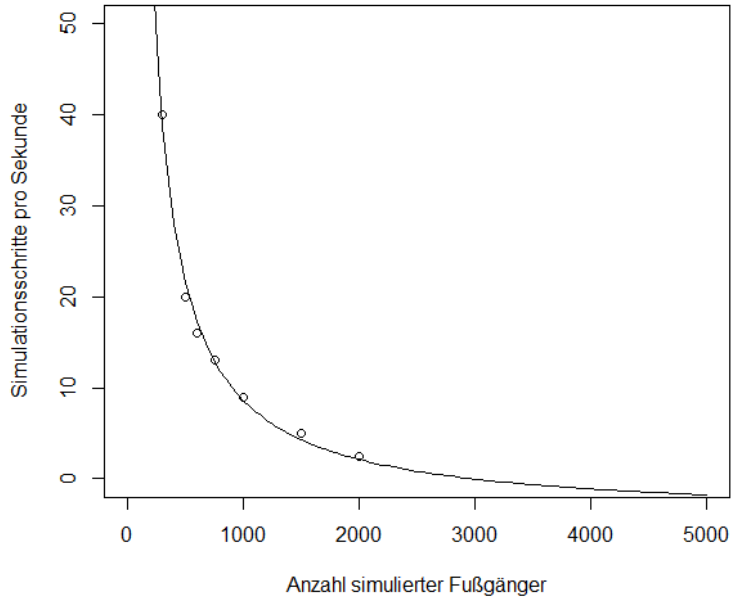


Abb. 4: Regressionskurve des Performancetests

Abbildung 4 zeigt die visualisierte Regressionskurve. Die x-Achse zeigt die Anzahl an Fußgängern und die y-Achse die Simulationsschritte pro Sekunde. Aufbauend auf den Ergebnissen aus Abbildung 4 und gleichzeitigen Tests der Simulation konnten wir feststellen, dass das Programm derzeit auf einem gebräuchlichen Arbeitsplatz PC mit bis zu 1.000 Fußgängern problemfrei einsetzbar ist. Der benutzte Testrechner läuft mit einem 64-bit Windows 7 System und einem 4-kernigen Intel-Prozessor mit 3,2 GHz Leistung. Weiterhin stand ein Arbeitsspeicher von 8 GB zur Verfügung. Bei einer Menge von 1.000 bis 2.000 Personen treten zunehmend Probleme auf, welche sich durch Ruckeln bemerkbar machen. Ab 2.000 Fußgängern läuft die Simulation nicht mehr flüssig genug, um verwendbare Ergebnisse zu produzieren. Außerdem werden die einzelnen Zeitschritte für die Simulation zu groß. Eine Verbesserung der Performance sehen wir als wichtige Verbesserung für zukünftige Arbeiten.

In Bezug zu OpenStreetMap sehen wir noch einen weiteren Untersuchungsbedarf. Dieser besteht einerseits darin Automatisierungsmöglichkeiten zu testen und andererseits die Primärdaten von OpenStreetMap analog zu den Geofabrik-Daten zu untersuchen. Konzeptuell besteht zudem das Ziel Anwendungslogik und Darstellung in der Bibliothek zu trennen, so dass eine Integration in andere Anwendungen leicht umsetzbar ist.

Weitere Forschungsarbeiten sind in das Projekt „TARGET“[3] (EU H2020) eingebettet. Ein Aspekt dieser Arbeiten wird sein, das Modell so anzupassen, einsatztaktisch relevante Vorgehensweisen zu modellieren. Beispielsweise, wenn die Polizei versucht eine Menschenmenge zu stoppen – die Menschenmenge könnte stehen bleiben, die Polizisten angreifen oder sich auflösen, um in alle Richtungen davonzustürmen.

7 Dank

Die Forschungsarbeiten, die zu den vorgestellten Ergebnissen geführt haben, wurden durch das Forschungsrahmenprogramm H2020 der Europäischen Union unter dem Kennzeichen 653350 (TARGET) gefördert.

8 Literaturverzeichnis

BURSTEDDE, C., KLAUCK, K., SCHADSCHNEIDER, A. & ZITTARTZ, J., 2001: Simulation of pedestrian dynamics using a two-dimensional cellular automaton. Physica A: Statistical Mechanics and its Applications, **295**(3-4), 507-525.

CHRAIBI, M., SEYFRIED, A. & SCHADSCHNEIDER A., 2010: Generalized centrifugal-force model for pedestrian dynamics. Physical Review, **E82**(4), 046111, 1-9.

DUCKHAM, M., KULIK, L., WORBOYS, M. & GALTON, A., 2008: Efficient generation of simple polygons for characterizing the shape of a set of points in the plane. Pattern Recognition, **41**(10), 3224-3236.

EDELSBRUNNER, H., KIRKPATRICK, D. & SEIDEL, R., 1983: On the shape of a set of points in the plane. IEEE Transactions on Information Theory, **29**(4), 551-559.

HELBING, D., 1990: Physikalische Modellierung des dynamischen Verhaltens von Fußgängern (Physical Modeling of the Dynamic Behavior of Pedestrians). PhD Thesis, Georg-August-Universität zu Göttingen.

HELBING, D. & MOLNÁR, P., 1995: Social force model for pedestrian dynamics. Physical Review, **E51**(5), 4282-4286.

HELBING, D., FARKAS, I., MOLNÁR, P. & VICSEK, T., 2002: Simulation of pedestrian crowds in normal and evacuation situations. Pedestrian Evacuation Dynamics, Schreckenberg, M. & Sharma, S.,D. (Hrsg.), Springer, Berlin, 21-58.

HELBING, D., BUZNA, L., JOHANSSON, A. & WERNER, T., 2005: Self-Organized Pedestrian Crowd Dynamics. Experiments, Simulations, and Design Solutions. Transportation Science, **39**(1), 1-24.

HELBING, D. & MUKERJI, P. 2012: Crowd disasters as systemic failures. Analysis of the Love Parade disaster. EPJ Data Science, **99**(1), 1-40.

HOOGENDOORN, S. & DAAMEN, W., 2007: Microscopic Calibration and Validation of Pedestrian Models: Cross-Comparison of Models Using Experimental Data. Traffic Granular Flow‘05, Schadschneider, A., Pöschel, T., Kühne, R., Schreckenberg, M. & Wolf, D. E. (Hrsg.), Springer, Berlin, 329-340.

[3] http://www.target-h2020.eu/

HUGHES, R. L., 2000: The flow of large crowds of pedestrians. Mathematics and Computers in Simulation, **53**(4-6), 367-370.

JOHANSSON, A., HELBING, D. & SHUKLA, P. K., 2007: Specification of the Social Force Pedestrian Model by evolutionary adjustment to video tracking data. Advances in Complex Systems, **10**(2), 271-288.

JOHANSSON, A., 2009: Data-driven modeling of pedestrian crowds. PhD Thesis, Technische Universität Dresden.

KEMLOH, U., CHRAIBI, M. & SESSER, F., 2016: Jupedsim/Jupedsim: Jupedsim V0.8.1, DOI: 10.5281/zenodo.592209.

KORMANOVÁ, A. 2013: A Review on Macroscopic Pedestrian Flow Modelling. Acta Informatica Pragensia, **2**(2), 39-50.

KRETZ, T., 2007: Pedestrian Traffic. Simulation and Experiments. PhD Thesis, Universität Duisburg-Essen.

LANG, H.W., 2016: Konvexe Hülle. Unter: http://www.iti.fh-flensburg.de/lang/algorithmen/geo/convex.htm (abgerufen am 05.01.2018).

MOUSSAÏD, M., HELBING, D., GARNIER, S., JOHANSSON, A., COMBE, M. & THERAULAZ, G., 2009: Experimental study of the behavioral mechanisms underlying self-organization in human crowds. Proceedings. Biological sciences, **276** (1668), 2755-2762.

PREDTECHENSKII, V.M. & MILINSKII, A.I., 1978: Planning for Foot Traffic Flow in Buildings. Amerind Publishing Company Inc., New Delhi.

PREPARATA, F. & SHAMOS, M., 1985: Convex Hulls: Basic Algorithms. Computational Geometry. Springer, New York, 95-149.

RAMM, F., 2017: OpenStreetMap Data in Layered GIS Format. Unter: https://www.geofabrik.de/data/geofabrik-osm-gis-standard-0.7.pdf (abgerufen am 16.01.2018).

TOFFOLI, T. & MARGOLUS, N., 1991: Programmable matter: concepts and realization. *Physica D,* **47**, 263-272.

TOGAWA, K., 1955: Study of Fire Escapes Based on the Observation of Multiple Currents. Report of the Building Research Institute, **14**, Tokyo, 1-29.

TORDEUX, A., CHRAIBI, M. & SEYFRIED, A., 2015: Collision-free speed model for pedestrian dynamics. Traffic and Granular Flow ’15, Knoop V. & Daamen, W. (Hrsg.), Springer, Berlin, 1-10.

Prediction of feed-in power from photovoltaics in local distribution networks

MOSTAFA ELFOULY[1], ANDREAS DONAUBAUER[1] & THOMAS H. KOLBE[1]

Abstract: Due to the shift of power production from highly centralised power plants to decentralised power production with renewable energy, the role of local distribution networks has changed. Originally, these networks were designed to distribute electric power from a centralised source to consumers. However, nowadays, the increasing number of climate-dependent decentralised energy production systems connected to local distribution networks, such as photovoltaics, affect the network stability. Thus, network operators need decision support methodology in order to cope with the network uncertainty. As a central part of the decision support system, climate-dependent forecast for feed-in power is required with high spatial resolution. In this paper, a method for the prediction of feed-in power production from photovoltaics, based on georeferenced power production and meteorological data is introduced.

1 Introduction

For decades, centralised power plants have been the operating model for electricity generation across Europe. This means, that huge quantities of electricity are being generated by power plants, that are physically clustered in a specific area or region located far away from the consumer (GREEN & SONNREICH 2014). The electricity being generated by the centralised power plants is being distributed through the electric power grid to multiple users (ACKERMANN et al. 2001).

Based on the Department of Economic and Scientific Policy within the EU parliament in June 2010, there has been a paradigm shift from centralised power plants to a more decentralised energy system. This shifts the narrative from passive to active consumers, in the sense that they can act as power producers as well (ALTMANN et al. 2010; COSSENT et al. 2009). End consumers, in turn, are often installing solar panels to cover their power needs and the excessive power are to be fed-in to the grid (CARLEY 2009). However, these climate-dependent decentralised energy production systems affect network stability. Hence, network operators need decision support tools for coping with network instabilities.

In this contribution, a method for feed-in power prediction from photovoltaics, based on georeferenced power production and climate data is described. The method takes into account that photovoltaic power production is influenced by various parameters: a) geographical location on earth and orientation of the panels, b) time of the year (changing sun elevation over the year), c) time of the day (HOSTE et al. 2009) (changing sun position over each day), d) climate (e.g. weather, wind, cloud, atmospheric conditions, etc.), e) shadowing effects caused by local topography (SALVATORE & FRANCISCO 2015), and f) panel-specific parameters (e.g. efficiency of the solar

[1] Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstr. 21, 80333 München, E-Mail: [mostafa.elfouly, andreas.donaubauer]@tum.de

panel, condition like dust, dirt or snow cover). Whereas, the aforementioned parameters (a-e) do play a pivotal role in the global horizontal irradiance (GHI), diffuse horizontal irradiance (DHI), and shortwave radiation received on the Earth surface, the panel-specific parameters are highly individual.
Within the course of this research, diverse meteorological aspects that play a role in the feed-in power have been investigated. As part of a state-of-the-art analysis, two approaches to compute or predict the feed-in power from photovoltaics have been identified: An analytical approach and an observation-based one.

2 Related Work

The amount of solar irradiation that reaches the surface of the solar panels determines how much power can be produced by a specific solar panel. Additionally, the panel-specific parameters (see section 1) contribute to the amount of power being produced and fed into the network. In this contribution, we have identified two approaches for tackling the issue of predicting the feed-in power from photovoltaics: an analytical approach and an observation-based approach.

2.1 Analytical Modelling

In the analytical approach, each and every aspect that plays a role in the computation of the amount of power being generated by the solar panel is to be thoroughly studied and a computational formula is to be obtained. As earlier mentioned, there are a variety of factors that do influence the feed-in power from photovoltaics.
One of the areas that have been studied in depth was the cloud coverage and its impact on the solar irradiation. In a research study by LUMB (1963) founded on a former study by KIMBALL (1928), an empirical correlation function between the average daily short-wave radiation and the fraction of sky covered by cloud has been introduced. However, that empirical formula that has been introduced did not take into account the different types of clouds. LUMB (1963) complements KIMBALL'S (1928) work by classifying and identifying different types of clouds. In his study, clouds were divided into nine categories, mainly based on their intensity (amount of cloud) and altitude. In addition to that, scattering of solar irradiation also varies based on, the cloud optical depth, cloud geometry, and the direction of the incident solar radiation (MCKEE & COX 1974; AIDA 1977).
However, the data required for these computations are typically not available, which makes it not practical in deducing a computational formula for estimating the short-wave radiation that is received at the solar panel. Also, due to the diversity of the domains that ought to be covered to precisely compute the amount of feed-in power from photovoltaics, an analytical model considering all the aforementioned factors would be very complex. Furthermore, from a practical point of view, the panel-specific parameters (see section 1) cannot be obtained for individual households. Hence, the second approach, which seems to be more applicable and more examined recently, is the observation-based approach.

2.2 Observation-based Modelling

In contrast to analytical modelling, observation-based approaches rely on correlating the observations of different phenomena with each other using statistical methods. In a study by YUAN et al. (2015), a correlation between the normalised solar radiation difference and the cloud optical thickness was built. An exponential correlation between the amount of solar radiation increase and the cloud optical thickness decrease was observed. In another research work by YANG et al. (2012), a forecast analysis between the Global Horizontal Irradiance (GHI) and Direct Normal Irradiance (DNI) based on a cloud cover index was developed. These studies rely on an observation-based approach to detect the correlation between cloud coverage and solar irradiation using NASA MODIS data (PLATNICK et al. 2015).

By taking into account these research studies, we concluded that an observation-based approach does overcome individual missing or prone to error parameters, as of the case in the analytical one. Additionally, various meteorological parameters can be correlated to the solar irradiation. This, in turn, allows for studying certain meteorological phenomena and their direct or indirect impact on the photovoltaic power generation (JEREZ et al. 2015).

As a result, it has been decided that we further investigate the observation-based approach in order to predict the power output from photovoltaic cells under certain weather conditions.

However, the studies mentioned above do not take into account local topography and panel-specific parameters (see section 1), which are important for predicting the feed-in power in local distribution networks, but which are not obtainable in most cases. In our research we propose an observation-based approach, which tries to overcome these problems.

3 Proposed Observation-based Approach

Our approach relies on statistical analysis of georeferenced historic power production measurements from photovoltaic systems installed in local distribution networks and georeferenced historic meteorological data. Figure 1 gives an overview of our approach. The numbers shown in the Figure correspond to the numbering of the sections in this paper.

3.1 Acquisition of power production data from photovoltaics stations

The first step in this research project was to obtain a reliable and continuous stream of photovoltaic power production data. Online services for providing crowd sourced photovoltaic power data, such as PVOutput[2] were inspected and analysed. More than 50 photovoltaic stations in the area of Utrecht (Netherlands) were picked from PVOutput and selected for further investigation (see Figure 2). Power production data acquisition was limited to a period of 12 months back, because of PVOutput limitations. However, although PVOutput provides free online service for photovoltaic output data, it does not guarantee their consistency or sustainability.

In order to take the observation-based approach a step further and correlate power production to meteo-rological factors, a con-tinuous stream of power readings was essential.

[2] http://www.pvoutput.org

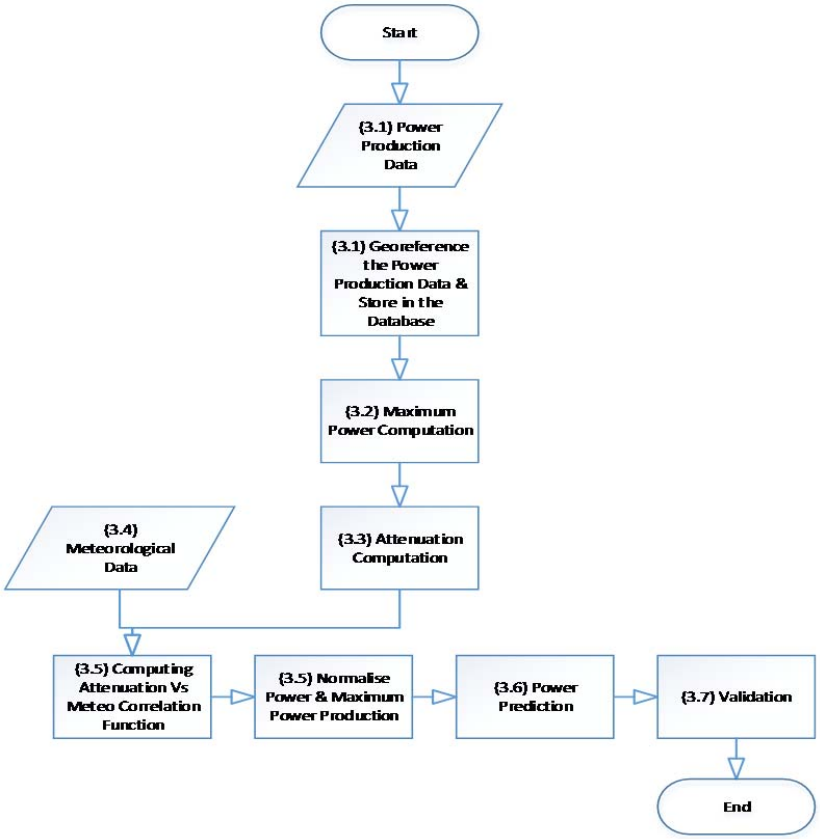


Fig. 1. Overview of the proposed approach

Therefore, the German energy company EnBW ODR provided us with photovoltaic power production data for 22 stations (spatial distribution see Figure 3) over the course of 13 months with 15 minutes temporal resolution. The same workflow has been applied on the EnBW ODR and the PVOutput data and the readings have been geo-referenced and stored in a geo-database in order to allow for spatial analysis, such as spatial autocorrelation of the power production data, spatial interpolation, and spatial correlation with meteorological data (see 3.5). In the following sections, the further steps of our approach are described. The outcomes are illustrated using either the crowd sourced PVOutput data or the data provided by the energy company EnBW ODR.

3.2 Maximum Power Estimation for each Station

In order to overcome the individual panel specific characteristics, we needed to derive a parameter, which accounts for all the influences and makes different photovoltaic systems comparable. Hence, we needed to identify the maximum expectable power that can be generated by a specific station at a particular point in time on a particular day. The maximum generated power varies on daily basis, because of the factors, such as change of inclination of the sun during the year (shadowing situation), and season specific atmospheric conditions. Therefore, we had to create the maximum power curve for each individual day of the year. In order to measure the maximum power at each point in time produced by a certain solar panel, we resorted to historic measurements. The maximum power produced for a specific station was computed for each day of the

year applying a moving window approach as follows: for each day of the year, for every 15 minutes time step, the generated power is compared to its relevant time values for +/-15 days, then the maximum reading is being selected (see Figure 4). The period of +/-15 days was chosen, because we assumed, that within this period the probability is high to find clear sky conditions. The maximum generated power was then assumed to be the maximum expected power that can be obtained from a particular solar panel under clear sky conditions at this point in time.

Fig. 2: PVOutput stations

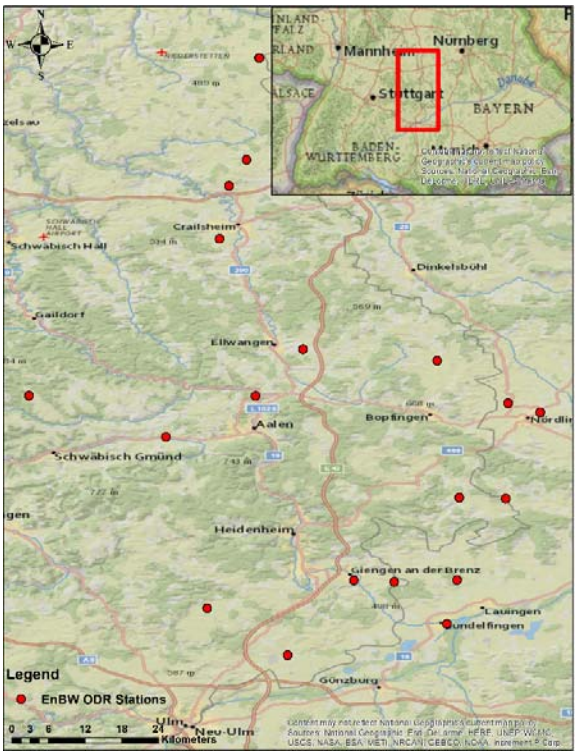


Fig. 3: EnBW ODR stations

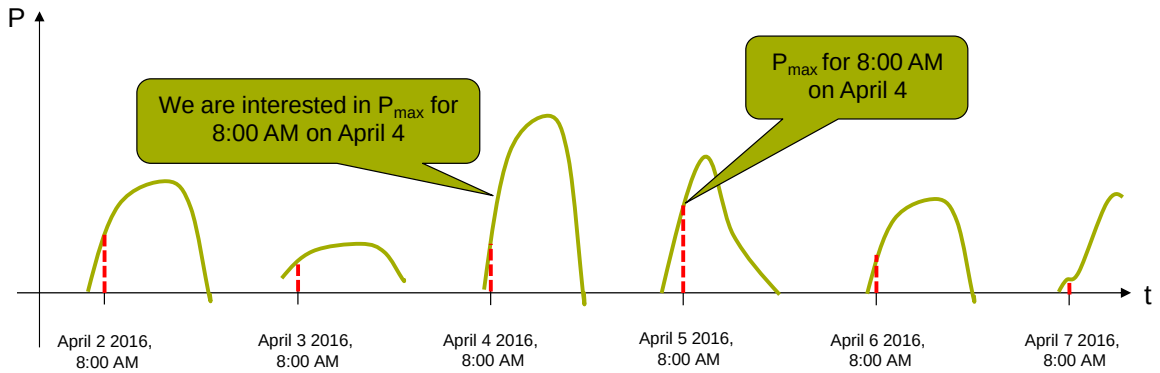


Fig. 4: Moving window approach for maximum power computation

3.3 Attenuation Computation and Spatial Interpolation

Considering the fact that various meteorological conditions may reduce the amount of power being generated by a certain percentage and by learning the maximum expected power production for each station at a specific point in time (see section above), we derived a specific attenuation function for each station along the year. Attenuation is defined as the percentage of power being lost, due to meteorological conditions. An attenuation value of 0% means that a specific station generates the maximum expectable energy for this point of day and time, taking into account the station-specific non-meteorological factors, such as date and time, efficiency, panel-condition and local topography. An attenuation value of 100% means that this station produces no power at all. In order to determine the current attenuation percentage, a ratio between the measured and maximum expectable power at a specific point in time has to be built. The attenuation accounts for all the influences and the afore-mentioned parameters – apart from meteorological factors - and therefore makes different stations comparable:

$$\text{Attenuation } A\ [\%] = \left(1 - \frac{P}{P_{max}}\right) * 100$$

where P is the measured power at a specific point in time and P_{max} is the maximum expectable power at this point in time under clear sky conditions (see 3.2).

Then, we generated daily-based, seasonal-based, and station-based, diagrams for plotting the actual power that is being produced by a specific station, maximum power at each point in time, and the attenuation percentage (see Figure 5).

Then, three scenarios have been considered, in order to observe various correlation aspects. The first one is to indicate the meteorological influence, such as cloud coverage and its impact on the power production. Thus, we decided on two consecutive dates where one is cloudy and the other one has clear sky conditions (see Figure 5). It can be noted, that there is severe power drop on the cloudy day, compared to the clear sky day. However, there might have been other meteorological factors that have played a role also. In the second scenario, we checked for spatial auto-correlation between nearby stations, in order to observe for certain meteorological phenomena, such as cloud movement, that occur over a certain region (see Figure 6). It can be noticed that the power curve is

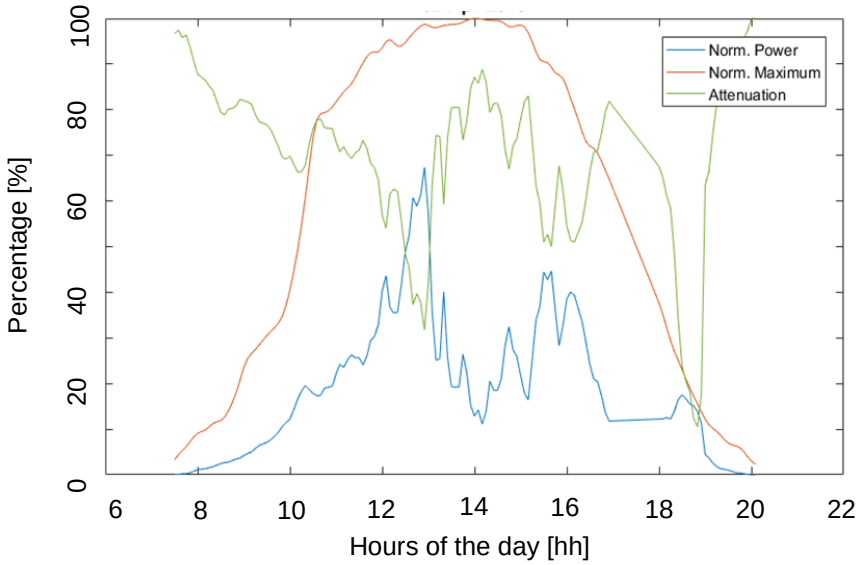
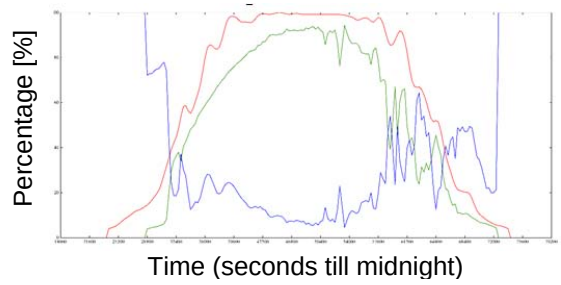


Fig. 5: An example for actual power production, maximum power, and attenuation percentage for Pasij station on April 2nd 2016. Example using PVOutput

almost the same for both nearby stations, because they do get exposed to similar conditions at this point in time. The last scenario was studying the seasonal effect and its correlation to the amount of power that can be produced by a certain station (see Figure 7). In that last scenario (see Figure 8), we can study seasonal effects of photovoltaic power production.



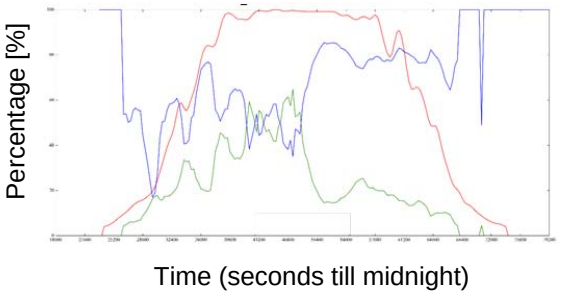
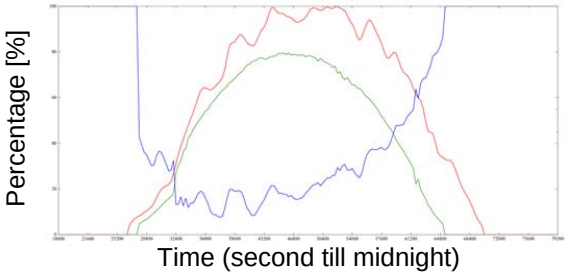


Fig. 6: Clear Sky vs Cloud for Moorstenweg station on 9th and 10th of May consecutively – examples using PVOutput data. The green curves differ substantially.



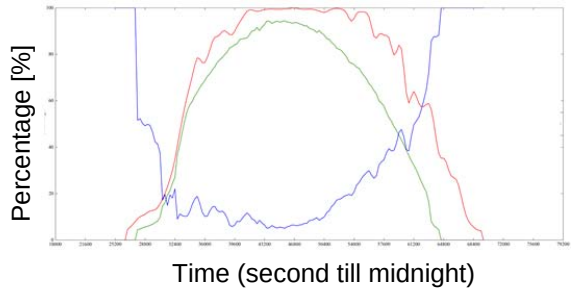
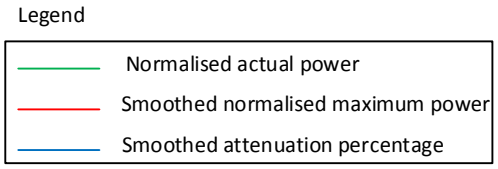
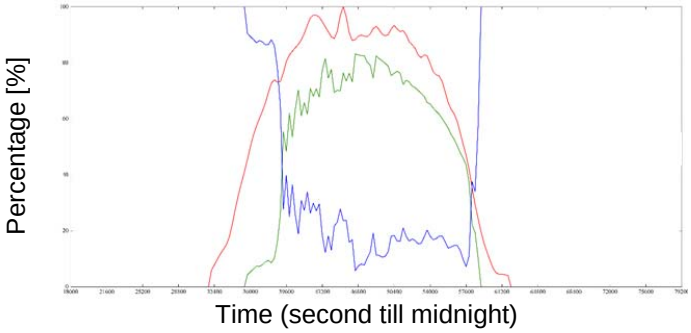


Fig. 7: Spatial auto-correlation on Fregatstraat & Moorsterweg on March 17th 2016 consecutively using PVOutput data. The green curves have nearly the same shape.





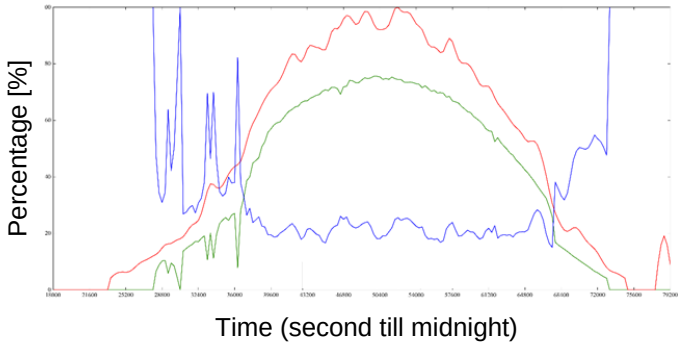


Fig. 8: Seasonal effect for Fregatstraat station on January 25th and July 18th 2016 consecutively using PVOutput data

Afterwards, a spatial interpolation between the stations has been performed, in order to deduce attenuation fields, which show spatial distribution and the change of attenuation throughout the day, month, and year for the whole area of study (see Figure 9). The spatial interpolation was carried out using the natural neighbours interpolation algorithm (SIBSON 1981), because of it creates a smooth field, whereas the observations at the known data points are preserved.

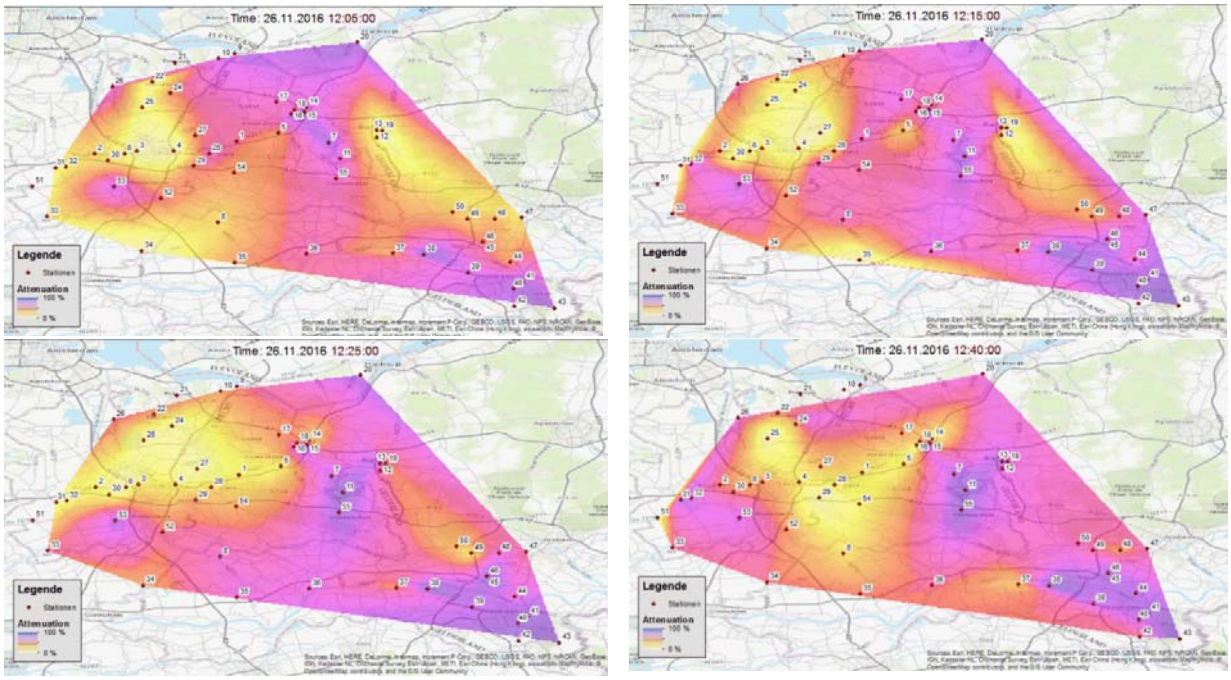


Fig. 9: Attenuation field of four consecutive points in time in the Utrecht area: Top left 12:05 pm, top right: 12:15 pm, bottom left: 12:25 pm, bottom right: 12:40 pm. Data source: PVOutput

The attenuation field now allows to predict the feed-in power for a station installed at an arbitrary location within the extent of the field and for a specific point in time. This use case is related to a real-world problem, that energy companies face. Whereas, only few stations have smart meters installed, the energy companies need to predict feed-in power for all stations connected to their network. Thus, the attenuation field can be used to predict the potentiality of power production for stations with no smart meters installed, by extracting the attenuation values for the geographic locations of the power stations from the field. The amount of power to be produced is then to be computed from the extracted attenuation values. In order to do so, the maximum power that is ought to be generated from this particular panel still has to be known.

3.4 Meteorological Data Acquisition

Acquiring meteorological data for the same spatial region as of the photovoltaics stations is crucial. Also, the time span and temporal reference system for the various meteorological attributes has to match with that of the power production readings. Hence, within the scope of this project, we obtained satellite-based meteorological data for our test region with a spatial resolution of

3x3km and a temporal resolution of 15 minutes from the company METEOBLUE[3]. The obtained data varied from cloud coverage and temperature to Global Horizontal Irradiance (GHI) and Diffuse Horizontal Irradiance (DIF). In order to prepare for correlation with power production data, the meteorological data had to be geo-referenced. Then, as a last step before correlating the meteorological readings with the power production readings, both of those readings had to be transformed to the same time base (UTC offset of one hour was chosen).

3.5 Finding a correlation function

By analysing the historic observations and meteorological data, a correlation function between the attenuation and the various meteorological factors derived from satellite observations, such as cloud cover and solar irradiance could be determined. The derived correlation function, allows to predict the attenuation field for a specific future point in time from weather forecast.
As a first step of the correlation process, geo-location and date-time of the power production, attenuation and meteorological data were used for joining the readings in a common database table. Then, scatter plots were produced to examine for potential correlation. It is to mention, that night time was excluded from our study, since it gives us a false indicator of strong correlation where readings from power production as well as solar irradiation are zero.

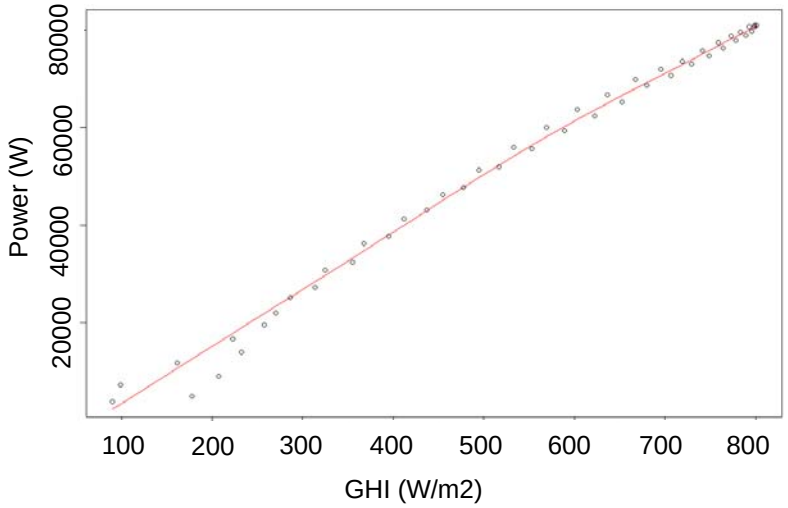


Fig. 10: Scatter plot between GHI & power for Spraitbach station in August 24th 2016. Data source: Power production: EnBW ODR; Meteorological data: Meteoblue

By analysing Figure 10, a mostly linear correlation between the GHI (W/m2) and the produced power (W) for a specific station can be observed. A linear regression function was then generated and checked thoroughly. In addition to that, we inspected the correlation between the attenuation and GHI. However, a linear regression in that case was not the best fit (see Figure 11). Thus, we generated a polynomial regression function of the 5th degree to correlate GHI with the attenuation percentage.

[3] https://content.meteoblue.com/

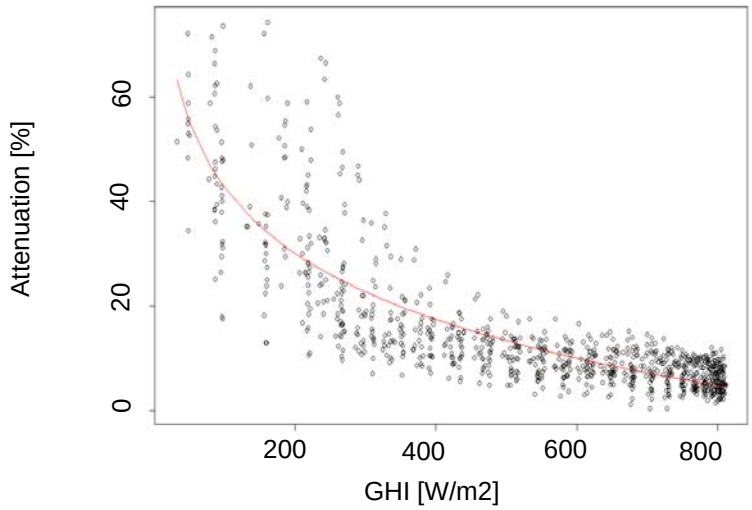


Fig. 11: Scatter plot between GHI & attenuation for all the stations on September 24th 2016. Data source: Power production data: EnBW ODR; Meteorological data: Meteoblue

By examining the correlation for longer periods of time, we concluded that there is a higher correlation between the GHI and the attenuation, than between the GHI and the power production, specifically under non-clear-sky-conditions. Hence, as a last step before forecasting, we extended our correlation over a 12-months timespan to assess the goodness of fit of the model. Furthermore, we studied the correlation separately for each season, in order to evaluate the fitting model and determine if the correlation varies from season to season.

As a result, of that, the goodness of the fitting model across the whole year is estimated to be 0.74 R squared. The fitting model operates during autumn with 0.85 R squared and the least during summer with 0.63 R squared. Additionally, it was found out that by taking into account 12 months of historical data, a linear regression equation could be deduced, which is in contrast to the results when we take only small sample data as seen in Figure 11.

We plotted the diverse meteo-rological aspects that we have obtained, such as cloud coverage and temperature against the attenuation to examine if there is a noticeable correlation between them (see Figure 11). However, temperature did not play a significant role in the amount of power being generated. As for cloud coverage, while it plays a role in the amount of solar radiation received on the earth surface, as we explained in the introduction, there was no direct correlation between the power production and the cloud coverage. This is because, in the meteorological data available, cloud coverage is expressed as a percentage representing the fraction of a 3×3 kilometres tile covered by clouds with 15 minutes temporal resolution. This means, that clouds can be formed within this time period without being detected. Also, different formations of clouds react to solar radiation differently as earlier discussed in section 2.2, and cannot be discriminated in the meteorological data available. Based on the regression model that has been generated in section 3.5 and by feeding GHI as input to the model, we could predict the attenuation percentage on that particular day. As demonstrated in Figure 13, the blue line in the middle is the estimated fitting line, while the red and the green ones are the lower and upper boundaries consecutively of the 95% confidence interval. Thus, the correlation function enables us to predict the attenuation field for a specific future point in time from weather predictions. Moreover, it allows us to predict the feed-in power for a specific station and for a specific future point in time from the predicted attenuation field, by reversing the attenuation formula:

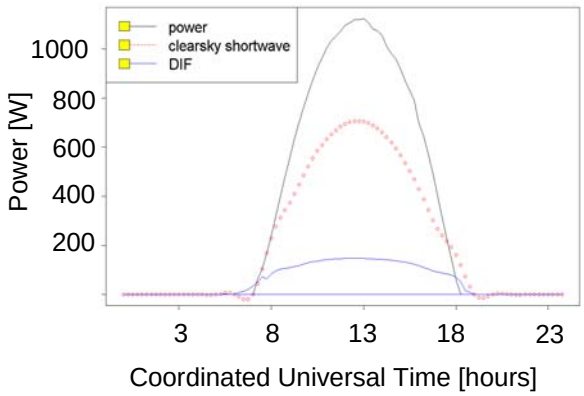


Fig. 12: Meteorological factors vs photovoltaics power production for a clear sky day for Altheim (Alb) on September 14th 2016. Data source: Power production data: EnBW ODR; Meteorological data: Meteoblue

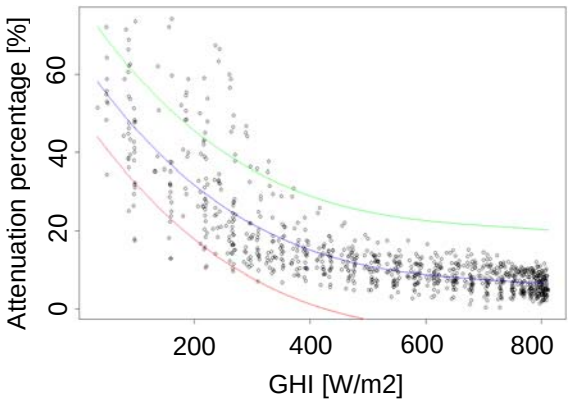


Fig. 13: Estimated attenuation percentage for Schrozberg on July 11th 2016. Data source: Power production: EnBW ODR; Meteorological data: Meteoblue

$$\text{Power } [W] = \left(1 - \frac{A[\%]}{100_{max}}\right) * Pmax[W]$$

Validation

In order to evaluate the estimation, we divided our data into training data and test data. By taking into account the seasonal influences, we used training data from different seasons, in order to be able to have a proper estimation. Validation in our approach went through different steps. In a first step the acquired data was validated. In order to determine if the power production and meteorological data are in the same time zone, we generated diagrams for various meteorological factors against power (see Figure 12). It can be clearly observed that the power production (W) and the clear sky shortwave (W/m2) do overlap and there is no time lag. Also, it can be clearly observed that the given day has clear sky conditions.

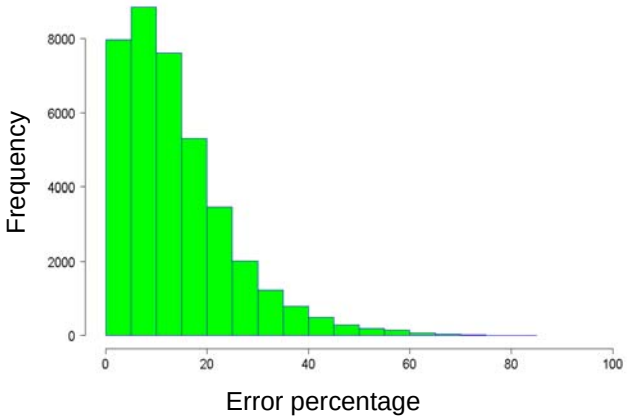


Fig. 14: Histogram showing the frequencies of the absolute differences between measured and predicted attenuation in Winter. Data source: Power production: EnBW ODR; Meteorological data: Meteoblue

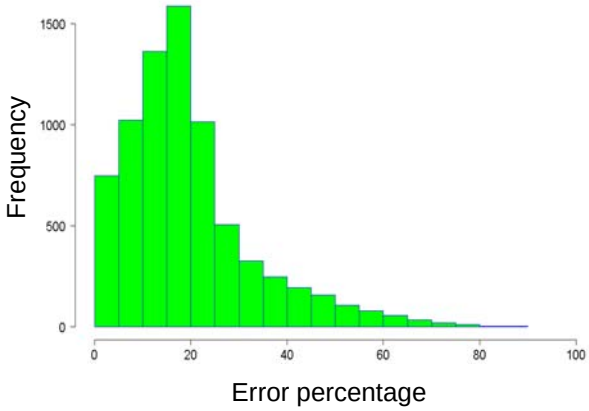


Fig. 15: Histogram showing the frequencies of absolute error percentages for power estimation in Spring. Data source: Power production: EnBW ODR; Meteorological data: Meteoblue

In order to validate the prediction model that has been created in section 3.5, we extracted a table containing the predicted attenuation, the actual measured attenuation, and the attenuation difference of each of the 812842 data items. In addition to that, we computed the estimated power and compared it to the actual power reading, in order to measure the power difference and generate the error percentage. We then generated histograms for the attenuation difference across different seasons to evaluate which seasons have higher correlation than others (see Figure 14). By observation, we can conclude that in winter, most of the attenuation differences lie between 0% and 20%. As a last step, the table below was created to provide statistical description of the validation.

Tab. 1: Statistical description of the validation results: differences between measured and predicted power (absolute percentage)

Season	Mean	Median	Standard deviation	Min	Max
Winter	31.3	27.9	21.3	0.0	93.3
Summer	19.9	14.5	16.9	0.0	91.7
Spring	23.0	18.6	17.8	8.0	91.0
Autumn	19.4	17	13.5	0.0	86.0
All seasons	25.2	20.1	19.6	8.0	93.3

4 Conclusions & Outlook

In this paper, the correlation between photovoltaic power production and heterogeneous meteorological and topographical aspects have been studied. Two approaches have been investigated, an analytical-based approach and an observation-based approach. In the first approach even though it gives a clear depiction of the amount of solar irradiation received on the Earth surface, the data sources required for these computations are generally not available. Additionally, it does not take into account the local topography and panel-specific parameters, such as efficiency and condition and its impact on the amount of power being produced by the solar panels. Furthermore, there is a plethora of domains that do play a role in the amount of power that is being produced, and predicting power values requires deep knowledge in all these various domains.
In the observation-based approach, we studied the actual power that is being produced by solar panels and its correlation to meteorological factors. This allowed us to take into account all the influences on the solar panel by analysing the actual power that is being produced. Hence, by computing the attenuation factor, as well as normalising the power, we could compare different PV systems and overcome station-specific characteristics that might be hard to obtain in an analytical approach. By using historical data for training our system, we could then deduce a correlation function between GHI and attenuation that counts for all these factors and their influence on the power production. We could apply this function for predicting power production from GHI forecast. In addition to that, the attenuation field that has been created enables us to do the prediction also for stations that do not have smart meters installed and also to locate solar panels, in order to generate the most power.

The accuracy of the prediction was determined by a cross-validation approach showing acceptable results. However, taking into account the use case of decision support for operators of local distribution networks, the spatio-temporal resolution of the weather data available today is still insufficient. This leads to a potential future research area, trying to perform now-casting of energy production from smart meter readings of georeferenced stations by applying methods like Gaussian Markov random fields. In this future approach, the attenuation fields could be extrapolated, which would have the strength to overcome the necessity of weather information in the prediction process.

The research presented in this paper was carried out in the project VerNet-LEM funded by the Bavarian Ministry of Economic Affairs in cooperation with the company AED-SICAD, based on a geospatial database system with data from EnBW ODR.

5 Bibliography

ACKERMANN, T., ANDERSSON, G. & ODER, L.S., 2001: Distributed generation: a definition. Electric Power Systems Research 57, 195-204.

ALTMANN, M., BRENNINKMEIJER, A., LANOIX, J.-CH., ELLISON, D., CRISAN, A., HUGYECZ, A., KORENEFF, G. & HÄNNINEN, S., 2010: Decentralised Energy Systems. Directorate General for Internal Policies Policy Department A: Economic and Scientific Policy, Industry, Research and Energy, European Parliament.

AIDA, M.A., 1977: Scattering of solar radiation as a function of cloud dimension and orientation. Journal Quantitative Spectroscopy and Radiative Transfer, **17**, 303-310.

BRUMMITT, N. & LUGHADHA, E. N., 2003: Biodiversity: Where's Hot and Where's Not. Conservation Biology, **17**(5), 1442-1448.

CARLEY, S., 2009. Distributed generation: an empirical analysis of primary motivators. Energy Policy, **37**, 1648-1659.

COSSENT, R., GOMEZ, T. & FRIAS, P., 2009: Towards a future with large penetration of distributed generation: Is the current regulation of electricity distribution ready? Regulatory recommendations under a European perspective. Energy Policy, **37**, 1145-1155.

YANG, D., JIRUTITIJAROEN, P. & WALSH, W.M., 2012: Hourly solar irradiance time series forecasting using cloud cover index, Solar Energy, **86**(12), 3531-3543, Dec. 2012.

ENVIRONMENTAL PROTECTION AGENCY, UNITED STATES 2017. Centralized Generation of Electricity and its Impacts on the Environment.

GREEN, D. & SONNREICH T., 2014: Centralised to De-centralised Energy – What Does it Mean for Australia. Infrastructure for 21st Century Australian Cities. Australian Davos Connection, limited, 2014, 177

HOSTE, G., DVORAK, M. & JACOBSON, M.Z., 2009: Matching hourly and peak demand by combining different renewable energy sources. A case study for California in 2020, Stanford University, Department of Civil and Environmental Engineering. http://www.etopia.be/IMG/pdf/HosteFinalDraft.pdf.

JEREZ, S., TOBIN, I., VAUTARD, R., MONTAVEZ, J.P., LOPEZ-ROMERO J.M., THAIS, F., BARTOK, B., CHRISTENSEN, O.B., COLETTE, A., DEQUE, M., NIKULIN, G., KOTLARSKI, S., VAN MEIJGAARD, E., TEICHMANN, C. & WILD, M., 2015: The impact of climate change on pho-

tovoltaic power generation in Europe. Nature Communications, **6**, http://dx.doi.org/10.1038/ncomms10014

KIMBALL, H., 1928: Monthly Weather Review, Washington, **89**, 147.

KOKHANOVSKY, A.A., 2003: The influence of cloud horizontal inhomogeneity on radiative characteristics of clouds: An asymptotic case study, IEEE Trans. Geoscience Remote Sensing, **41**, 817-825.

LUMB, F.E., 1963: The influence of cloud on hourly amounts of total solar radiation at the sea surface. Quarterly Journal of the Royal Meteorological Society, 90(383), 43-56, 10.1002/qj.49709038305

MCKEE, T.B. & COX, S.K., 1974: Scattering of Visible Radiation by Finite Clouds. Journal of Atmospheric Science, **31**, 1885-1892.

PLATNICK, S., ACKERMAN, S., KING, M.D., MEYER, K., MENZEL, W.P., HOLZ, R. E., BAUM, B.A. & YANG, P., 2015: MODIS atmosphere L2 cloud product (06_L2), NASA MODIS Adaptive Processing System, Goddard Space Flight Center.

SIBSON, R., 1981: Interpolating Multivariate Data: Chapter 2: A Brief Description of Natural Neighbor Interpolation, John Wiley & Sons, New York.

Task-oriented Data Classification of Choropleth Maps for Preserving Local Extreme Values

JUIWEN CHANG[1] & JOCHEN SCHIEWE[1]

Abstract: It is well known that the selection of class breaks for choropleth maps, as a powerful tool widely used to illustrate values among polygons by geographers and cartographers, have a significant influence on user perception. Furthermore, conventional approaches do not take spatial context into consideration. This contribution presents an innovative algorithm for the task of preserving local extreme values and compares it to other classification methods. With that, we cannot only prove the enhanced visual effect but also the quantitative improvement in terms of preserved local extreme values in choropleth maps.

1 Introduction

In thematic maps, e.g. for representing population densities, the attribute values of single regions (like districts) can be categorized into representative classes and respective colors. The visual impact of these choropleth maps – and with that derived interpretations and decisions – is influenced by the actual value distribution of the input data, the data classification method and the selection of color ramps.

Conventional data classification methods (like equidistant, quantiles, or Jenks) work data-driven. With that they explicitly do not consider any spatial context within the dataset. If, for example, a state that contains a local extreme value is put into the same category as some of its neighbor states, this important information gets lost. Moreover, neglecting geographical relationships among adjacent polygons may also influence other patterns like hot spots or clusters.

Instead of the conventional data-driven approach, this contribution will present a new task-oriented classification approach that has been developed and implemented for the task of preserving local extreme values.

2 Method

For the task of preserving polygons showing local extreme values, the following three-step procedure has been implemented and tested.

2.1 Identification of local extreme values

A local extreme value is defined as a polygon that shows a larger (or smaller) value compared to all its neighboring polygons. Neighbors are defined as polygon pairs having a common boundary or sharing a common vertex. An R-Tree spatial index was used to accelerate the search for neighboring polygons. Among all neighbors of an actual extreme value candidate, the minimum of the absolute difference values is identified and stored. In fig. 1, for example, this difference

[1] HafenCity University Hamburg, Lab for Geoinformatics and Geovisualization (g2lab), Überseeallee 16, D-20457 Hamburg, Germany, E-Mail: [juiwen.chang, jochen.schiewe]@hcu-hamburg.de

value is 10.17. The reason why we select the minimum absolute difference is because we want to find one break value that separates this local extreme polygon from all its neighbors. In the case of fig. 1, any value between 7.52 and 17.69 can be used for such a separation.

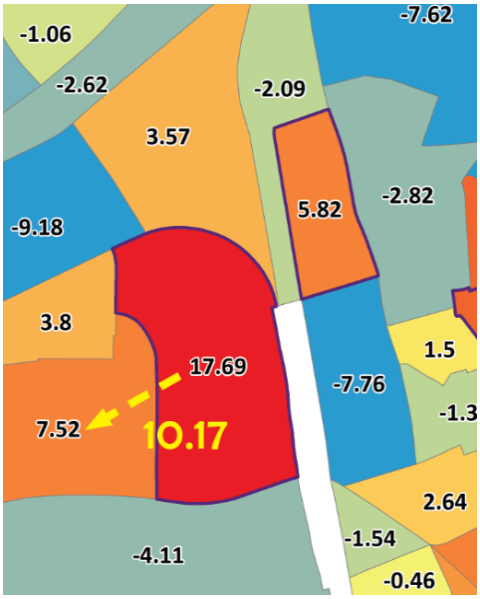


Fig. 1: Identified polygon with local extreme value (value 17.69) and the "closest" neighbor (value 7.52)

2.2 Sorting all local extreme value polygons according to significance

For all local extreme values the corresponding value intervals (bounded by the values of the extreme polygon and the numerically closest neighboring polygon as derived in step 2.1) are plotted and sorted from largest to smallest interval widths (Fig. 2). The larger the interval width (e.g., value 10.17 in Fig. 1) is, the more important or "significant" this local extreme value is in comparison to other candidates.

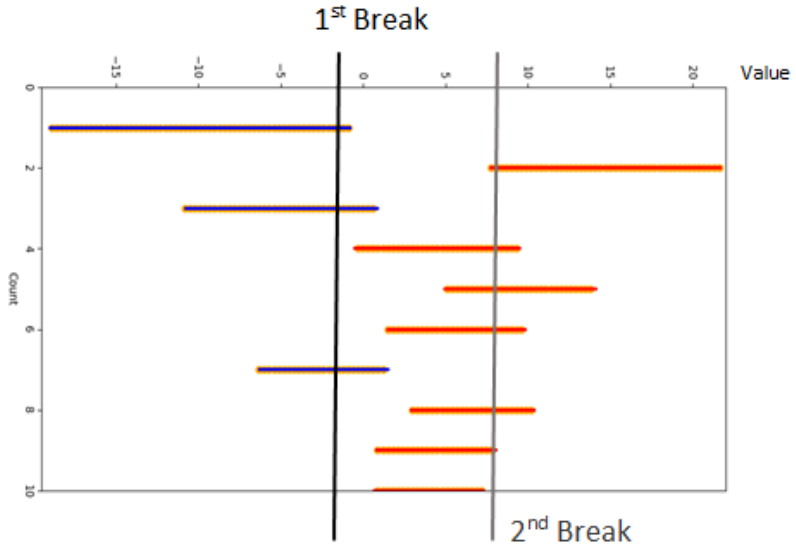


Fig. 2: Intervals between extreme value polygons and closest neighbors (blue: local minimum, red: local maximum) and setting of class breaks

2.3 Setting class breaks

In the last step, a Plane Sweep Algorithm is carried out in order to retrieve optimal intersections (i.e., class boundaries; Fig. 2). For this, a vertical line is swept from left (global minimum) to right (global maximum) in order to retrieve total number of intersection counts for each line. Starting with the most significant extreme value polygon (rf. step 2.2), a class break is set within the corresponding value interval where the number of intersections has its maximum. After identifying the first class break, those segments which intersect this one, are excluded and then the next sweep is performed. This algorithm repeats the above steps until the total number of classes (predefined by the user) have been met or all intervals have been integrated.

3 Results and Evaluation

In this chapter, the task-oriented local extreme method was implemented and tested through datasets with different characteristics. An evaluation was performed with QGIS in order to retrieve a better understanding regarding the relationship between the preservation rate of local extreme values and numbers of classes. In this example, the deprived index from the City of Hamburg at so-called statistic district level in the year 2016 is used (see section 7 for a description of this dataset). It comprises social monitoring indices including unemployment rate, children with immigration background and single parents, long-term social help applicants (SGB II recipients) and education level. The original dataset consists of 847 records. Indices were nominalized and ranged from -19.03 to 21.62 with a standard deviation of 2.926.

3.1 Visual comparison

For comparison purposes, Fig. 3 shows visualizations of this data set using different data classification approaches (with 7 classes) and the same color scheme.
It demonstrates how different classification approaches can influence the recognition of neighboring geographical elements. A local maximum (red) (value 4.14) is preserved with the task-oriented method (aChor), similarly for nature breaks (Jenks). Conversely, equal interval and quantile classification are not able to strictly preserve the local maximum value in this case. The local minimum (blue) (value -4.58) is well preserved with the task-oriented method whereas the other three methods are not able to distinguish this value from its surrounding elements.

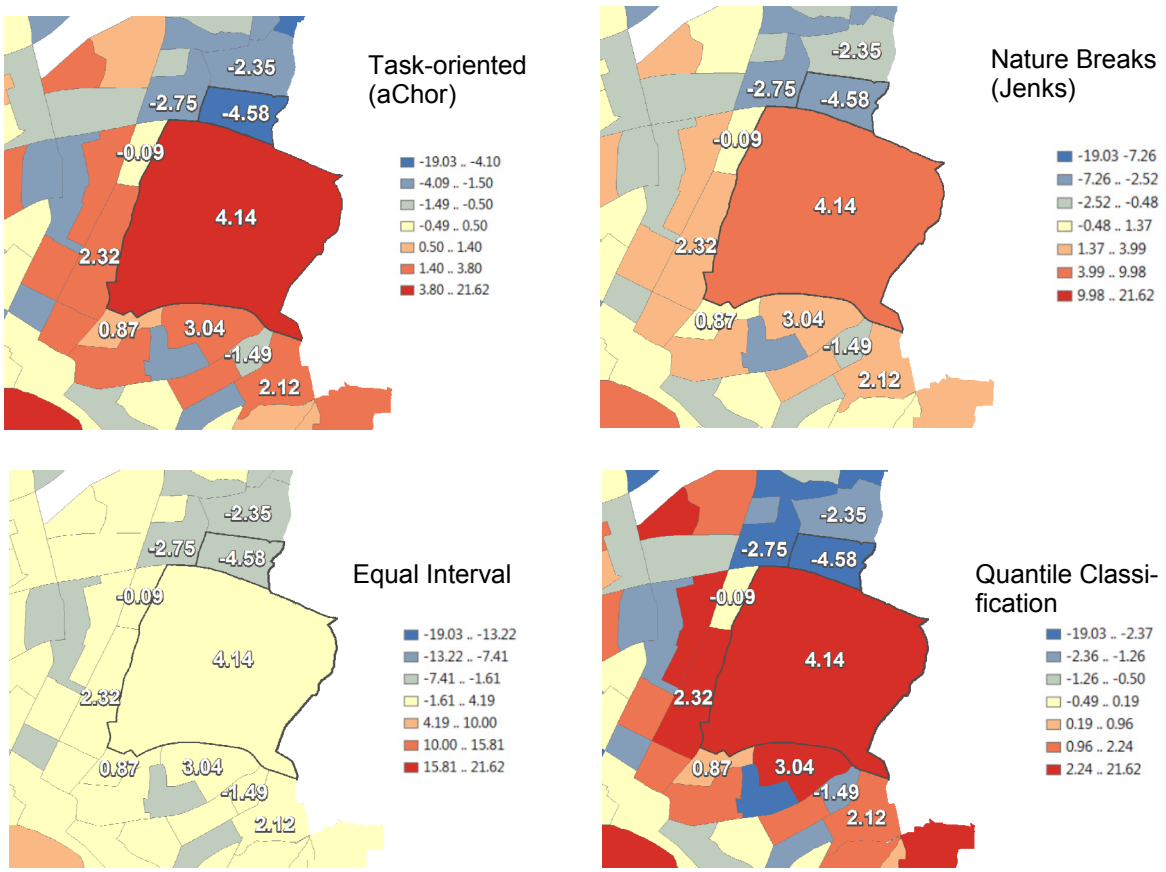


Fig. 3: Comparison of task-oriented (aChor) with conventional data classification methods

3.2 Numerical comparison

In order to get an overall understanding of the local extreme value preservation rate for different classification methods, Table 1 compares aChor with other methods for a fixed number of seven classes.

Tab. 1: Task-oriented data classification in comparison with conventional method (7 classes)

	No. of local maximum values (Max)	No. of local minimum values (Min)	Preserved Max	Preserved Min	Total %
Task-oriented method (aChor)	136	135	91	90	66.8%
Quantile Classification			76	86	59.8%
Natural Breaks Jenks			73	68	52.0%
Geometrical Interval			70	65	49.8%
Equal Interval			37	53	33.2%

Extending this evaluation, Fig. 4 shows dependency of the preservation rate on the number of classes. In this example, aChor reaches a 100 % preservation rate when 36 classes are used. In general, the quantile classification method has a rather good performance with a smaller number of classes; nevertheless it shows no advantage anymore when compared with Jenks with class numbers larger than 20.

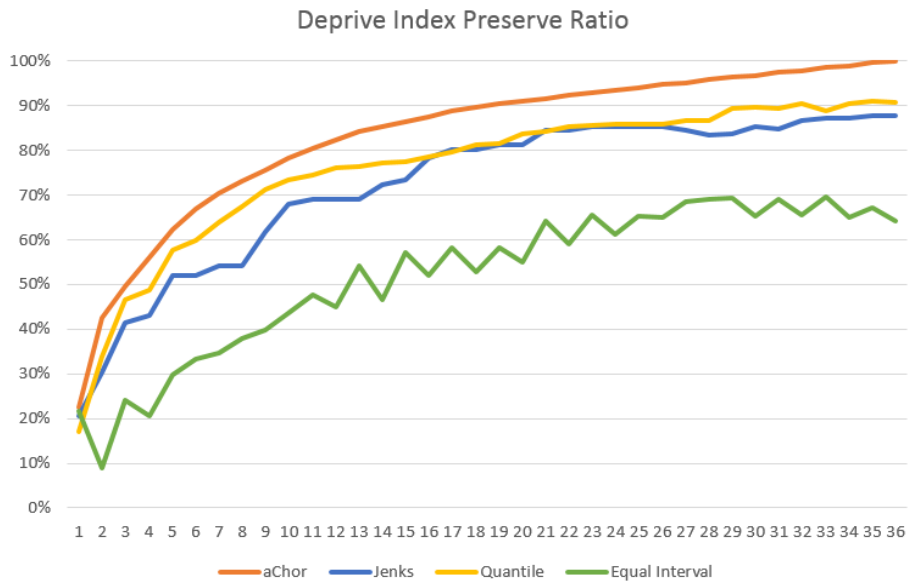


Fig. 4: Preservation rate within different data classification method with the same dataset

4 Conclusion and Discussion

We presented a method that preserves spatial relationships, in particular local extreme values. Our studies (beyond the aforementioned datasets, further examples have been tested (see section 7) and have proven visually and numerically that the rate of preservation for these features at any given number of classes is significantly better compared to conventional data classification methods. Also a 100 % preservation rate is reached with a smaller number of classes compared to all other methods.
Nevertheless, presently this method needs to go through further implementation and adjustment as well as an application of other datasets that show different statistical distributions.

5 Acknowledgements

The work presented here has been conducted within the project *"Aufgabenorientierte Datenklassifikation und Gestaltung von Choroplethenkarten (aChor)"*, funded by the German Research Foundation (DFG).

6 Literature

SCHIEWE, J., 2016a: Preserving attribute value differences of neighboring regions in classified choropleth maps. International Journal of Cartography, **2**(1), 6-19. DOI: 10.1080/23729333.2016.1184555

SCHIEWE, J., 2016b: Aktuelle Forschung und Entwicklung in der Kartographie – das Beispiel der aufgabenorientierten Gestaltung von Choroplethenkarten. In: Harzer, C. (Ed.): GIS-Report 2016/17, Harzer-Verlag: 9-16.

SCHIEWE, J., 2017: Data Classification for Highlighting Polygons with Local Extreme Values in Choropleth Maps. In: Peterson M.P. (ed.): Advances in Cartography and GIScience, Lecture Notes in Geoinformation and Cartography, Springer: 449-459.

7 Attachment: Dataset description for testing purposes

Deprived index:

http://suche.transparenz.hamburg.de/dataset/sozialmonitoring-integrierte-stadtteilentwicklung-bericht-2016-anhang

Yearly average age of the inhabitants of German communities in 2011:

https://www.destatis.de/DE/Methoden/Zensus_/Downloads/2F_BevoelkerungAlterGeschlecht.html

SPD second votes in the Hamburg districts in the federal elections 2017: https://www.statistik-nord.de/wahlen/wahlen-in-hamburg/bundestagswahlen/2017/

Hot-Spot-Analyse zur Unterstützung der Datenklassifizierung – ein Versuch zur Minimierung der Subjektivität in thematischen Karten am Beispiel des Projekts AlpES (Alpine Ecosystem Services)

PETER FRECKMANN[1]

Zusammenfassung: Die Erstellung thematischer Karten setzt in vielen Fällen die Klassifizierung der zu visualisierenden Geodaten voraus. Je nach Wahl einer Klassifizierungsmethode unterscheiden sich die erstellten thematischen Karten mehr oder weniger stark voneinander. Der Kartenersteller muss entscheiden, welche der verschiedenen Klassifizierungsmethoden die Realität am besten abbildet. Die Entscheidung für eine bestimmte Methode impliziert damit eine subjektive Komponente. Der Beitrag zeigt am Beispiel von ausgewählten Indikatoren aus dem Projekt AlpES auf, wie mit Hilfe einer Hot-Spot-Analyse die Datenklassifizierung unterstützt werden kann, um das Ergebnis objektiver zu gestalten.

1 Einleitung

Um Geodaten in thematischen Karten zu visualisieren, ist es häufig notwendig, die Daten zu klassifizieren, um durch quantitative Zusammenfassung räumliche Strukturen erkennbar zu machen. Je nach Wahl einer bestimmten Methode unterscheiden sich die erstellten thematischen Karten mehr oder weniger stark voneinander. Es besteht das Problem, aus verschiedenen Klassifizierungsmethoden die jeweils geeignetste auszuwählen, um die Realität bestmöglich abzubilden. Ein Thema kann also auf unterschiedliche Weise visualisiert und interpretiert werden. Die subjektive Komponente durch den Kartenersteller spielt damit eine nicht zu vernachlässigende Rolle (HRUBY 2016). Der Beitrag soll zeigen, wie mit Hilfe räumlicher, im Gegensatz zu nicht-räumlicher statistischer Verfahren die Datenklassifizierung unterstützt werden kann, um die Subjektivität in thematischen Karten zu minimieren.

2 Nicht-räumliche Datenklassifizierung

Die Klassifizierung der Ausgangsdaten für eine thematische Karte ist notwendig, um Informationen über räumliche Verteilung und Muster visualisieren zu können. Das gilt sowohl für Choroplethenkarten (Variation von Farbton, Helligkeit, Sättigung) als auch für punkt-, linien- oder flächenbezogene Diagrammsignaturen (gestufte Größen). Dabei sollte die Klassenbreite so gewählt werden, dass sie mit der Verteilung der Daten korrespondiert. Da thematische Karten heute fast ausschließlich mit Geoinformationssystemen (GIS) oder entsprechenden Tools erstellt

[1] Hochschule Karlsruhe, Fakultät für Informationsmanagement und Medien, Moltkestr. 30, D-76133 Karlsruhe, E-Mail: peter.freckmann@hs-karlsruhe.de

werden, werden dafür in erster Linie die dort standardmäßig zur Verfügung stehenden Methoden der Klassifizierung verwendet. Dazu gehören insbesondere:

- Äquidistante Klassen – die Klassenbreite ist für jede Klasse gleich groß.
- Natural Breaks – die Klassenbreite richtet sich nach Sprüngen in der Datenreihe, dort wo diese auftreten werden die Klassengrenzen gesetzt. Dadurch werden die Unterschiede in den Werten innerhalb der Klassen minimiert.
- Quantile – allen Klassen ist die gleiche Anzahl an Objekten zugeordnet, d.h. sie haben die gleiche Häufigkeit.

Hinzu kommt die Möglichkeit, Klassengrenzen unabhängig von den vorgegebenen Methoden auch interaktiv durch den Kartenersteller festzulegen. Wie schon häufig in der Fachliteratur diskutiert, führt der Einsatz der genannten Methoden zu unterschiedlichen Ergebnissen in der Kartendarstellung, die dadurch z.T. erheblichen Interpretationsspielraum bieten. (ARMSTRONG et al. 2003; BREWER & PICKLE 2002; MONMONIER 1996; MONMONIER 2005). Allen Methoden gemeinsam ist die Tatsache, dass hier nicht-räumliche statistische Verfahren zur Klassenbildung herangezogen werden (HRUBY 2016). Damit spiegelt der Karteninhalt entweder die Vorstellung des Kartenerstellers über die reale Situation wider, oder es besteht die Gefahr, dass die im GIS standardmäßig voreingestellte Methode übernommen wird, ohne zu hinterfragen, ob diese die Realität am besten abbildet.

3 Räumliche Datenklassifizierung

Im Gegensatz zur nicht-räumlichen Statistik wird in der räumlichen Statistik die subjektive Komponente bei der Bildung von Gruppen reduziert, indem das Erkennen und Zusammenfassen einander ähnlicher und räumlich zusammenhängender Gebiete unterstützt wird (HRUBY 2016). Auch hierzu sei auf die einschlägige Literatur verwiesen (ARMSTRONG et al. 2003; CHEN et al. 2008; DUQUE et al. 2006; KRONENFELD & WONG 2017). Im Folgenden soll ein pragmatischer Ansatz vorgestellt werden, der es erlaubt, mit Hilfe entsprechender GIS-gestützter Funktionen eine Klassenbildung unter Berücksichtigung räumlicher Zusammenhänge durchzuführen (BENNETT et al. 2010).

3.1 Räumliche Autokorrelation – Moran’s I Index

Die geographische Zusammenfassung der Objekte (Punkte, Linien, Flächen) eines Untersuchungsgebietes zu Gruppen, erfolgt in zwei Schritten. Zunächst kann mit dem Moran’s I Index die räumliche Autokorrelation unter den Objekten bezogen auf die Lage und gleichzeitig auf die Merkmalsausprägung gemessen werden. Der Indexwert weist darauf hin, ob die von den Objekten gebildeten Muster zufällig verteilt oder gruppiert sind (BAHRENBERG et al. 2008; GETIS & ORD 1992; GOODCHILD 1986). Der Moran’s I Index basiert auf einer schlussfolgernden Statistik, d.h. die Ergebnisse der Analyse werden immer im Kontext der Null-Hypothese interpretiert. Bei der Moran's I Statistik gibt die Null-Hypothese an, dass das analysierte Attribut zufällig auf die Objekte im Untersuchungsgebiet verteilt wird. Wenn der von der Analyse zurückgegebene p-Wert statistisch signifikant ist, kann die Null-Hypothese abgelehnt werden.

Der Moran's I Indexwert nimmt dabei Werte zwischen -1 und + 1 an. Ein positiver Indexwert zeigt eine Tendenz zur Clusterbildung und ein negativer Indexwert eine Tendenz zur Streuung an. Der Indexwert lässt sich wie folgt berechnen:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2} \qquad (1)$$

mit

z_i	Abweichung des Attributwertes des Objekts i vom Mittelwert
$\omega_{i,j}$	räumliche Gewichtung (Distanz) zwischen den Objekten i und j
n	Anzahl der Objekte
S_0	Summe über alle Distanzen

3.2 Hot-Spot-Analyse – Getis-Ord Gi*-Statistik

Zeigt der Moran's I Indexwert eine Tendenz zur Gruppierung, lassen sich im zweiten Schritt die Objekte über eine Clusteranalyse zu zusammenhängenden Gebieten aggregieren. Mit Hilfe der Getis-Ord Gi*-Statistik lässt sich feststellen, wo Objekte mit hohen oder niedrigen Werten zu räumlicher Clusterbildung neigen (GETIS & ORD 1995). Dabei wird jedes einzelne Objekt unter Berücksichtigung der benachbarten Objekte bewertet:

$$G_i^* = \frac{\sum_{j=1}^{n} \omega_{i,j} x_j - X \sum_{j=1}^{n} \omega_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} \omega_{i,j}^2 - \left(\sum_{j=1}^{n} \omega_{i,j}\right)^2\right]}{n-1}}} \qquad (2)$$

mit

x_j	Attributwert des Objekts *j*
$\omega_{i,j}$	räumliche Gewichtung (Distanz) zwischen den Objekten *i* und *j*
n	Anzahl der Objekte
X	Arithmetischer Mittelwert des Attributs x
S	Standardabweichung

Ein statistisch signifikanter Hot-Spot ist ein Objekt mit einem hohen Attributwert, das außerdem von anderen Objekten mit hohen Werten umgeben ist. Für die identifizierten Cold-Spots verhält es sich umgekehrt. Das Ergebnis der Analyse zeigt damit gleichzeitig die für die jeweilige Forschungsfrage relevanten Gebiete. Damit ist also auch ein Regionalisierungseffekt verbunden.

4 Anwendungsbeispiel - AlpES

Das beschriebene Verfahren wurde im Interreg Alpine Space Projekt AlpES – Alpine Ecosystem Services getestet (http://www.alpine-space.eu/projects/alpes/en/home). Im Rahmen dieses Projekts werden Ökosystemleistungen als Hauptsäulen einer grünen Wirtschaft im Alpenraum untersucht, um Ökosysteme und ihre Leistungen für die Gesellschaft zu schützen und nachhaltig

nutzbar zu machen. Unter Ökosystemleistungen versteht man die Bereitstellung von Gütern und Leistungen durch die Natur. Darunter fallen sowohl Versorgungsleistungen mit Verbrauchsgütern wie etwa Nahrungsmittel oder Holz, als auch die von der Natur bereitgestellten Regulierungsleistungen wie die Reinigung von Luft und Wasser (TAPPEINER & TASSER 2017).
Ziel des Projektes ist es, ein Verständnis für Ökosystemleistungen zur Steuerung regionaler und transnationaler Umweltpolitik zu etablieren. An dem Projekt sind 10 Partner aus allen Alpenanrainerstaaten unter der Leitung der EURAC Research (Bozen) beteiligt. Das Gesamtbudget beträgt 2,3 Mio. Euro. Das Untersuchungsgebiet umfasst 16.785 Gemeinden, für die 8 umweltrelevante Indikatorengruppen ausgewertet werden (Abb. 1). Einen wichtigen Teil des Projekts bildet das Work Package „Mapping & Assessment“. Hier werden auf Basis der Indikatoren eine Vielzahl an thematischen Karten erstellt, aus denen Handlungsempfehlungen für die Umweltpolitik abgeleitet und die in einem WebGIS der Öffentlichkeit zugänglich gemacht werden. Beides setzt thematische Karten voraus, die möglichst wenig Spielraum für Fehlinterpretationen bieten sollten. Das gilt sowohl für die Nutzung der Karten durch die Projektbeteiligten als auch für die späteren Nutzer der Projektergebnisse mit unterschiedlich fachlichen Ausrichtungen und regionalen, insbesondere grenzüberschreitenden Kenntnissen.

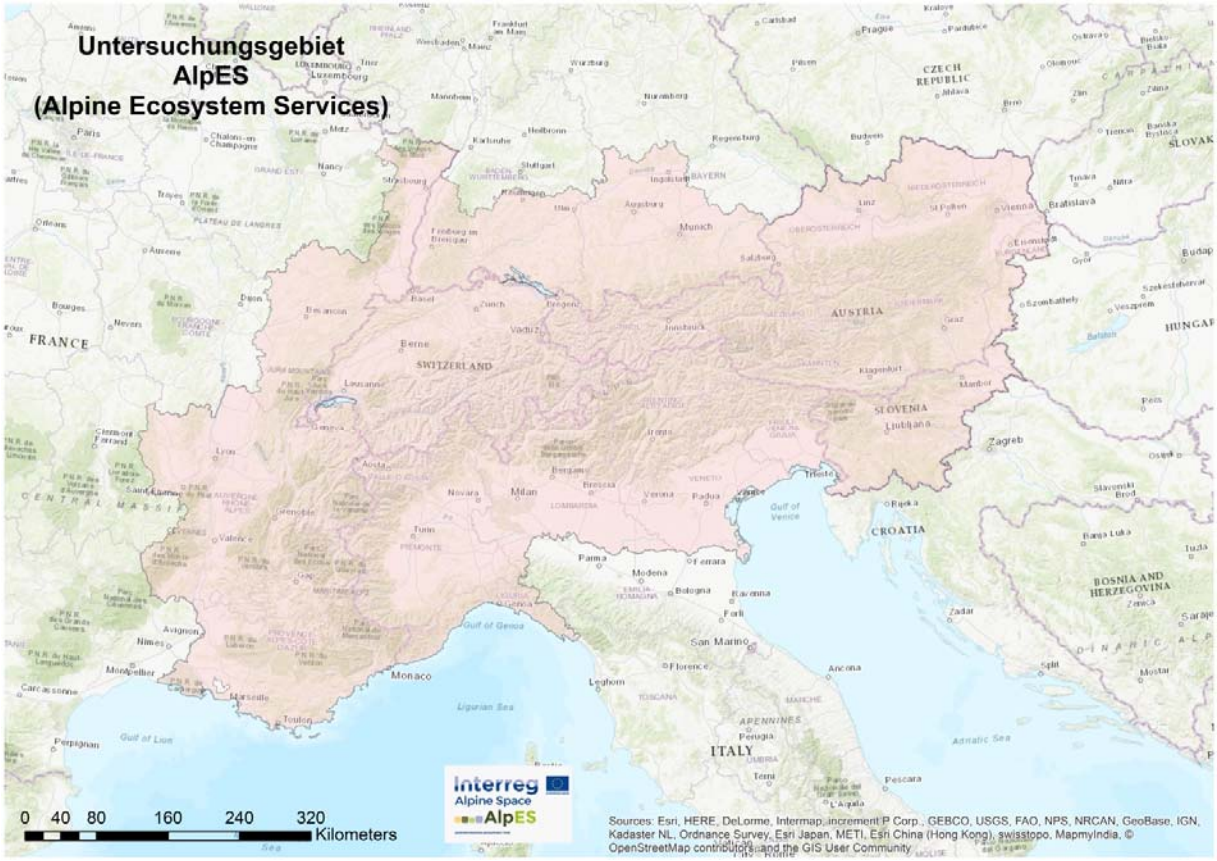


Abb. 1 AlpES Untersuchungsgebiet

4.1 Verfahrensablauf

Abb. 2 verdeutlicht den Ablauf zur Erstellung thematischer Karten. Bei Verwendung nicht-räumlicher Statistik werden die Daten mit den standardmäßigen Methoden auf Klassen verteilt und unter Verwendung kartographischer Modelle in Karten visualisiert. Beim Einsatz räumlicher Statistik wird zunächst der Moran's I Index für die darzustellende Variable bestimmt. Zeigen die Objekte eine Tendenz zur Gruppenbildung wird im nächsten Schritt die Hot-Spot-Analyse durchgeführt. Als Ergebnis erhält man zusammenhängende Flächen mit hohen bzw. niedrigen Attributwerten sowie Gebiete die keine signifikanten Gruppen bilden. Wie später an den Anwendungsbeispielen noch gezeigt wird, lassen sich aus der Analyse Klassengrenzen ableiten, die die Wertespanne der Merkmalsausprägung zwischen den Hot- und Cold-Spots darstellen.
Man hat jetzt die Möglichkeit, die standardmäßige Klassenbildung mit dem Ergebnis der Hot-Spot-Analyse zu vergleichen, um auf dieser Grundlage, die Karte mit der am besten geeigneten Klassifizierungsmethode zu verwenden oder aber die Klassen mit den aus der Hot-Spot-Analyse abgeleiteten Klassengrenzen zu definieren. Damit hätte man eine Karte erstellt, deren Datenklassifizierung auf räumlicher Statistik beruht.

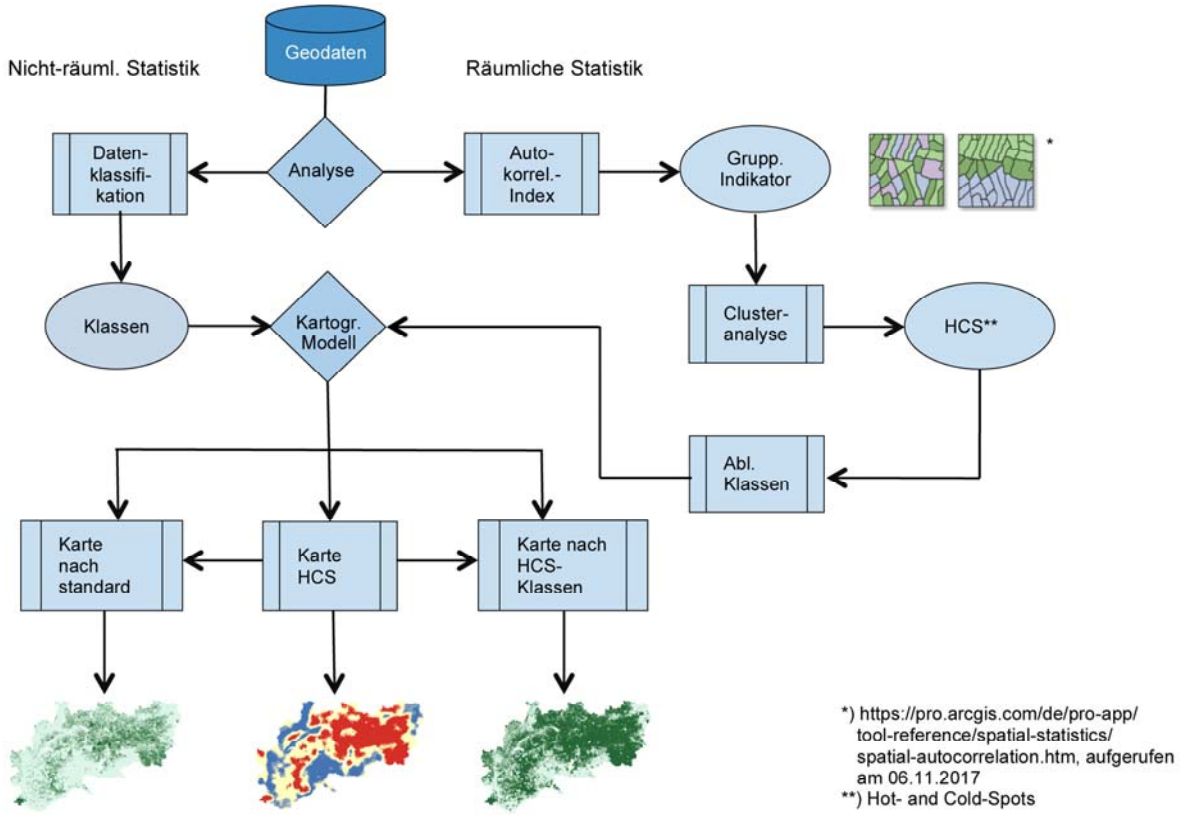


Abb. 2: Erstellung thematischer Karten mit nicht-räumlicher und räumlicher Statistik

4.2 Ergebnisse für ausgewählte Indikatoren

Für die Anwendung der Hot-Spot-Analyse wurden aus den 8 Indikatorengruppen vier Einzel-Indikatoren ausgewählt. Tab. 1 zeigt den für diese Indikatoren jeweils berechneten Moran's I Indexwert. Alle Werte sind positiv und bewegen sich zwischen 0,72 und 0,03. Daraus lässt sich schließen, dass die Gemeindegebiete mehr oder weniger stark zur Gruppenbildung tendieren.

Tab. 1: Indikatoren

AlpES ausgewählte Indikatoren	**Moran's I Indexwert**
Annual Average Available Water Runoff with Drinking Water Quality	0,72
Population Density (Inhabitants/km^2)	0,55
Fodder Production of Alpine Grassland, Supply of DM per Area	0,03
Grassland (ha)	0,17

Abb. 3 gibt einen Überblick über die Häufigkeitsverteilung der Daten. Dabei wird deutlich, dass sich insbesondere bei den Indikatoren „Population Density" und „Fodder Production" der größte Teil der Merkmalsausprägungen im Bereich kleiner Werte liegt und nur wenige Gemeinden sehr hohe Werte aufweisen, wodurch die Klassenbildung mit den Standard-Klassifizierungsmethoden erschwert wird.

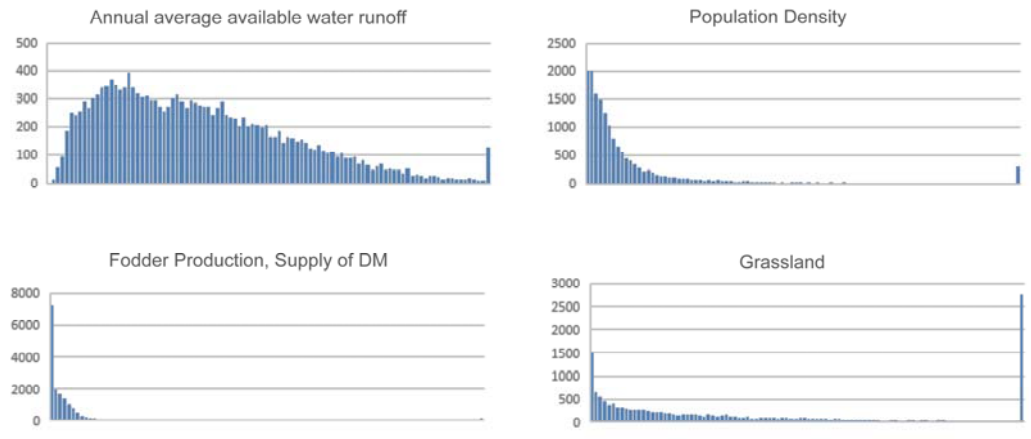
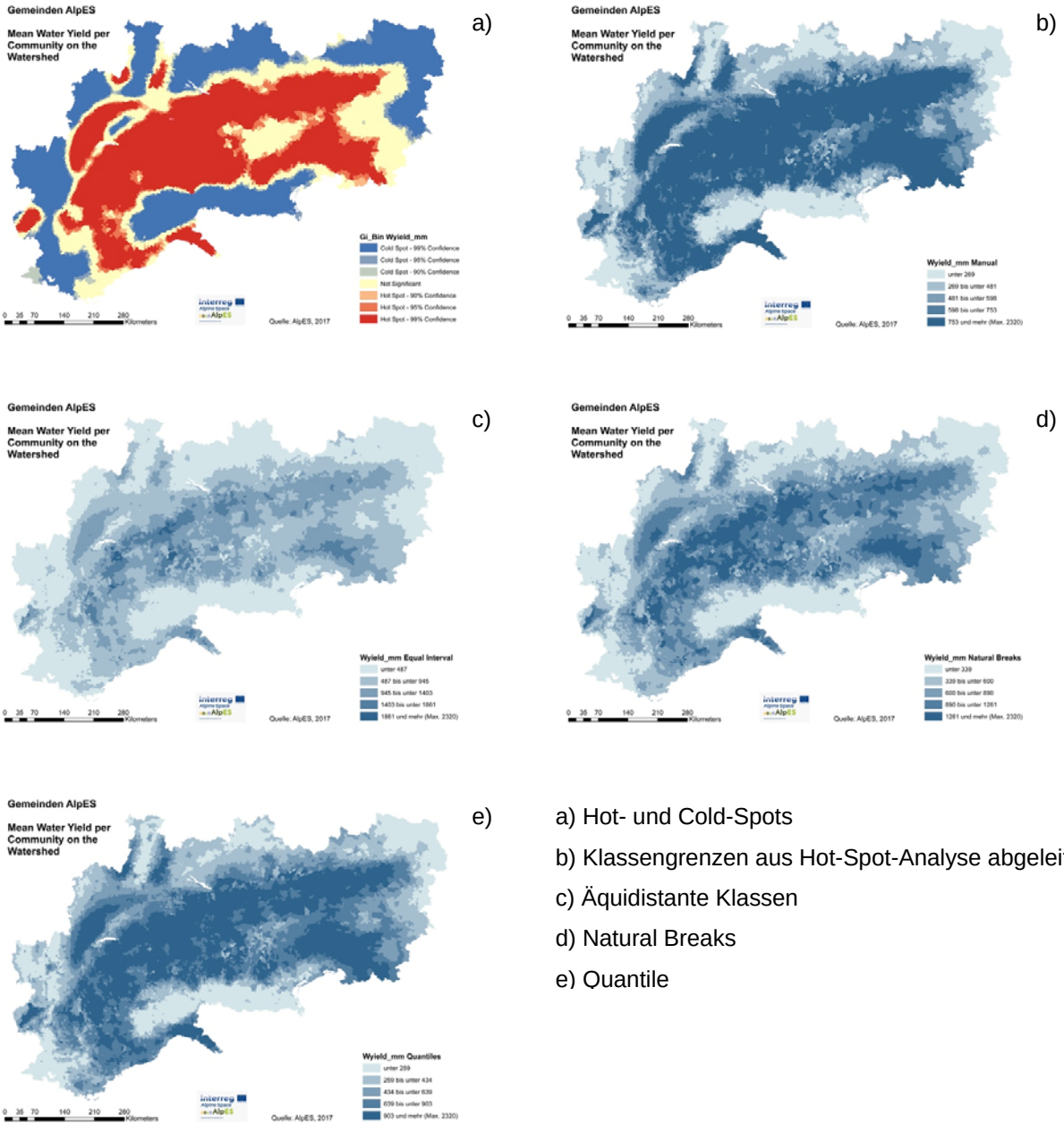


Abb. 3: Verteilung der ausgewählten Indikatoren

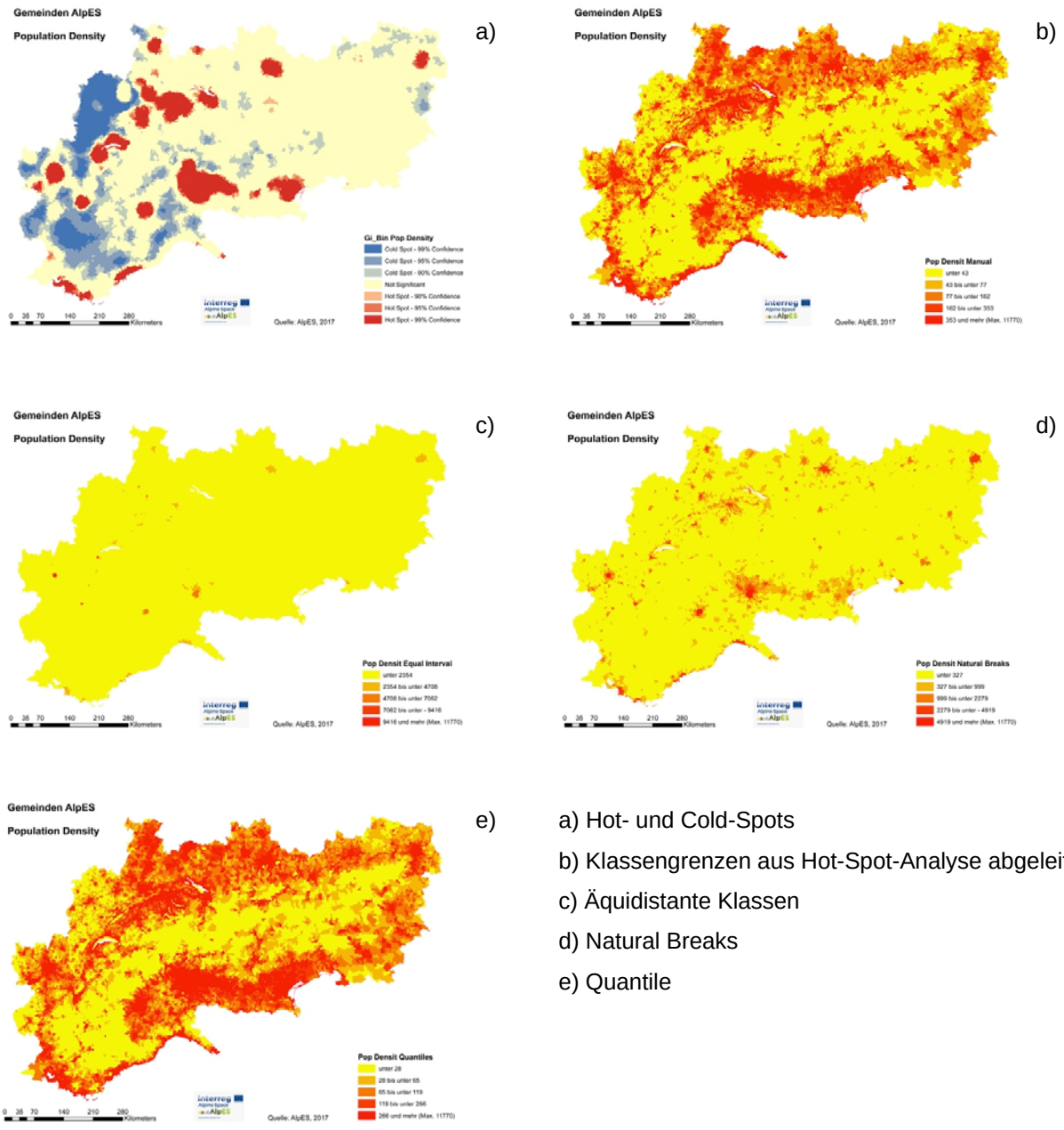
Im Folgenden wurden für die vier ausgewählten Indikatoren jeweils 5 thematische Karten erstellt, die das Ergebnis der Hot-Spot-Analyse und verschiedene Klassifizierungsmethoden zeigen. Alle Berechnungen und die Erstellung der Karten wurden mit der ArcGIS Spatial Statistics Toolbox durchgeführt. Vergleichbare Verfahren stehen natürlich auch in anderen Geoinformationssystemen zur Verfügung, so dass das verwendete Tool austauschbar ist.
Abb. 4 zeigt den Mittleren Oberflächenwasserabfluss im Untersuchungsraum. Die Karte der Hot-Spot-Analyse spiegelt das Relief exakt wider. Insbesondere im Bereich der zentralen Alpen liegen die Hot-Spots, da hier der höchste Oberflächenwasserabfluss auftritt. ArcGIS ermittelt die Hot- und Cold-Spots jeweils für das 90, 95 und 99% Signifikanzniveau. Durch die Aggregation der im Untersuchungsraum verteilten Hot- bzw. Cold-Spots für jedes Signifikanzniveau und gleichzeitiger Berechnung des Mittelwertes der jeweiligen Merkmalsausprägung für die aggregierten Flächen, erhält man Mittelwerte, die die Wertespanne der Merkmalsausprägung zwischen den Hot- und Cold-Spots repräsentieren. Diese können als Grundlage für eine Klassenbildung verwendet werden. Damit ergibt sich, wie in Karte b) zu sehen ist, ein fast identisches Bild wie in der Karte der Hot- und Cold-Spots. Die Klassifizierungsmethoden Äquidistante Klassen, Karte c) und Natural Breaks, Karte d) zeigen die räumliche Struktur weniger deutlich. Dagegen entspricht die Karte mit der Methode Quantile sehr genau dem Ergebnis der Hot-Spot-Analyse.
Abb. 5 enthält die Karten für den Indikator Einwohnerdichte. Die auf Basis der Hot-Spot-Analyse festgelegten Klassengrenzen, Karte b) zeigen die auch in Karte a) zu sehenden Kernstädte im Untersuchungsraum und darüber hinaus eine detaillierte Differenzierung der sie umgebenden Agglomerationsräume. Die Klassifizierungsmethoden Äquidistante Klassen und Natural Breaks liefern dagegen ein Ergebnis, dass die reale Situation nicht besonders gut wieder gibt. Quantile dagegen zeigen wiederum eine Situation, die der Realität sehr nahe kommt.
Abb. 6 zeigt die Karten für den Indikator Futterproduktion pro ha. Die Hot-Spot-Analyse ergab nur einige wenige kleine Gebiete mit hohen Werten. Der überwiegende Teil der Gemeinden ist nicht signifikant hinsichtlich einer Gruppierung zu Hot- oder Cold-Spots. Verwendet man die auf Basis der Hot-Spot-Analyse ermittelten Klassengrenzen, ergibt sich ein interessantes Bild. Die Hot-Spots finden sich als einzelne Gemeindeflächen wieder. Hier handelt es sich um Grünlandflächen innerhalb dieser Gemeinden, die sowohl von der Höhenlage als auch den klimatischen Bedingungen besonders gute Bedingungen haben und damit eine hohe Produktivität aufweisen, aber nur vereinzelt auftreten. Die in Deutschland und in der Schweiz zu erkennenden größeren zusammenhängenden Flächen mit höheren Werten, die man in der Karte der Hot-Spot-Analyse nicht sieht, gehören zu größeren Almengbieten, die als Futterflächen mit hoher Energieausbeute ausgewiesen worden sind. Insofern ergibt sich eine realitätsnahe Abbildung. Deutlich wird auch, dass Äquidistante Klassen und Natural Breaks keine brauchbaren Ergebnisse liefern und die Klassifizierung mit Quantilen die Situation visuell überbewertet.
Wurden mit den drei beschriebenen Indikatoren relative Werte betrachtet, soll mit Abb. 7 auch auf Absolutwerte eingegangen werden, da auch Diagramme in gestuften Größen klassifizierte Werte repräsentieren können. Die Karten b) bis e) zeigen einen Ausschnitt des Untersuchungsraums, da die Kreisdiagramme in der Gesamtansicht kaum einzeln sichtbar wären. Zum besseren Vergleich wurden in den Kartenausschnitten die Hot- und Cold-Spots als weitere Ebene unter die Kreisdiagramme gelegt. Deutlich wird hier, dass die Klassifizierung auf Basis

der Hot-Spot-Analyse wieder sehr gut geeignet ist. Dort wo die Hot-Spots liegen, findet man die größten Kreise und umgekehrt. Äquidistante Klassen geben diese Struktur nicht wieder, Natural Breaks nur eingeschränkt. Wohingegen Quantile in etwa der Karte b) entsprechen, also die Situation gut abbilden.



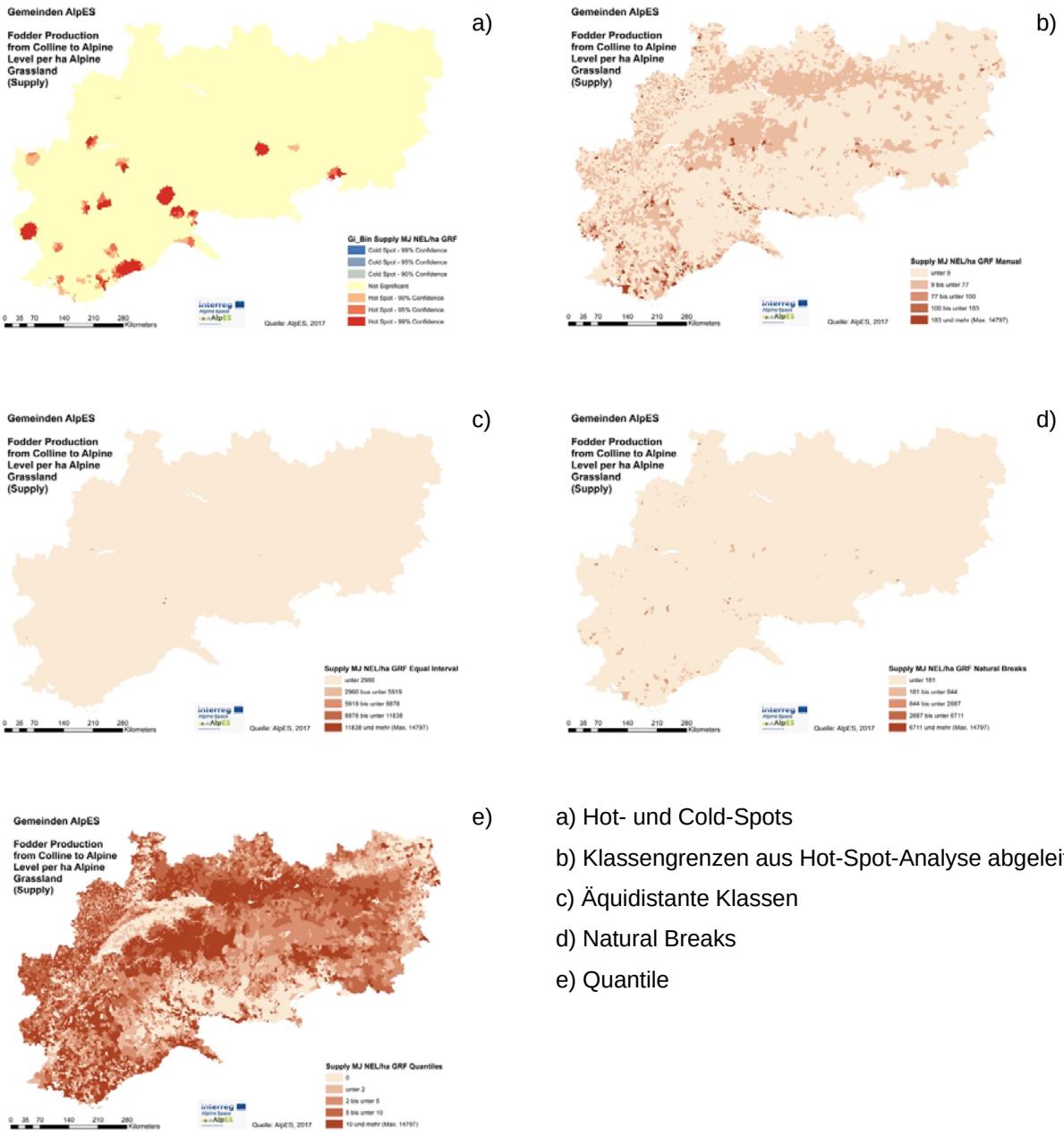
a) Hot- und Cold-Spots
b) Klassengrenzen aus Hot-Spot-Analyse abgeleitet
c) Äquidistante Klassen
d) Natural Breaks
e) Quantile

Abb. 4: Thematische Karten für den Indikator Annual Average Available Water Runoff



a) Hot- und Cold-Spots
b) Klassengrenzen aus Hot-Spot-Analyse abgeleitet
c) Äquidistante Klassen
d) Natural Breaks
e) Quantile

Abb. 5: Thematische Karten für den Indikator Population Density



a) Hot- und Cold-Spots

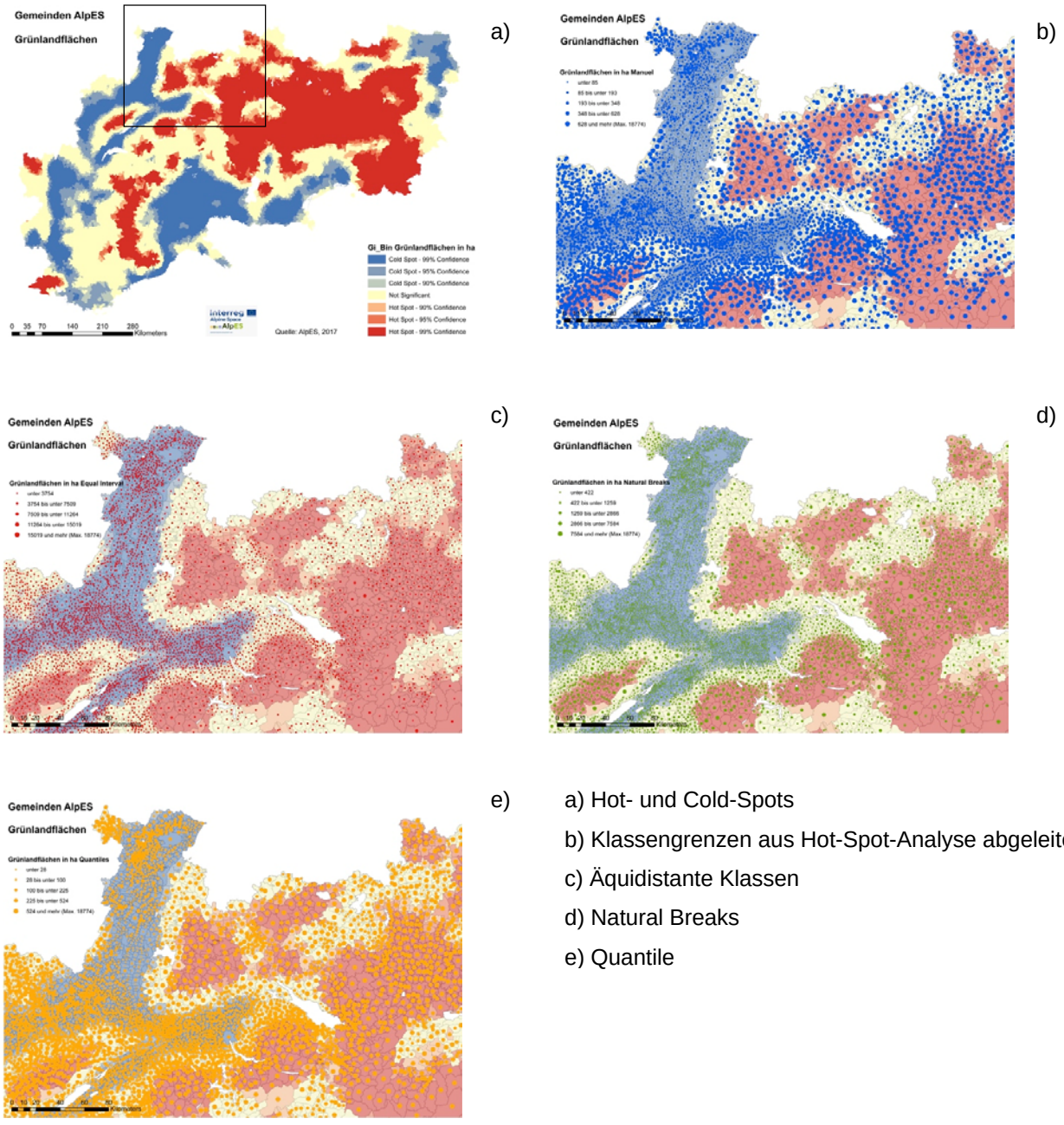
b) Klassengrenzen aus Hot-Spot-Analyse abgeleitet

c) Äquidistante Klassen

d) Natural Breaks

e) Quantile

Abb. 6: Thematische Karten für den Indikator Fodder Production of Alpine Grassland



a) Hot- und Cold-Spots
b) Klassengrenzen aus Hot-Spot-Analyse abgeleitet
c) Äquidistante Klassen
d) Natural Breaks
e) Quantile

Abb. 7: Thematische Karten für den Indikator Grassland

5 Zusammenfassung

Betrachtet man die oben beschriebenen Ergebnisse, lässt sich bezogen auf die verwendeten Beispieldaten festhalten, dass die Hot-Spot-Analyse eine Klassifizierung mit den Standardmethoden Äquidistante Klassen, Natural Breaks und Quantile dahingehend unterstützt, dass die Subjektivität, die mit den nicht-räumlichen Klassifizierungsmethoden verbunden ist, durch die Gegenüberstellung mit den Ergebnissen der Hot-Spot-Analyse, reduziert werden kann. Im Einzelnen lassen sich folgende Vorteile erkennen:

- Mit der Hot-Spot-Analyse lässt sich, wie generell mit Clusterverfahren, eine Regionalisierung innerhalb des Untersuchungsraumes durchführen. Hot- und Cold-Spots bilden bezogen auf ein bestimmtes Merkmal weitgehend homogene Gebiete. Das ist umso ausgeprägter, je höher der Moran's I Index ausfällt.

- Bei vorliegender Autokorrelation erlaubt es die Hot-Spot-Analyse die standardmäßig zur Verfügung stehenden Klassifizierungsmethoden dahingehend zu überprüfen, ob diese die vorhandenen räumlichen Verteilungen und Muster wieder geben und hilft dabei eine fundiertere Entscheidung für eine der Standardmethoden zu treffen.

- Unabhängig davon besteht auch die Möglichkeit, Klassengrenzen auf Basis der Hot-Spot-Analyse zu bestimmen. Wie die Anwendungsbeispiele gezeigt haben, ergibt sich damit, vorallem bei einer deutlichen Tendenz zur Gruppenbildung der Objekte, ein Ergebnis, das den tatsächlichen realen räumlichen Strukturen näher kommt als insbesondere mit den Standardmethoden Äquidistante Klassen und Natural Breaks.

6 Literaturverzeichnis

ARMSTRONG, M.P., XIAO, N. & BENNETT, D.A., 2003: Using Genetic Algorithms to Create Multicriteria Class Intervals for Choropleth Maps. Annals of the Association of American Geographers, **93**(3), 595-623.

BAHRENBERG, G., GIESE, E., MEVENKAMP, N. & NIPPER, J., 2008: Statistische Methoden in der Geographie. Multivariate Statistik, **2**, 352-370.

BENNETT, L., D'ACOSTA, J. & VALE, F., 2010: Spatial Data Mining: A Deep Dive into Cluster Analysis. https://blogs.esri.com/esri/arcgis/2010/07/13/spatial-statistics-resources/, letzter Zugriff 25.09.2017.

BREWER, C. & PICKLE, L., 2002: Evaluation of Methods for Classifying Epidemiological Data on Choropleth Maps in Series. Annals of the Association of American Geographers, **92**(4), 662-681.

CHEN, J., ROTH, R.E., NAITO, A.T., LENGERICH, E.J. & MACEACHREN, A.M., 2008: Geovisual analytics to enhance spatial scan statistics interpretation; an analysis of U.S. cervical cancer morality. International Journal of Health Geographics, **7**(57), 1-18

DUQUE, J.C., RAMOS, R. & SURINACH, J., 2007: Supervised Regionalization Methods: A Survey. Research Institute of Applied Economics 2006, Working Papers 2006/**8**, 1-31.

GETIS, A. & ORD J.K., 1992: The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis **24**(3), 189-206.

GETIS, A. & ORD J.K., 1995: Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. Geographical Analysis **27**(4), 286-306.

GOODCHILD, M.F., 1986: Spatial Autocorrelation. Catmog 47, Geo Books.

HRUBY, F. 2016: 190 Jahre Choroplethenkarten – Ein Zwischenresümee. Kartographische Nachrichten **2**/16, 58-65

KRONENFELD, B.J. & WONG, D.W.S., 2017: Visualizing statistical significance of disease clusters using cartograms. International Journal of Health Geographics, **16**(19), 1-13

MONMONIER, M. 1996: How to Lie with Maps, Second Edition.

MONMONIER, M. 2005: Lying with Maps, Statistical Science, **20**(3), 215-222.

TAPPEINER, U. & TASSER, E., 2017: Ökosystemleistungen in den Alpen. Natur und Landschaft. DOI: 10.17433/9.2017.50153509.440-445.

Hochdurchsatz-Phänotypisierung von Weizen unter Hitze- und Trockenstress mit Hilfe von UAVs

KAROLIN KUNZ[1], YUNCAI HU[1] & URS SCHMIDHALTER[1]

Zusammenfassung: Um die Auswirkung des Klimawandels auf die Nahrungsmittelproduktion schon heute abschätzen und angepasste Getreidesorten selektieren zu können, wurde in der Saison 2016/17 ein Feldversuch in der Republik Moldau durchgeführt. Im dortigen kontinentalen Klima mit sehr trockenen und heißen Sommern wurden 40 Weizensorten verschiedener Herkunft angebaut und deren Wachstum mit unterschiedlichen Methoden der Hochdurchsatz-Phänotypisierung erfasst. Durch Messungen von Spektralreflexion und Pflanzentemperatur konnte der Einfluss des abiotischen Stresses analysiert werden. Hierbei konnten Unterschiede zwischen Züchtungen aus Osteuropa und Deutschland erkannt und verglichen werden. Die deutschen Sorten – Linien und Hybride – zeigten durch die Einwirkung von Hitze und Trockenheit u.a. starke Ertragseinbußen.

1 Einleitung

Weizen ist eine der wichtigsten Getreidearten weltweit und wird jährlich auf mehr als 220 Mio. ha angebaut (SHIFERAW et al. 2013). Allerdings werden Weizenproduzenten weltweit vor große Herausforderungen gestellt. Bis zum Jahr 2050 wird die Weltbevölkerung auf über 9 Milliarden Menschen und die Nachfrage nach landwirtschaftlichen Produkten um bis zu 50 % ansteigen. Gleichzeitig nimmt die landwirtschaftlich nutzbare Fläche durch Erosion und Versiegelung ab. Um eine ausreichende Versorgung mit Nahrungsmitteln gewährleisten zu können, müssen daher die Erträge pro Fläche deutlich gesteigert werden. Die Erträge werden jedoch durch viele biotische und abiotische Faktoren limitiert. Hierzu zählt auch der Klimawandel, einhergehend mit zunehmenden Hitze- und Trockenheitsereignissen (FAO 2017). Um eine zufriedenstellende Produktion sicher zu stellen, müssen Sorten aller landwirtschaftliche Kulturen, u.a. Weizen, gezüchtet werden, die an hohe Temperaturen und eine knappe Wasserversorgung angepasst sind und unter diesen Bedingungen hohe Flächenerträge erzielen können. Hierzu ist es einerseits notwendig, Sorten zu sichten, die schon heute in ariden Gebieten angebaut und gezüchtet werden, andererseits Hochertragssorten aus gemäßigten Klimazonen abiotischem Stress in Form von Hitze und Trockenheit auszusetzen und jeweils die Entwicklung und das Wachstum der Pflanzen, sowie den Ertrag zu erfassen. Je nach Anzahl der getesteten Sorten und Größe der Anbaufläche kann die Phänotypisierung sehr arbeits- und zeitaufwändig sein. Eine Erleichterung wird hier durch verschiedene Methoden der Hochdurchsatz-Phänotypisierung geboten. Diese umschließt die Erfassung verschiedener pflanzenphysiologischer Parameter mit Hilfe von Spektrometern und Thermalkameras, die entweder händisch oder mit Hilfe von Drohnen (Unmanned Aerial Vehicle – UAV) durchgeführt werden kann.

[1] Technische Universität München, Lehrstuhl für Pflanzenernährung, Emil-Ramann-Straße 2, D-85354 Freising, E-Mail: [kunz, hu, schmidhalter]@wzw.tum.de

2 Material und Methodik

2.1 Feldversuch

In der Saison 2016/2017 wurde ein Feldversuch im Norden der Republik Moldau in der Stadt Bălți durchgeführt. 40 verschiedene Sorten Winterweizen wurden in drei Wiederholungen angebaut, die Parzellen waren jeweils 5 × 1 m groß. 20 der getesteten Weizensorten stammen aus osteuropäischen Ländern (Rumänien, Ukraine, Bulgarien und Republik Moldau), bei den anderen 20 Sorten handelt es sich um deutsche Hochertragssorten, davon 16 Linien und vier Hybriden.

Tab. 1: Vergleich der Standorte Freising (D) und Bălți (MD)

	Freising	**Bălți**
Breitengrad	47,46	48,24
Längengrad	27,56	11,44
Boden	Schluffiger Lehm	Schwarzerde
Jährlicher Niederschlag ∅ (mm)	530	800
Jahresmitteltemperatur ∅ (°C)	10,3	7,5
Niederschlag während Vegetationsperiode ∅ (mm)	200	350
Lufttemperatur im Sommer ∅ (°C)	18,5 - 21	13,5

Während der Vegetationsperiode ist die Lufttemperatur im Norden Moldaus im Durchschnitt 5-8°C höher (bis 21°C) als in Freising (13,5°C), der Niederschlag jedoch deutlich geringer (Bălți 200 mm, Freising 350 mm) (Tabelle 1). Zum Zeitpunkt der Körnerreife des Weizens herrschen somit Bedingungen, an die die deutschen Weizensorten nicht angepasst sind und unter starkem Hitze- und Trockenstress leiden. Dadurch kommt es zu starken Ertragseinbußen. Die osteuropäischen Sorten sind für solche klimatischen Bedingungen gezüchtet, erbringen allerdings geringere Erträge als die deutschen Sorten in gemäßigtem Klima. Durch die Lage auf nahezu dem gleichen Breitengrad eignet sich der moldawische Standort gut, die Wachstumsbedingungen und die Pflanzenentwicklung mit der in Freising zu vergleichen.

2.2 Messung von Spektralreflexion und Temperatur der Pflanzen

Während des Wachstums wurden von den einzelnen Parzellen regelmäßig mit einem passiven Spektralsensor (tec5 HandySpec Field 11-01360xx) die Reflexion der elektromagnetischen Strahlung zwischen 302 und 1148 nm erfasst. Dieses breite Messspektrum ermöglicht die Berechnung sog. Vegetationsindizes, welche Informationen über den Zustand der Pflanzen, Nährstoff- und Wasserhaushalt geben können. Berechnet wurden der Water Index WI (R900/R970) (PEÑUELAS et al. 1997), der Normalized Differenced Vegetation Index NDVI ((R780−R670)/(R780 + R670)) (ROUSE et al. 1974) und der Red Edge Inflection Point REIP (700 + 40((R670 + R780)/2−R700)/(R740 −R700)) (GUYOT et al. 1974). Der Water Index dient zur Beurteilung des Pflanzenwassergehaltes, der NDVI erlaubt Rückschlüsse über den

Chlorophyllgehalt und die Vitalität der Pflanze. Der REIP ist hingegen abhängig von der Biomasseentwicklung und zeigt zudem einen deutlichen Zusammenhang mit dem Stickstoff (N) -Gehalt der Pflanzen.
Zusätzlich wurden mit Thermalkameras Bilder der Pflanzen gemacht. Hierbei kamen die tragbare Kamera Fluke Ti400 und die Drohne albris der Firma senseFly zum Einsatz. Die technischen Daten sind in Tabelle 2 aufgeführt. Mit der Fluke-Kamera wurde jede Parzelle in einem Winkel von ca. 45° fotografiert, so dass möglichst viel Pflanzenoberfläche erfasst werden konnte. Anschließend wurden mit der Software LabView Fluke (National Instruments, Version 12.0f3) die Pixel von Boden und Pflanzen identifiziert und getrennt, sodass nur die Temperatur der Pflanzenoberfläche in Betracht gezogen wurde. Die Drohne flog in einer Höhe von 17.2 m und fotografierte mit einer Überlappung von 80 %. Dadurch wurde jede Parzelle mehrfach abgebildet. So konnte von jeder Parzelle mindestens eine Thermografie in Nadirposition gemacht werden. Eine Auftrennung nach Boden- und Pflanzentemperatur war hier nicht möglich, die angegebenen Werte sind jeweils Durchschnittswerte für die Temperatur einer ganzen Parzelle.

Tab. 2: Technische Daten der verwendeten Thermalkameras

	Fluke Ti400	**Drohne Albris**
Auflösung	320 × 240 Pixel	80 × 60 Pixel
Sichtfeld (FOV)	24° × 17°	50° (horizontal)
RGB-Kamera	38 MP	5 MP

3 Ergebnisse und Diskussion

Die berechneten Indizes zeigten deutliche Unterschiede zwischen den Weizensorten aus Deutschland und aus den osteuropäischen Ländern (Abb. 1).
Die Werte des WI und NDVI zeigen über die Vegetationsperiode hinweg einen recht ähnlichen Verlauf: Beide zeigen am 31.5.2017 zur Blüte (Zadoks Skala 55-61) (ZADOKS et al. 1974) ihre höchsten Werte und sinken danach ab. Dabei ist auffällig, dass die Werte der osteuropäischen Sorten jeweils früher und schneller absinken als die der deutschen Linien. Beim WI zeigen die Hybridsorten ab Mitte Juni die geringsten Werte, beim NDVI liegen sie zwischen den osteuropäischen und den deutschen Kultivaren. Auffällig ist, dass bei beiden Indizes die deutschen Sorten am längsten höhere Reflexionswerte aufweisen.

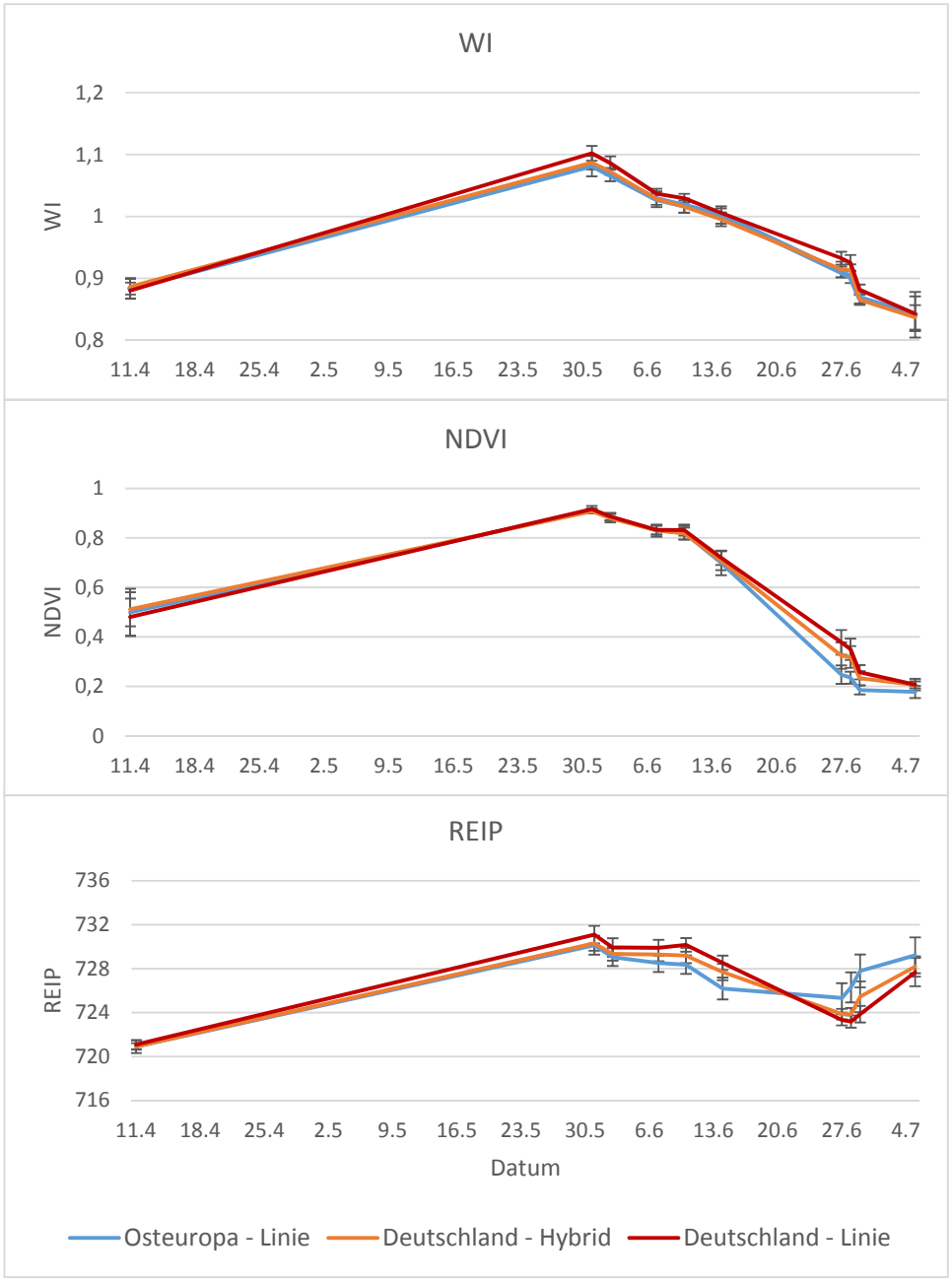


Abb. 1: Water Index, NDVI und REIP osteuropäischer und deutscher Weizensorten im Verlauf der Vegetationsperiode

Die Werte des REIP-Index zeigen einen anderen Verlauf. Zwar erreichen auch sie am 31.5. zur Blüte ihr Maximum, jedoch sinken sie anschließend nicht ganz so schnell ab und zeigen sogar ab Ende Juni wieder einen raschen und deutlichen Anstieg. Hierbei ist auffällig, dass die Werte der deutschen Linien und Hybriden stärker absinken und diesen Unterschied bis zur Ernte nicht mehr kompensieren können. Hingegen bleiben die Werte der osteuropäischen Linien etwas stabiler,

sinken nicht ganz so weit ab und sind am Ende der Vegetationsperiode auf einem deutlich höheren Level als die deutschen Sorten. Auch hier liegen die Hybridsorten wieder zwischen den deutschen und den osteuropäischen Linien. Es ist zu vermuten, dass die REIP-Werte der osteuropäischen Sorten am Ende der Vegetationsperiode höher sind als die der deutschen Sorten, da letztere einen geringeren N-Gehalt aufweisen. Die entsprechenden Untersuchungen sind zum Zeitpunkt der Publikation noch nicht abgeschlossen. Daher kann diese Vermutung noch nicht bestätigt werden.

Da die deutschen Sorten für ein gemäßigtes Klima gezüchtet wurden, sind sie auch an eine längere Vegetationsperiode angepasst. Das heißt, dass v.a. der Zeitraum zwischen Blüte und Reife länger ist. Dies spiegelt sich in den Vegetationsindizes wider. Besonders am WI und NDVI kann man erkennen, dass die deutschen Sorten, sowohl Linien als auch Hybride, länger grün bleiben. Dadurch sinkt der NDVI später ab. Durch die Züchtung auf einen späteren Erntezeitpunkt, eine spätere Abreife und damit einen längeren Wachstumszeitraum, bleibt auch der Wassergehalt der deutschen Sorten länger auf einem höheren Niveau. Die Hybridsorten scheinen besser mit den Stressbedingungen des moldawischen Klimas zurecht zu kommen. Ihre Indizes verhalten sich ähnlicher wie die der osteuropäischen Sorten. Durch den Heterosiseffekt der Hybridzüchtung sind diese Sorten robuster und weniger empfindlich gegenüber abiotischen Stressoren.

Zwischen 29.6.2017 und 5.7.2017 kann man eine rasche Abnahme des NDVI der deutschen Sorten, v.a. Linien, erkennen, während die Werte der osteuropäischen Sorten nahezu gleichbleiben. Der Unterschied zwischen den beiden Gruppen wird geringer. Das ist ein Hinweis auf die sog. Notreife in den deutschen Weizensorten. Dabei trocknet das Getreide ab, die Körner bleiben klein und leicht, da vor der Abreife nicht mehr ausreichend Stärke eingelagert werden kann (vgl. Abb. 2). Dadurch kommt es zu Ertragseinbußen. Ursache für Notreife kann eine Unterbrechung der Nährstoffzufuhr sein, wie sie bei Trockenheit eintreten kann.

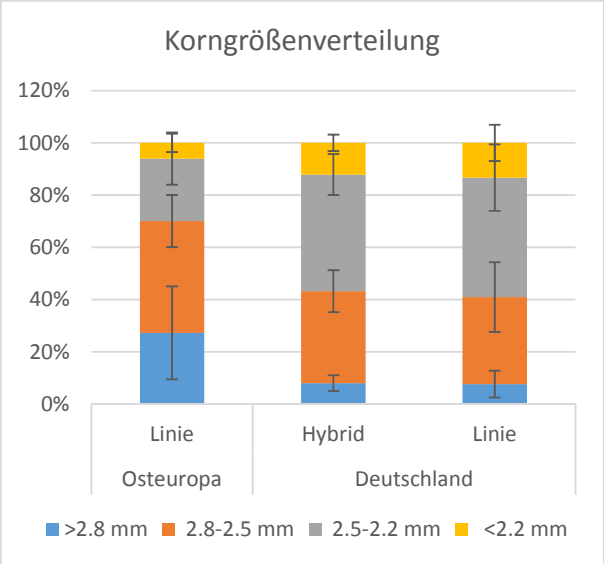


Abb. 2: Anteile der einzelnen Korngrößen am Gesamtertrag (%)

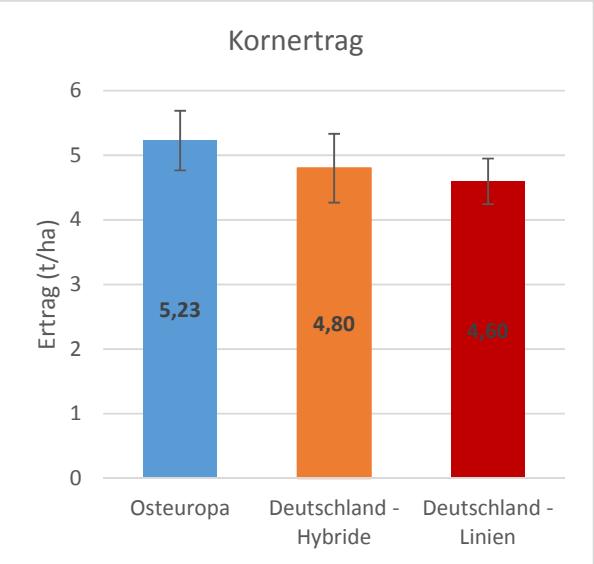


Abb. 3: Kornertrag (t/ha) der getesteten Weizensorten

Die Ertragseinbußen sind in Abbildung 3 deutlich erkennbar. Die osteuropäischen Sorten zeigen normale Ertragslevel, während die deutschen Sorten deutlich weniger Kornertrag bringen, verglichen mit einem Anbau im gemäßigten Klima Süddeutschlands. Dort sind je nach Sorte bis zu 9 t/ha möglich.
Korrelationen zwischen den Vegetationsindizes und dem Kornertrag zeigten sich vor allem zu Beginn und ganz am Ende der Vegetationsperiode (Abb. 4).

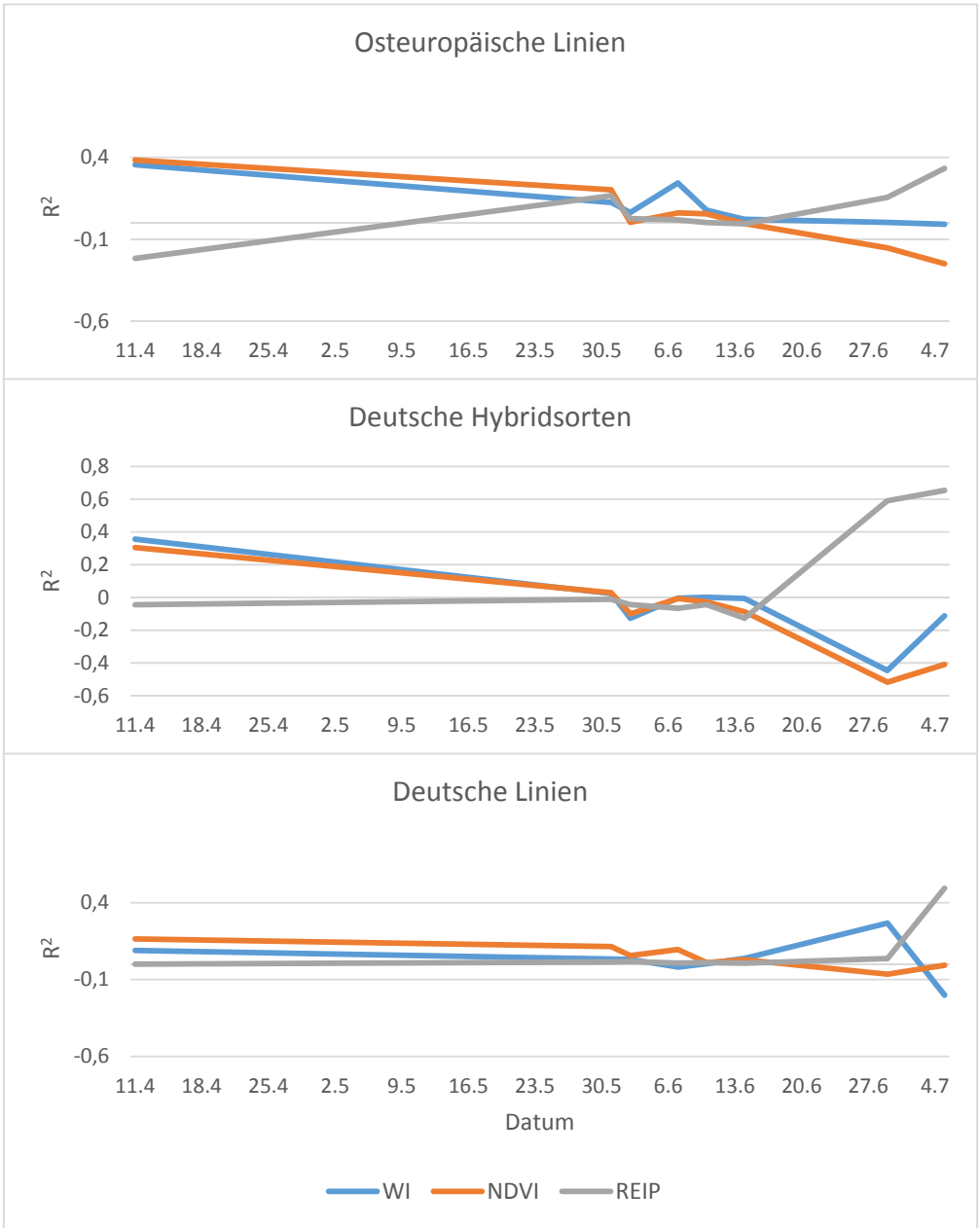


Abb. 4: Verlauf der Korrelation zwischen Ertrag (t/ha) und ausgewählten Spektralindizes nach Herkunft der Züchtungen

Auch bei den Korrelationen zeigt sich wieder bei den osteuropäischen Sorten und den deutschen Hybriden ein ähnlicher Verlauf. Zu Beginn zeigen der WI und NDVI einen Zusammenhang mit dem Ertrag, der dann während der Vegetationsperiode immer geringer wird. Nach dem Stadium der Milchreife (ZS 79, Messung vom 20.6.2017) werden die Werte negativ. Je höher der nun ermittelter NDVI- bzw. WI-Wert ist, mit desto geringeren Kornerträgen muss gerechnet werden. Die NDVI- und WI-Werte sind hoch, so lange die Pflanze noch grün und vital ist. Wenn dies zu einem späten Zeitpunkt noch der Fall ist, wird der Effekt der Notreife stärker und die Erträge sinken umso stärker ab. Der REIP verhält sich umgekehrt, ist gegen Ende der Vegetationsperiode deutlich stärker korreliert als zu Beginn. Dies bekräftigt den Zusammenhang mit der Biomasseentwicklung. Allerdings ist eine solche Entwicklung des REIP für eine frühzeitige Vorhersage der zu erwartenden Kornerträge nicht hilfreich. Die deutschen Linien zeigen nur kurz vor der Getreideernte nennenswerte Korrelationen zwischen den Vegetationsindizes und dem Ertrag. Es ist also davon auszugehen, dass die Vorhersage des Ertrages schwieriger und ungenauer wird, je stärker der abiotische Stress auf die Pflanze einwirkt.
Mit Hilfe der Thermalkameras sollte untersucht werden, ob sich die Temperaturen der Weizensorten je nach Herkunft unterscheidet und ob die Hybridsorten auch hier einen physiologischen Vorteil aufzeigen. Abbildung 5 zeigt die gemessenen Temperaturen der zwei Thermalkameras im zeitlichen Verlauf.

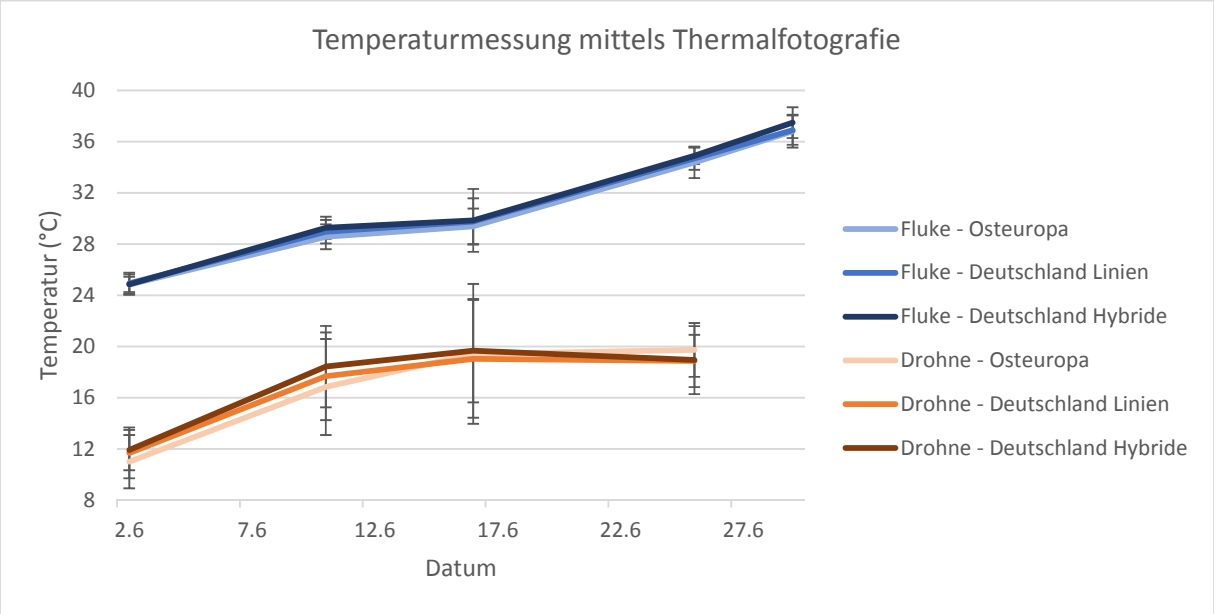


Abb. 5: Gemessene Temperatur der Weizensorten mit Fluke Ti400 bzw. Drohne "albris"

Wie in Abbildung 5 ersichtlich ist, sind die gemessenen Temperaturen zwischen den beiden Thermalkameras sehr unterschiedlich. Die Differenzen zwischen Fluke und Drohne lagen für die einzelnen Parzellen zwischen 1-20 °C. Die handgehaltene Thermalkamera Fluke hat eine um ein Vielfaches höhere Auflösung und kann dadurch viel genauer differenzieren.

Während des Wachstums, solange die Pflanze noch grün ist, dient Wasser als Kühlmittel, um die Temperatur der Pflanze zu regulieren. Zudem werden die Spaltöffnungen geöffnet, damit ein ausreichender Gasaustausch zusätzlich vor zu starkem Temperaturanstieg schützt. Im Falle anhaltender Trockenheit schließt die Pflanze die Spaltöffnungen der Blätter, um Wasserverlust durch Transpiration zu verringern (TAIZ et al. 2015). Bei einer Kombination aus Trockenheit und Hitze, wie sie während des Feldversuches auftrat, kommen die Pflanzen in physiologischen Stress, da es an Wasser mangelt, um die Temperatur zu regulieren, die Stomata aber geschlossen bleiben, um zusätzlichen Wasserverlust zu verhindern. Weizensorten, die an ein trockenes und heißes Klima angepasst sind, kommen unter solchen Bedingungen nicht so schnell in physiologischen Stress, als Sorten, die an kühlere Standorte gewöhnt sind. Demnach war zu erwarten, dass die osteuropäischen Weizensorten eine geringere Oberflächentemperatur zeigen, als die deutschen Sorten. In Abbildung 4 ist eine entsprechende Tendenz erkennbar, sowohl bei Messungen mit der Fluke als auch mit der Drohne. Statistisch signifikant sind die Unterschiede jedoch nicht. Der deutliche Anstieg der Temperatur, die mit der Fluke-Kamera gemessen wurde, ab 16. Juni (ZS 73-75, Milchreife) geht mit der beginnenden Seneszenz einher. Die Pflanzen beginnen abzusterben und vertrocknen. Dadurch ist immer weniger Wasser zur Temperaturregulation vorhanden und die Pflanzentemperatur steigt an.
Die Temperaturwerte, welche mit der Drohne albris erfasst wurden, zeigen diesen Anstieg nicht. Sie liegen über den gesamten Zeitraum der Messungen deutlich unter jenen der Fluke-Messungen.
Parallel stattfindende Untersuchungen zeigten die Schwierigkeiten auf, die sich aus der geringen Auflösung der Thermalkamera der Drohne ergeben (BARMEIER & MÖNICH 2017). Dabei wurden Versuchsobjekte mit 1 bzw. 0.5 m^2 Grundfläche aus verschiedenen Entfernungen thermografiert und die Überblendung der Aufnahmen verglichen (Abb. 6). Als Überblendung wird die Wahrnehmung der Wärmestrahlung über das Objekt hinaus bezeichnet.

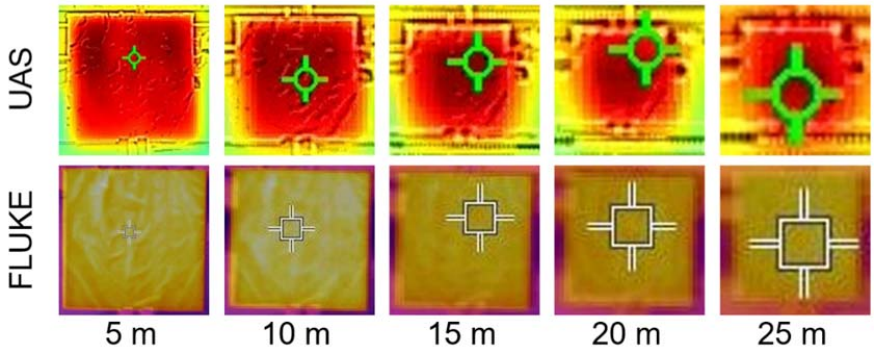


Abb. 6: Thermalaufnahmen einer Fläche von 1 m^2 nach verwendeter Kamera und Entfernung

Hierbei wurde deutlich, dass bei den Aufnahmen mit der Fluke auch in 25 m Entfernung noch klare Kanten und die Form eines Quadrates erkennbar waren. Bei der Drohne hingegen war die Form des Quadrates ab einer Entfernung von 20 m nur noch schwer erkennbar, die Wärmequelle scheint die Form eines Kreises anzunehmen. Die Überblendung war bei der Drohne deutlich stärker ausgeprägt als bei den Aufnahmen mit der Fluke.

Eine klare Abgrenzung des fotografierten Bereiches ist auch in der Auswertung der Weizenparzellen schwierig. Die Flughöhe der Drohne betrug 17 m. In Abbildung 6 wird erkennbar, dass auch bei dieser Entfernung die Überblendung deutlich negativen Einfluss auf die Qualität der Aufnahme hat. An den Aufnahmen mit der Fluke sieht man, dass eine höhere Auflösung hier Abhilfe schaffen kann. Dadurch wird die Auswertung der einzelnen Temperaturbereiche vereinfacht und die Abgrenzung der Parzellen zueinander eindeutiger, was Fehler und Abweichungen verringert.
Bei den Thermalfotografien konnte keine starke Korrelation zum Ertrag festgestellt werden. Dies deutet darauf hin, dass die Temperatur keinen konstant negativen Einfluss auf den Kornertrag hat.

4 Fazit & Ausblick

Verschiedene Ansätze der Hochdurchsatz-Phänotypisierung bieten die Möglichkeit, wichtige pflanzenphysiologische Parameter nicht-destruktiv und schnell zu erfassen. Dies spielt eine wichtige Rolle in der Beurteilung des Einflusses abiotischer Stressfaktoren, wie hohe Temperaturen, auf Pflanzen. Dabei stehen vor allem der geringere Arbeitsaufwand und der flexible Einsatz im Vordergrund. Im durchgeführten Feldversuch konnten Unterschiede zwischen den Weizensorten je nach Herkunft der Züchtung festgestellt werden. Eine Vorhersage des Kornertrages war aber schwierig und unpräzise. Es waren deutliche Ertragseinbuße bei den deutschen Weizensorten vorhanden, jedoch waren diese nur zu einem geringen Grad mit den Spektral- bzw. Thermalmessungen korreliert.
Für die Erfassung von Spektralreflexionen und Oberflächentemperatur können entsprechende Sensoren auf UAVs angebracht werden, so dass größere Flächen in sehr kurzer Zeit beurteilt werden können. Spektralindizes können ohne großen Aufwand errechnet und ausgewertet werden und werden häufig zur Vorhersage des Kornertrages oder auch Nährstoffbedarfes herangezogen. Bei der Thermalfotografie ist zu beachten, dass die Kameras eine ausreichende Auflösung haben, damit der Grad der Überblendung möglichst gering ist und die zu untersuchenden Flächen scharf abgegrenzt abgebildet werden können. Allerdings ist bei diesen bildgebenden Verfahren eine zeitaufwändige Nachbereitung und die Verfügbarkeit und Kenntnis entsprechender Software notwendig. Wenn jedoch die technischen Voraussetzungen geschaffen sind können Drohnen bei der Beurteilung des Zustandes von Getreidepflanzen zielgerichtet eingesetzt werden.

5 Literaturverzeichnis

BARMEIER, G. & MÖNICH, T., 2017: Können Drohnen traditionelle Boniturmethoden im Feldversuchswesen ersetzen? Lehrstuhl für Pflanzenernährung, Technische Universität München.

FAO, 2017: The future of food and agriculture – Trends and Challenges. Rome.

GUYOT, G., BARET, F. & MAJOR D.J., 1988: High Spectral Resolution: Determination of Spectral Shifts Between the Red and Near Infrared. International Archives of Photogrammetry and Remote Sensing, **11**, 750-760

PEÑUELAS, J., PIÑOL, J., OGAYA, R. & FILELLA, I., 1997: Estimation of plant water concentration by the reflectance Water Index WI (R900/R970). International Journal of Remote Sensing **18**(13) 2869-2875.

ROUSE, J.W., HAAS, J.R.H., SCHELL, J.A. & DEERING, D.W., 1974: Monitoring vegetation systems in the Great Plains with ERTS. Third ERTS-1 Symposium NASA, Washington, DC, 309-317.

SHIFERAW, B., SMALE, M., BRAUN, H.J., DUVEILLER, E., REYNOLDS, M. & MURICHO, G., 2013: Crops that feed the world 10. Past success and future challenges to the role played by wheat in global food security. Food Security, **5**, 291-31.

TAIZ, L., ZEIGER, E., MOLLER, I.M. & MURPHY, A., 2015: Plant Physiology and Development. Sinauer Associates, Incorporated.

ZADOKS, J.C., CHANG, T.T. & KONZAK, C.F., 1974: A decimal code for the growth stages of cereals. Weed Research, **14**(6), 415-421.

Bestimmung der Erdoberflächenbewegung mittels Radarinterferometrie nahe der Agua Blanca Störung, Baja California

ANIKA RIEDEL[1], BJÖRN RIEDEL[1], CHRISTINE SCHOTTMÜLLER[1], WOLFGANG NIEMEIER[1] & MARKUS GERKE[1]

Zusammenfassung: Das BMBF-geförderte ECOAQUA Projekt untersucht die Folgen der bio-ökonomischen Ausbeutung von Grundwassersystemen in ariden Küstengebieten und analysiert die beteiligten Prozesse und ihre Wechselwirkungen. Das Untersuchungsgebiet umfasst die Stadt Ensenada, Baja California, Mexiko mit ihren umliegenden Tälern. Die extensive Grundwasserentnahme führt zu Salzwasserintrusionen an der Pazifikküste und damit zur Verschlechterung der Wasserqualität im Grundwasserleiter.

Dieser Beitrag stellt die aus radarinterferometrischen Beobachtungen abgeleiteten Oberflächen-bewegungen nahe der Agua Blanca Störung, südlich von Ensenada vor. Die Bewegungsraten dienen als ein Parametersatz für ein Vorhersagemodell, das zukünftige Veränderungen des Küstenökosystems beschreibt und mögliche Risiken abschätzt.

Das Untersuchungsgebiet hat eine Größe von 670 km² und umfasst Bergketten mit Höhen bis 1000 m, kleine Ortschaften und landwirtschaftlich genutzte Täler. Zur Auswertung wurden Radaraufnahmen aus Envisat im auf- und absteigenden Bahnbogen ausgewertet. Der aufsteigende Bahnbogen umfasst die Jahre 2005 bis 2010 mit 20 Aufnahmen und der absteigende Bahnbogen liefert 45 Bilder der Jahre 2003 bis 2010. Des Weiteren wird die Zeitreihe mit Sentinel-1 Daten aus drei Jahren ergänzt. Die Daten wurden mit der Persistent Scatterer Interferometry (PSI) und Small Baseline Subset (SBAS) Methode bearbeitet.

Die PSI Auswertung zeigt keine zufriedenstellenden Ergebnisse, weil die Methode in den ländlichen Gebieten zu wenige permanente Rückstreuer über die Zeit nutzen kann. Im Gegensatz zur PSI Methode zeigen die SBAS Auswertungen auf den pliozänen Konglomeraten im Maneadero Tal und am San Tomas Bach Bewegungen, die vermutlich durch Erosions- und Ablagerungsprozesse aus Starkregenereignisse in den Wintermonaten verursacht wurden. Im Maneadero Tal sind punktuelle Bewegungen, vermutlich durch den Bau von Regenrückhaltebecken und Gewächshäusern, erkennbar.

Die Bewegungen der Agua Blanca Störung werden mit 4,5 ±1,4 mm/a für den aufsteigenden Bahnbogen und 4,1 ±1,3 mm/a für den absteigenden Bahnbogen in Schrägentfernungsgeometrie berechnet. ROCKWELL et al. (1993) bestimmt für die letzten 28.000 a Bewegungen von 4,1 mm/a an der Störung. DIXON et al. (2002) bestimmt aus relativen GPS Messungen eine Differenz über die Agua Blanca Störung von 2,5 mm/a und aus Modellrechnungen 6,2 mm/a. Die hier ermittelten Bewegungsraten stimmen in der Größenordnung mit den Literaturwerten überein.

1 Einleitung

Das Projekt ECOAQUA setzt sich mit den Folgen der bio-ökonomischen Ausbeutung von Grundwassersystemen in ariden Küstengebieten auseinander, um die Prozesse und ihre Wechselwirkungen besser verstehen zu können, ein Modell zu entwickeln um zukünftige Veränderun-

1 Institut für Geodäsie und Photogrammetrie, TU Braunschweig, Bienroder Weg 81, D-38106 Braunschweig, E-Mail: [a.riedel, b.riedel, c.schottmueller, w.niemeier, m.gerke]@tu-braunschweig.de

gen des Küstenökosystems vorherzusagen und mögliche Risiken besser abschätzen zu können (NIEMEIER et al. 2015). Dieser Beitrag liefert einen Überblick über die Georisiken in dem Gebiet. Das zu untersuchende Gebiet in Mexiko umfasst Bergketten, kleine Ortschaften und landwirtschaftlich genutzte Täler. In den Tälern besteht auf Grund der landwirtschaftlichen Nutzung und der wachsende Industrie- und Tourismusbrache ein erhöhter Wasserbedarf. Das Wasser wird aus grundwassergespeisten Brunnen bezogen und führt zur Ausbeutung der Aquifere (NIEMEIER et al. 2015). Die Übernutzung des Grundwassers scheint zu Salzwasserintrusionen am Pazifik und zur Verschlechterung der Wasserqualität im Grundwasserleiter zu führen (DAESSLÈ et al. 2014).
Das Untersuchungsgebiet, siehe Abb. 1, befindet sich ca. 100 km südlich der amerikanisch- mexikanischen Grenze im Norden der Halbinsel Baja California, Mexiko. Im Westen grenzt das Untersuchungsgebiet an den Pazifik und hat eine Fläche von 670 km² mit 23 km Ausdehnung von Ost nach West und 28 km Ausdehnung von Nord nach Süd. Im Nordwesten befindet sich die Stadt Ensenada mit rund 300.000 Einwohnern (IMIP & CODEEN 2011), südlich der Stadt schließt sich das Maneadero Tal an. Das Tal hat eine Fläche von ca. 80 km² und grenzt im Süden an die NW nach SE streichende Punta Banda Bergkette mit Höhen bis 1000 m. Die Bergkette hat eine Länge von ca. 30 km und durchtrennt fast das gesamte Untersuchungsgebiet. Die Berge im Osten des Untersuchungsgebiets mit Höhen zwischen 400 und 700 m gehören zur Peninsular Bergkette, die sich von Südkalifornien bis in den Süden Baja Californias zieht (HIRABAYASHI et al. 1996). Im Südosten des Untersuchungsgebietes liegt das San Tomas Tal, das durch den Santo Tomas Bach entwässert wird. Im Norden Baja Californias herrscht ein gemäßigtes, sommertrockenes, winterfeuchtes Subtropen- Klima.
In den Tälern von Maneadero und San Tomas, sowie in den Fluss- und Bachtälern befinden sich quartäre, fluviatile, schlecht verfestigte Kiese, Sande und Silte (Qhoal). An der Küste finden sich überwiegend schmale Streifen quartärer äolischer Ablagerungen (Qhoc), die Lagune im Maneadero Tal besteht aus lakustrinen Ablagerungen (Qhopa) (SGM 2003).
Die Halbinsel Baja California besteht im Osten aus Gesteinen des nordamerikanischen Basements, der Westen wurde aus magmatischen Bogengesteinen gebildet. Die Magmatite sind durch die Agua Blanca Störung unterteilt. Nördlich dieser Störung befindet sich die Santiago Peak Formation, südlich die Alisitos Formation. Die Agua Blanca Störung entstand in der Kreide, als der Alisitos Inselbogen mit dem magmatischen Bogen der Santiago Peak Arc Formation am nordamerikanischen Kontinentalrand kollidierte (SCHMIDT et al. 2002).
Im Miozän entstand aus der Transformstörung zwischen der pazifischen und der nordamerikanischen Platte die San Andreas Störung, die die Agua Blanca Störung reaktivierte (HENRY & ARANDA-GOMEZ 1998). Die heute dextrale Agua Blanca Störung ist die südlichste Störung des San Andreas Störungssystems (SPRINGER 2010).
DIXON et al. (2002) bestimmen die Bewegungsraten der Agua Blanca Störung aus GPS- Daten von 1993 bis 1998 und ermittelt daraus Geschwindigkeiten von 2,5 mm/a, viskoelastische Modellberechnungen führen zu Geschwindigkeiten von 6,2 mm/a für die Agua Blanca Störung.

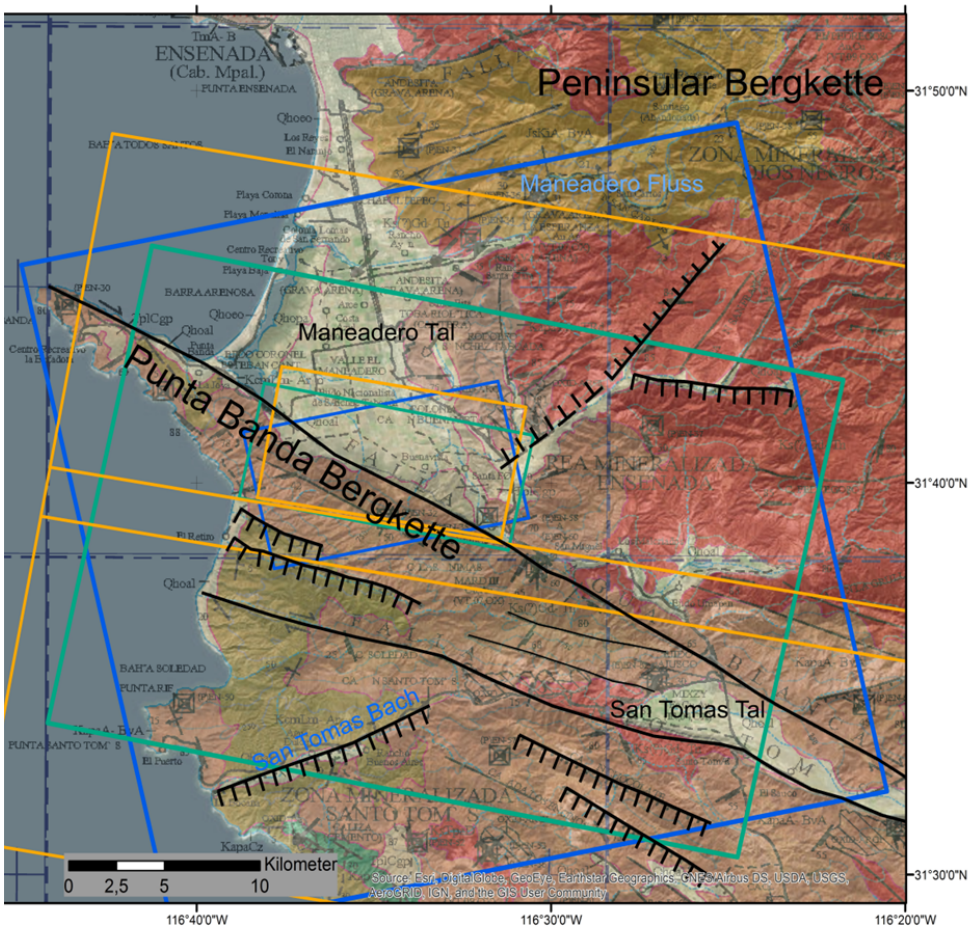


Abb. 1: Übersicht über das Untersuchungsgebiet mit der geologischen Karte (SGM 2003) unterlegt. Envisat: Gebiete des aufsteigenden Bahnbogens (blau), absteigend (grün), Sentinel-1a absteigend (gelb).

ROCKWELL et al. (1993) bestimmen anhand von linearen Regressionsgeraden aus Versatz und Alter für die Agua Blanca Störung an der Punta Banda Bergkette Bewegungsraten von 4,1 mm/a für die letzten 28 ka. Für die letzten 840 ka wird eine Geschwindigkeit von 2,3 mm/a berechnet, wobei vermutet wird, dass sich die Bewegungsrate im Quartär beschleunigt hat.
Im Norden Baja Californias beschränkt sich die Seismizität auf die obersten 20 km der Kruste. Die stärkste Aktivität wird nahe der San Miguel Vallecitos Störung, nordöstlich des Untersuchungsgebiets, in den obersten 15 km beobachtet. Die Agua Blanca Störung ist zurzeit seismisch still, DIXON et al. (2002) vermuten, dass sie nicht inaktiv ist, sondern sich in der interseismischen Phase befindet. Der Zeitpunkt des letzten Erdbebens ist nicht bekannt.
Die in diesem Beitrag ermittelten Bewegungsraten werden als ein Parametersatz für ein Vorhersagemodell, das zukünftige Veränderungen des Küstenökosystems beschreibt, verwendet um mögliche Risiken abzuschätzen.

2 Methodik

Die Oberflächenbewegungen im Untersuchungsgebiet werden aus radarinterferometrischen Beobachtungen abgeleiteten.
Die Radarfernerkundung ist ein aktives Messverfahren, bei dem ein Satellit elektromagnetische Wellen auf die Erde sendet und die rückgestreute Welle empfängt, dabei werden Amplitude und Phase aufgezeichnet. Beim interferometrischen SAR (InSAR) werden die Phasenunterschiede zweier getrennt voneinander aufgenommenen Datensätze bzw. angeordneter Antennen ausgewertet. Dieser Phasenunterschied kann i) zur Objekthöhenberechnung (Ableitung von Geländemodellen) oder ii) zur Bestimmung der geometrischen Veränderungen (Setzung, Verschiebungen) der Erdoberfläche genutzt werden (ALBERTZ 2007).
Multitemporale Auswertemethoden ermöglichen die Erstellung von Zeitreihen über die Erdoberflächenbewegungen im Untersuchungsgebiet. Es wurde die Persistent Scatterer Interferometrie (PSI) Methode nach FERRETTI et al. 2001 und die Small Baseline Subset (SBAS) Methode nach BERARDINO et al. 2002 durchgeführt. Die PSI Methode analysiert die Verschiebung von stabilen Streuern (Scatterer) über den Auswertezeitraum, die Auswahl dieser, kann auf Grund von Kohärenzgrenzwerten, Amplitudenwerten oder Phaseneigenschaften erfolgen. Aus den einzelnen Interferogrammen wird die Verschiebung der Streuer über die gesamte Zeit ermittelt. Diese Methode ist besonders geeignet für urbane Gebiete (MOTAGH 2016). Die SBAS Methode arbeitet mit verteilten Streuen (distributed scatterer) und ist deshalb besonders geeignet für ländliche Gebiete und gemischt bebaute Flächen. Die Methode ist deshalb in vielen landwirtschaftlich geprägten Gebieten besser geeignet als die PSI Methode (GONG et al. 2016).
Aus dem ESA Archiv standen 65 Aufnahmen von Envisat in auf- und absteigenden Bahnbögen der Jahre 2003 bis 2011 zur Verfügung. Der aufsteigende Bahnbogen umfasst die Jahre 2005 bis 2011 mit 20 Aufnahmen, der absteigende Bahnbogen liefert 45 Bilder der Jahre 2003 bis 2011. Zur Erweiterung der Zeitreihe wurden außerdem 55 Bilder des absteigenden Bahnbogens des Sentinel-1a verwendet, der seit 2014 operiert. Das Untersuchungsgebiet umfasst zwei Bursts, die gemeinsam ausgewertet wurden.
Die Daten werden mit der PSI und SBAS Methode bearbeitet. In einem ersten Schritt wurde das gesamte Untersuchungsgebiet mit einer Ausdehnung von 23 km x 28 km prozessiert, um einen Überblick über mögliche Bewegungen zu erhalten. Zudem wurde ein Detailausschnitt (12 x 6 km) über die Agua Blanca Störung bearbeitet um relative Bewegungsraten zu bestimmen und diese mit Literaturwerten zu vergleichen.
In der Abb. 2 sind die räumlichen und zeitlichen Verbindungen zwischen den einzelnen Radaraufnahmen in Form sogenannter Netzwerkgrafiken dargestellt. Die PSI Auswertungen sind links dargestellt und bilden nur Interferogramme zwischen einem Masterdatensatz und allen anderen Aufnahmen (slaves). Die SBAS Auswertung verwendet mehrere Masterdatensätze auf Basis einer maximalen zeitlichen und räumlichen Differenz. Die jeweiligen Werte sind der Tabelle 1 zu entnehmen.

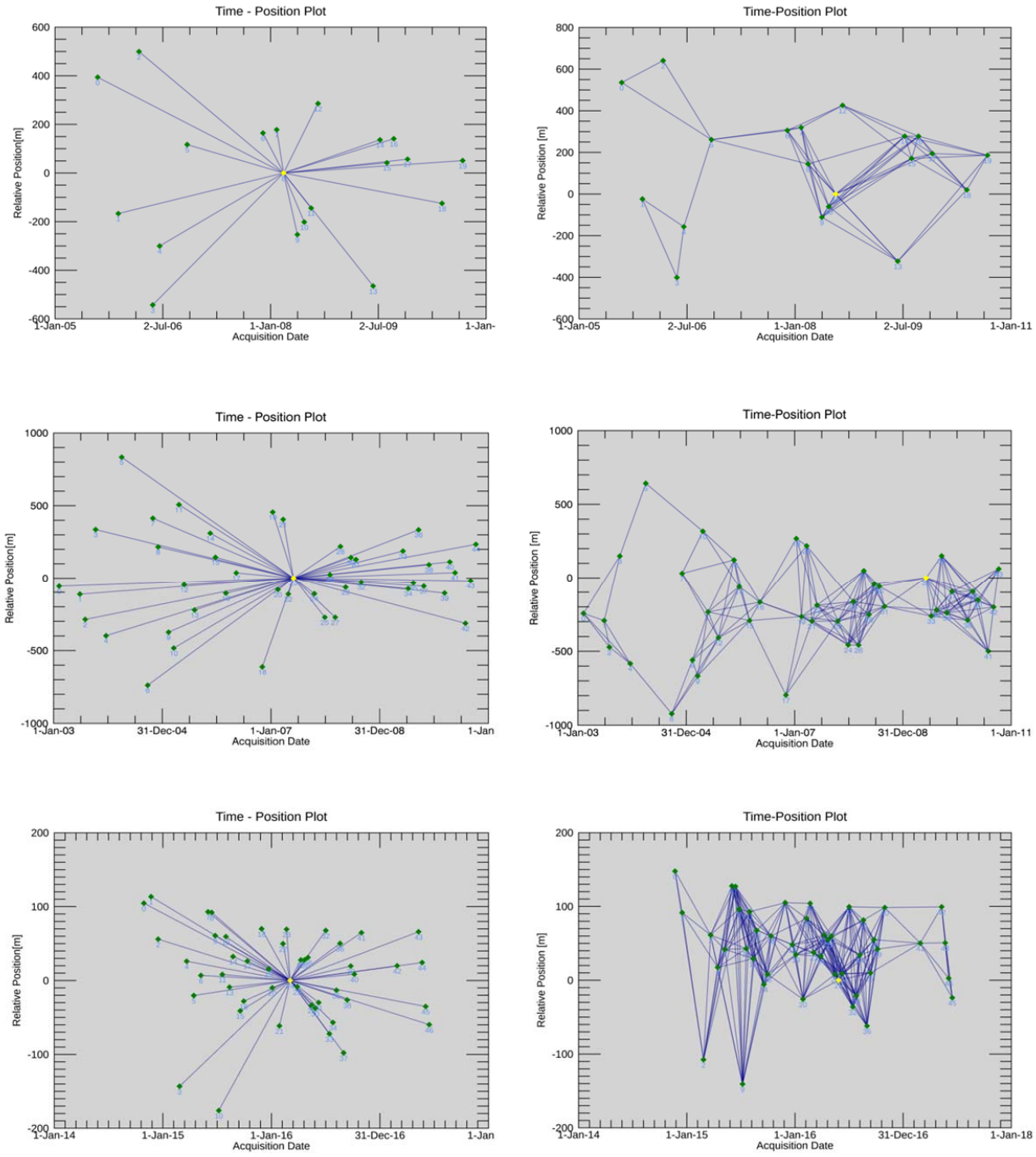


Abb. 2: Räumlichen und zeitlichen Verbindungen zwischen Master und Slaves für PSI Auswertung (links) und SBAS Auswertung (rechts). Netzwerkgrafiken: Obere Zeile Envisat aufsteigender Bahnbogen, mittlere Zeile Envisat absteigender Bahnbogen, unten Sentinel-1

Tab. 1: Übersicht über die verwendeten Parameter in der SBAS Auswertung

Satellit	Bahnbogen	Zeitintervall	Anzahl der Bilder	Gebiet	Max. kritische Basislinie (%)	Max. temporale Basislinie (d)	Anzahl Interferogramme
Envisat	Aufsteigend	08.2005- 09.2010	20	gesamt	60	500	63
Envisat	Aufsteigend	08.2005- 09.2010	20	Ausschnitt	50	400	37
Envisat	Absteigend	02.2003- 10.2010	45	gesamt	60	450	214
Envisat	Absteigend	02.2003- 10.2010	45	Ausschnitt	50	365	153
Sentinel-1a	Absteigend	11.2014-10.2017	55	gesamt	45	120	342
Sentinel-1a	Absteigend	11.2014-10.2017	53	Ausschnitt	45	60	170

3 Ergebnisse

Das Ergebnis der PSI Auswertung des aufsteigenden Bahnbogens aus Envisat, siehe Abb. 3 A, liefert im gesamten Untersuchungsgebiet keine verlässlichen Geschwindigkeiten. Der absteigende Bahnbogen aus Envisat zeigt in der PSI Auswertung, siehe Abb. 3 B, ein sehr homogenes Ergebnis, das keine signifikanten Bewegungsraten liefert.
Die Ergebnisse der PSI Auswertung aus Sentinel, siehe Abb. 3 C, lassen sich ebenfalls kaum vom Rauschen unterscheiden. Im Maneadero- und im San Tomas Tal sind die Geschwindigkeiten negativ in den höheren Regionen leicht positiv.

3.1 SBAS Ergebnisse im gesamten Untersuchungsgebiet

Die SBAS Auswertung des aufsteigenden Bahnbogens aus Envisat im Untersuchungsgebiet, siehe Abb. 3 D, zeigt Geschwindigkeitswerte zwischen -7 bis +6 mm/a. Im Maneadero Tal sind die Geschwindigkeiten überwiegend negativ, im Westen und Südwesten stark negativ mit Geschwindigkeiten zwischen -7 bis -2 mm/a. Im Osten des Maneadero Tals, auf Höhen um 300 m, steigen die Geschwindigkeiten auf bis +6 mm/a an. Höher gelegene Bereiche weiter östlich zeigen Geschwindigkeiten zwischen 0 bis +3 mm/a. Auf der Punta Banda Gebirgskette sind die Geschwindigkeiten mit -5 bis -2 mm/a stark negativ und auch Südlicher, am Pazifik, sind die Geschwindigkeiten zwischen -5 bis +2 mm/a im Durchschnitt, leicht negativ. Im Süden des Untersuchungsgebiets, entlang des San Tomas Baches, sind die Geschwindigkeiten meist positiv mit Geschwindigkeiten bis +8 mm/a. Nördlich und südlich des San Tomas Tals sind die Geschwindigkeiten zwischen -5 bis -2 mm/a.
Die ermittelten Geschwindigkeiten des absteigenden Bahnbogens aus Envisat, siehe Abb. 3 E, zeigen Werte zwischen -6 bis +7 mm/a im gesamten Untersuchungsgebiet. Im Maneadero Tal überwiegen die positiven Geschwindigkeiten zwischen 0 bis +3 mm/a, im Südwesten und Südosten des Tals sind negative Geschwindigkeiten zwischen -5 bis 0 mm/a zu beobachten. In der Peninsular Bergkette im Osten überwiegen die Geschwindigkeiten zwischen -2 bis 0 mm/a. Südlich der Punta Banda Bergkette, an der Pazifikküste, sind die Geschwindigkeiten zwischen -2 bis 0 mm/a negativ. Nahe des San Tomas Baches haben die Geschwindigkeiten Werten von 0 bis +7 mm/a und sind relativ hoch.

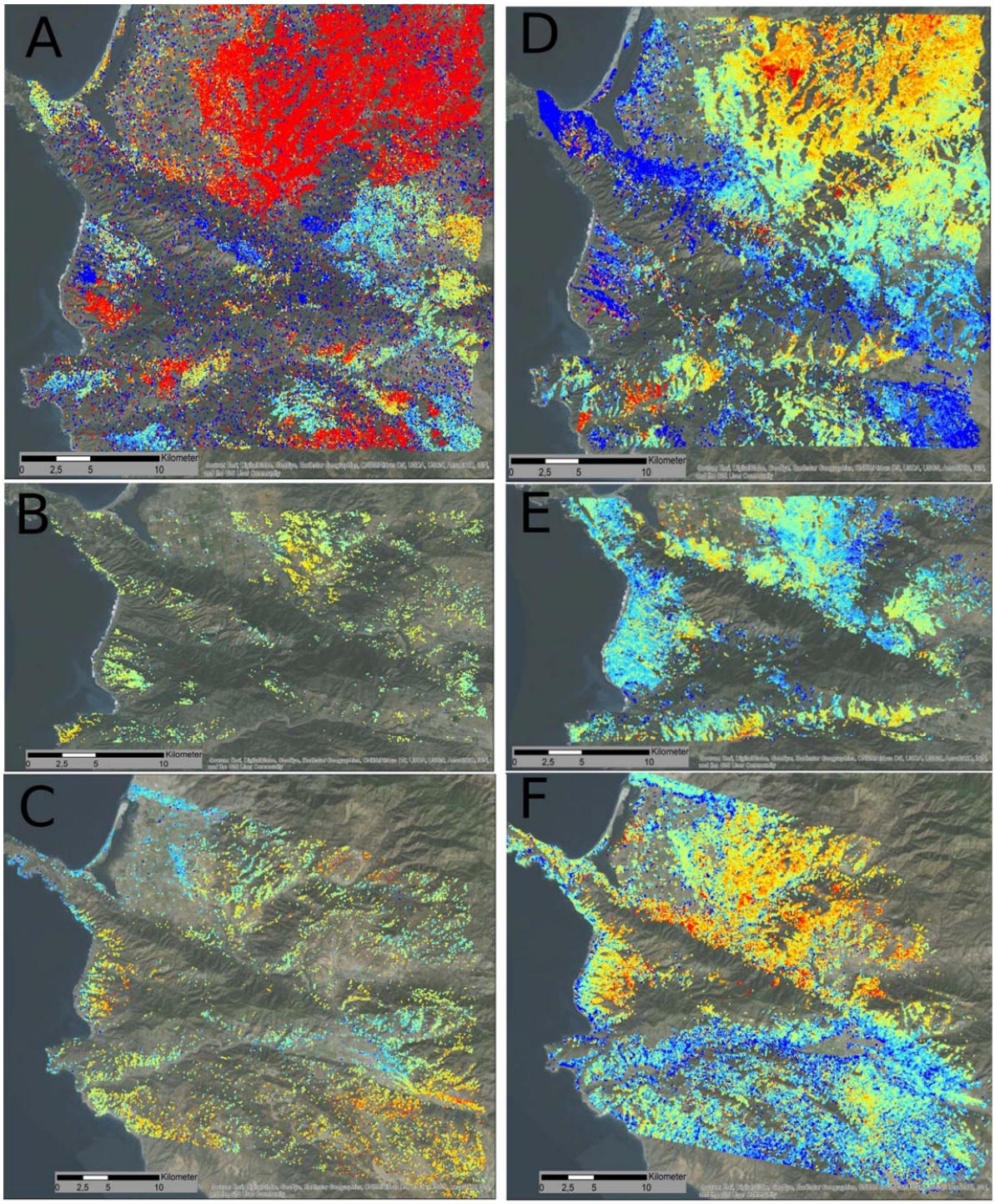


Abb. 3: Ergebnisse der PSI und SBAS Auswertungen des Untersuchungsgebiets. A: PSI Envisat aufsteigend, B: PSI Envisat absteigend, C: PSI Sentinel-1a, D: SBAS Envisat aufsteigend, E: SBAS Envisat absteigend, F: SBAS Sentinel-1a.

Insgesamt weist dieser Datensatz ein sehr homogenes Bild auf mit stark negativen Geschwindigkeiten an den südlichen Berghängen der Punta Banda und den westlichen Hängen der Peninsular Bergkette.
Die Ergebnisse für die SBAS Auswertung der Sentinel-1 Zeitreihe, siehe Abb. 3 E, weist Geschwindigkeiten zwischen +8 bis -9 mm/a auf. Im Maneadero Tal sind Geschwindigkeiten überwiegend zwischen -2 bis +2 mm/a festzustellen. In der Peninsular Bergkette steigen die Geschwindigkeiten zwischen +1 bis +8 mm/a an. Im landseitigen Teil der Lagune und an den Bergen der Punta Banda Bergkette sind positive Geschwindigkeiten zu beobachten. Südlich der Bergkette am Pazifik steigen die Geschwindigkeiten von -9 auf +8 mm/a an. Im Bereich des San Tomas Tals und Baches sind die Geschwindigkeiten negativ mit Werten zwischen 0 bis -8 mm/a.
Beim Vergleich der Envisat mit den Sentinel Ergebnissen muss darauf geachtet werden, dass sich die Aufnahmegeometrien der Satellitentypen unterscheiden. Envisat hat einen Einfallwinkel θ_{inc} von 17° bis 26° in den vorgestellten Daten. Bei Sentinel ist θ_{inc} zwischen 36° und 51°. Dadurch variiert auch der Line of Sight (LOS) Winkel. Die vorgestellten Ergebnisse sind jeweils in LOS dargestellt.

3.2 SBAS Ergebnisse für den Ausschnitt der Agua Blanca Störung

Die SBAS Ergebnisse des aufsteigenden Bahnbogens aus Envisat im kleinen Ausschnitt der Agua Blanca Störung, siehe Abb. 4 oben, zeigt Geschwindigkeiten zwischen -10 bis +8 mm/a. An den Berghängen der Punta Banda Bergkette sind die Geschwindigkeiten mit -10 bis -4 mm/a am größten. Im Nordwesten des Ausschnitts ist ein fast runder Bereich mit 500 m Durchmesser und Geschwindigkeiten zwischen -10 bis -4 mm/a erkennbar. Ringsherum steigen die Geschwindigkeiten kontinuierlich auf +2 mm/a an. In der Mitte des Ausschnitts zeigt sich ein Bereich negativer Geschwindigkeiten bis -4 mm/a. Es scheint als wechselten sich von West nach Ost negative und positive Geschwindigkeiten ab.
Die Geschwindigkeiten der SBAS Auswertung der Jahre 2003 bis 2010 des absteigenden Bahnbogens aus Envisat, siehe Abb. 4 Mitte, zeigen im Ausschnitt der Agua Blanca Störung Werte zwischen -8 bis +6 mm/a. Im Nordwesten des Ausschnitts sind negative Geschwindigkeiten zwischen -8 bis -3 mm/a erkennbar, die eine runde Form mit ca. 400 m Durchmesser zeigen. Die Geschwindigkeiten steigen rundherum auf bis +2 mm/a an, im Osten schließen sich Geschwindigkeiten zwischen +1 bis +3 mm/a an. In der Mitte des Ausschnitts sind die Geschwindigkeiten zwischen -8 bis 0 mm/a und im Südosten steigen sie auf 0 bis +3 mm/a an. An den Hängen der Peninsular Bergkette sind die Geschwindigkeiten mit Werten zwischen -3 bis +1 mm/a eher negativ. Auf der Südseite der Punta Banda Bergkette sind vereinzelte Geschwindigkeiten zwischen -8 bis +3 mm/a berechnet worden.
Die Auswertung der Daten aus Sentinel-1 für den kleinen Ausschnitt der Agua Blanca Störung, siehe Abb. 4 unten, zeigt Probleme in der Auswertung. Im Südwesten und Nordosten scheinen topographische Effekte aufzutreten und sehr hohe Geschwindigkeiten zu erzeugen. Im Südosten ist eine deutliche Phasenrampe sichtbar. Im mittleren Bereich der Auswertung sind nördlich der Agua Blanca Störung meist positive Geschwindigkeiten zwischen +1 bis +7 mm/a erkennbar. Auf der Störung und südlicher überwiegen die negativen Geschwindigkeiten zwischen 0 bis -10 mm/a.

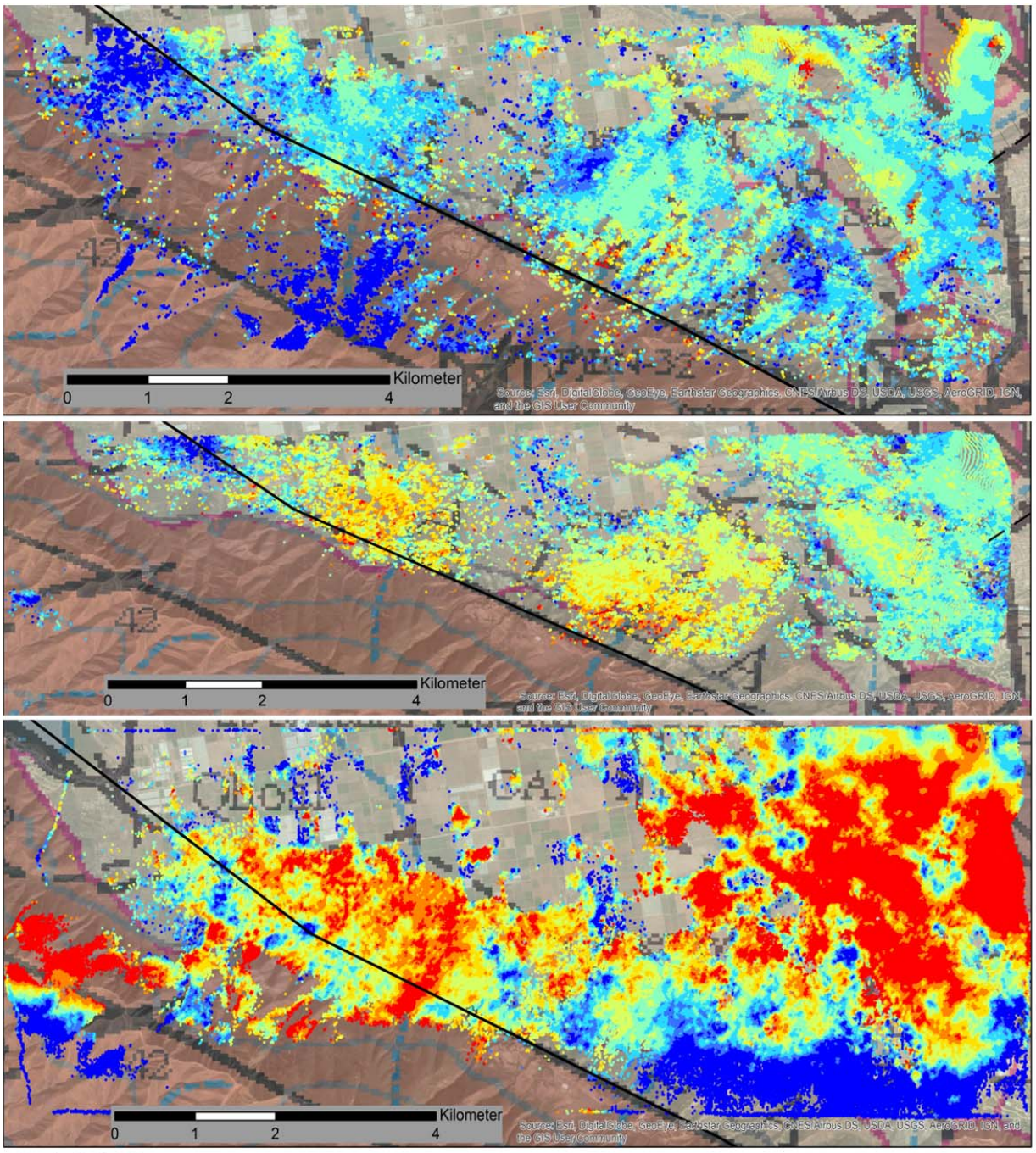


Abb. 4: Ergebnisse der SBAS Auswertungen mit der geologischen Karte (SGM 2003) unterlegt und der Agua Blanca Störung (schwarz) überlagert. Oben: Envisat aufsteigend, Mitte: Envisat absteigend, Unten: Sentinel-1a

4 Diskussion

Die Ergebnisse der PSI Auswertungen sind in den absteigenden Bahnbögen aus Envisat und Sentinel kaum vom Rauschen zu unterscheiden. Das Ergebnis des aufsteigenden Bahnbogens wirkt in den topographisch hohen Lagen sehr verrauscht und zeigt für die Peninsular Bergkette sehr hohe Geschwindigkeiten. Es lässt sich feststellen, dass die PSI Auswertungen im Untersuchungsgebiet keine belastbaren Ergebnisse liefern. Die PSI Methode scheint daher für das ländliche Untersuchungsgebiet nicht geeignet zu sein.
In Abb. 5 A sind die Ergebnisse der SBAS Auswertung des aufsteigenden Bahnbogens aus Envisat mit der geologischen Karte unterlegt. Westlich des San Tomas Tals sind die Geschwindigkeiten höher als südlich und nördlich davon. In der geologischen Karte (SGM 2003) ist ein Intrusionskörper (vermutlich Kreidezeitliche Granodiorite und Tonalite) sichtbar, der durch die San Tomas Störung versetzt wurde. Die San Tomas Störung könnte diese unterschiedlichen Geschwindigkeitsfelder herbeiführen.
Nordwestlich des San Tomas Tals befindet sich der San Tomas Graben, der in der Karte durch zwei parallele Verwerfungen gekennzeichnet ist. Die Geschwindigkeitsmuster scheinen sich an diesen zu orientieren. Südlich der Punta Banda Bergkette befindet sich ein negatives, verrauschtes Gebiet. Der Verlauf der San Tomas Störung ist hier sehr deutlich sichtbar und auch die nördlicheren Abschiebungen sind durch die Ergebnisse südlich erkennbar. Im Südwesten, am San Tomas Bach, befindet sich eine Abschiebung. Die hohen Geschwindigkeiten dort könnten durch Erosion und Sedimentablagerungen aus Starkregenereignissen entstanden sein.
Die Winter 2004/2005, 2008/2009, 2010/2011 und 2016/2017 waren sehr niederschlagsreich (CONAGUA 2017). Der Winter 2008/2009 fällt in den Auswertungszeitraum und könnte starke Erosionsprozesse begünstigt haben.
Die ermittelten Geschwindigkeiten im Maneadero Tal sind insgesamt negativ, weisen aber kleinräumige positive Geschwindigkeitsfelder auf. Im Tal wurden in den letzten Jahren vermehrt Becken für aufbereitetes Wasser und Gewächshäuser gebaut. Es lässt sich vermuten, dass durch die Auflast dieser Becken, der Boden in der Nähe lokal verformt wird und so dies zu kleinräumigen positiven Verformungsmustern führt. Die Gewächshäuser könnten sehr gute Streuer sein, weil ihre Objekteigenschaften sich nicht ändern.
In Abb. 5 B sind die Ergebnisse des absteigenden Bahnbogens aus Envisat dargestellt und mit der geologischen Karte unterlegt. Die räumliche Ausdehnung ist insgesamt geringer, als im aufsteigenden Bahnbogen, so dass die Bewegung der San Tomas Störung nördlich und südlich nicht erkennbar ist. Am San Tomas Bach sind die Geschwindigkeiten, vermutlich durch Erosions- und Ablagerungsprozesse erhöht.
Die Verwerfungen und Abschiebungen südlich der Punta Banda Bergkette sind aus den Ergebnissen nicht ableitbar. Im Maneadero Tal sind überwiegend positive Geschwindigkeiten zu beobachten und punktuell auch höhere, die durch Wasserbecken oder Gewächshäuser erklärt werden könnten, siehe oben.
In Abb. 5 C sind die Ergebnisse aus Sentinel-1a mit der geologischen Karte unterlegt. Die positiven Geschwindigkeiten auf der landzugewandten Seite der Lagune könnten durch Sedimentakkumulation entstanden sein.

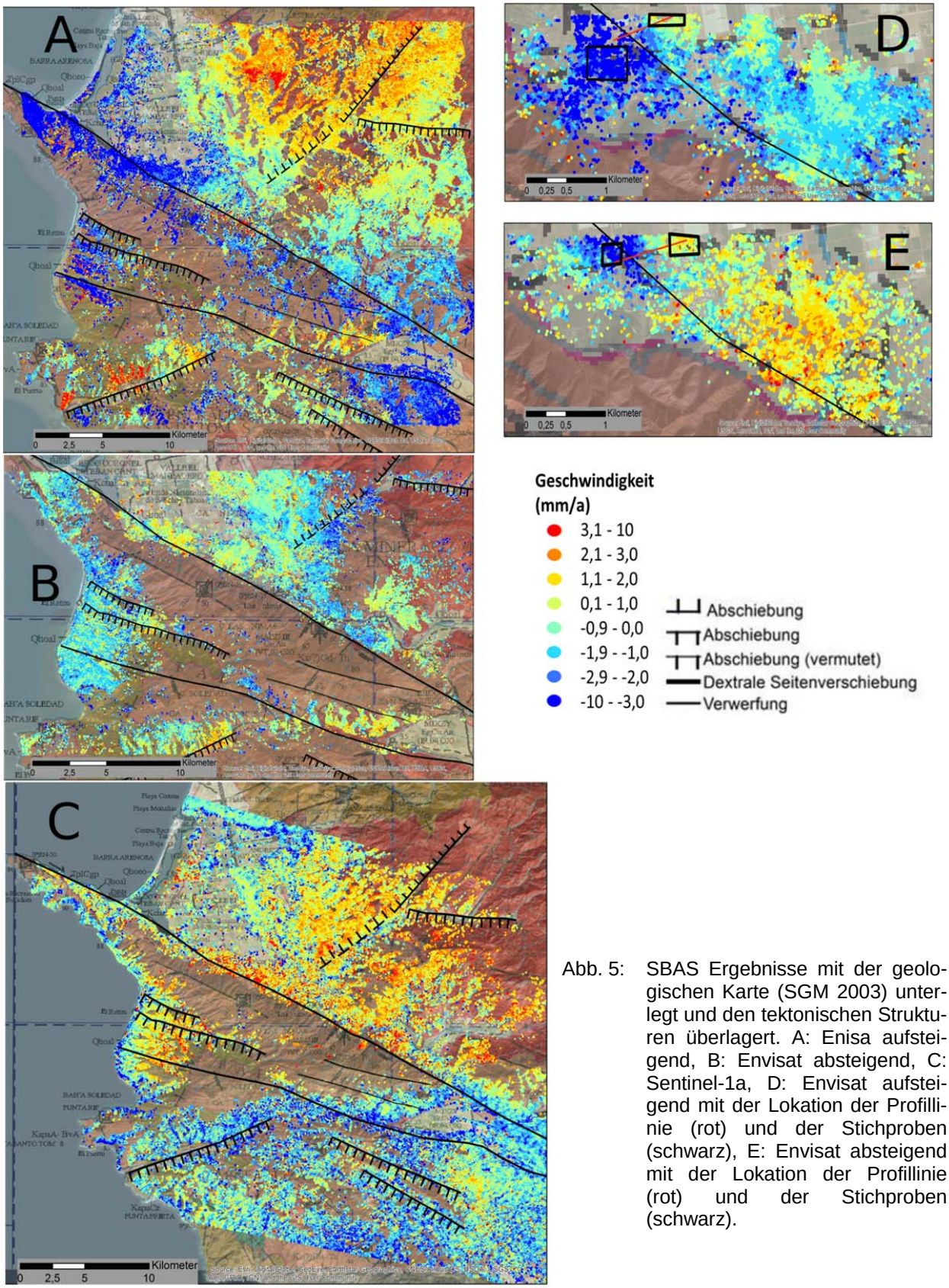


Abb. 5: SBAS Ergebnisse mit der geologischen Karte (SGM 2003) unterlegt und den tektonischen Strukturen überlagert. A: Enisa aufsteigend, B: Envisat absteigend, C: Sentinel-1a, D: Envisat aufsteigend mit der Lokation der Profillinie (rot) und der Stichproben (schwarz), E: Envisat absteigend mit der Lokation der Profillinie (rot) und der Stichproben (schwarz).

Die überwiegend negativen Geschwindigkeiten im Maneadero Tal könnten die extensive Nutzung des Grundwassers zeigen. Die Übernutzung des Aquifers könnte zu Setzungserscheinungen an der Erdoberfläche führen (DE LUNA et al. 2017). Die hohen Geschwindigkeiten in der Peninsular Bergkette und südlich der Punta Banda Bergkette am Pazifik sind topographische Effekte, die in einem weiteren Schritt korrigiert werden müssten (HANSSEN 2001).
Die negativen Geschwindigkeiten am San Tomas Bach könnten auch durch Erosion und Starkregenereignisse im Winter 2016/2017 erklärbar sein.
Beim Vergleich der Envisat mit den Sentinel Ergebnissen muss darauf geachtet werden, dass die Zeitreihen sehr unterschiedliche Zeiträume abbilden und in der Anzahl der verarbeiteten Aufnahmen stark variieren. Des Weiteren unterscheiden sich die Aufnahmegeometrien der Satellitentypen. Am San Tomas Bach sind in allen Auswertungen hohe Geschwindigkeiten erkennbar, allerdings variiert das Vorzeichen, dies könnte u.a. mit dem LOS und dem Azimut der Aufnahme zusammenhängen. Insgesamt fällt aber auf, dass die Ergebnisse im Maneadero Tal negativ sind und in der Peninsular Bergkette topographische Effekte auftreten und zu hohen Geschwindigkeiten führen, dies ist in den kürzeren Zeitreihen von Sentinel-1a und im aufsteigenden Bahnbogen von Envisat ausgeprägter.
In Abb. 4 oben sind die Ergebnisse des Ausschnitts der Agua Blanca Störung im aufsteigenden Bahnbogen von Envisat mit der geologischen Karte unterlegt und mit den Strukturen überlagert. Im Westen des Ausschnitts korreliert das negative Geschwindigkeitsfeld südlich der Agua Blanca Störung sehr genau mit der Grenze der quartären Ablagerungen (hell) zu den Unter-Kreide Andesiten und andesitischen Brekzien (rot). Nördlich der Agua Blanca Störung sind die Geschwindigkeiten überwiegend positiv. Der Geschwindigkeitswechsel scheint durch die Agua Blanca Störung verursacht zu sein, weil eine atmosphärische Filterung bereits erfolgt ist. Im Südosten des Tals befinden sich pliozäne polymiktischen Konglomerate (beige), dieses Gebiet weist schwächere Hangneigungen auf als die umliegenden Punta Banda Berge. In Luftbildern sind alluviale Fächer erkennbar, die durch Starkregenereignisse in den Wintermonaten entstanden sein könnten.
In Abb. 4 Mitte sind die Ergebnisse des absteigenden Bahnbogens mit der geologischen Karte unterlegt. Im Westen befindet sich das negative Geschwindigkeitsfeld fast ausschließlich südlich der Agua Blanca Störung, nördlich sind die Geschwindigkeiten positiv. Es könnte vermutet werden, dass die Lage der Störung nicht korrekt in der geologischen Karte erfasst ist und etwas nördlicher verläuft. Im Osten des Ausschnitts verläuft ein Flussbett, das in den Wintermonaten Wasser führen könnte und die negativen Geschwindigkeiten dort durch Erosionsprozesse erklären könnte. Die Konglomerate im Südosten sind durch positive Geschwindigkeiten gekennzeichnet.
In Abb. 4 unten sind die Ergebnisse von Sentinel dargestellt. Die Agua Blanca Störung ist durch die verbleibenden Effekte und die Phasenrampe nicht eindeutig erkennbar. Die Auswertung scheint zu Fehlerträchtig um die Störung deutlich abzubilden. Des Weiteren beträgt die Zeitreihe nur drei Jahre und ist damit auch sehr kurz im Vergleich zu den Ergebnissen von Envisat.
Der Vergleich der SBAS Auswertungen für Envisat zeigt ähnliche Geschwindigkeitsmuster und Amplituden im Maneadero Tal, die Geschwindigkeiten an den Berghängen im Osten sind unterschiedlich.

DIXON et al. (2002) berechnen aus 5 Jahren GPS-Daten, die in vier Mess-Kampagnen jeweils 3 bis 5 Tage aufgezeichnet wurden, eine horizontale Geschwindigkeitsdifferenz von 2,5 mm/a senkrecht zur Streichrichtung der Störung. In Abb. 5 D und E sind Profillinien (rot) senkrecht zur Agua Blanca Störung eingezeichnet um einen Eindruck über die Geschwindigkeitsverteilung an der Störung zu bekommen, die in LOS vorliegen. Die Abb. 6 links zeigt das Profil des aufsteigenden Bahnbogens mit einem relativ regelmäßigen Verlauf. Die Abb. 6 rechts zeigt den absteigenden Bahnbogens mit einer größeren Variation der Geschwindigkeiten.

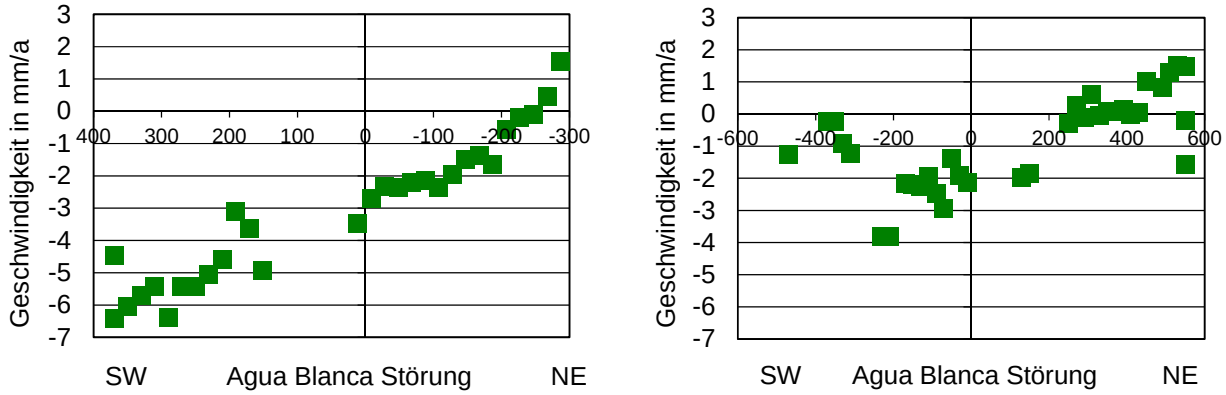


Abb. 6: Envisat-Profile senkrecht zur Agua Blanca Störung. Links: Aufsteigend, Rechts: Absteigend, Lokationen: siehe Abbildung 5 D und E.

In Abb. 5 D und E sind jeweils eine Punktwolke (schwarz) nördlich und südlich der Agua Blanca Störung markiert. Die Tabelle 2 zeigt die Berechnung von relativen Geschwindigkeiten für die Agua Blanca Störung aus diesen Punktwolken. Für den aufsteigenden Bahnbogen ergeben sich so im Mittel Geschwindigkeiten von 4,5 ±1,4 mm/a. In den Punktwolken des absteigenden Bahnbogens für Envisat ist die Geschwindigkeit im Mittel 4,1 ±1,3 mm/a.
ISYA et al. 2018 ermitteln aus den hier vorgestellten Daten des absteigenden Bahnbogens mit troposphärischer Power Law Korrektur eine Geschwindigkeit von 4,5 ±1,0 mm/a aus zwei Punkten. TS1 befindet sich nördlich der Agua Blanca Störung im Südwesten des Maneadero Tals, nahe der Lagune und weist eine Geschwindigkeit von -2,7 ±0,4 mm/a auf. Der Punkt TS2 südlich der Agua Blanca Störung befindet sich am Südhang der Punta Banda Bergkette nahe des Pazifiks und wurde mit 1,8 ±0,9 mm/a ermittelt.
ROCKWELL et al. (1993) ermitteln für die letzten 28 ka Bewegungen an der Agua Blanca Störung von 4,1 mm/a. Die von DIXON et al. (2002) berechneten Geschwindigkeiten aus GPS Daten von 2,5 mm/a senkrecht zur Störung sind niedriger, als die hier ermittelten Werte. Es muss allerdings beachtet werden, dass die hier vorgestellten Ergebnisse in LOS vorliegen und nicht in horizontale und vertikale Komponenten zerlegt wurden.

Tab. 2: Geschwindigkeitsberechnungen aus Envisat für die Agua Blanca Störung

Punktgruppe im Norden			
Bahnbogen	Stichproben-größe	Mittelwert mm/a	Standard-abweichung mm/a
Aufsteigend	75	0,1	0,8
Absteigend	130	0,6	0,6

Punktgruppe im Süden			
Bahnbogen	Stichproben-größe	Mittelwert mm/a	Standard-abweichung mm/a
Aufsteigend	328	-4,4	1,2
Absteigend	73	-3,5	1,1

Differenz von Nord und Süd			
Aufsteigend		4,5	1,4
Absteigend		4,1	1,3

4 Fazit und Ausblick

Die Ergebnisse der SBAS Methode zeigen am San Tomas Bach, im Süden des Untersuchungsgebiets, Bewegungen, die vermutlich durch Erosions- und Ablagerungsprozesse aus Starkregenereignissen verursacht wurden. Dies ist auch auf den pliozänen Konglomeraten im Südosten des Maneadero Tals erkennbar. Im Maneadero Tal sind insgesamt negative Bewegungen zu beobachten, die durch die Übernutzung des Aquifers erstanden sein könnten. Die punktuellen positiven Bewegungen im Maneadero Tal sind, vermutlich durch den Bau von Wasserbecken und Gewächshäusern, erklärbar. Die San Tomas Störung zeigt nahe dem Tal gegenläufige Geschwindigkeitsmuster, die durch die Störung verursacht sein müssen, da der Versatz einer Intrusion aus der geologischen Karte erkennbar ist.

Aus den Ergebnissen der Envisat Auswertungen lassen sich für die Agua Blanca Störung Bewegungen in der Größenordnung von 4,5 ±1,4 mm/a für den aufsteigenden Bahnbogen und 4,1 ±1,3 mm/a für den absteigenden Bahnbogen in LOS ermitteln. Korrigiert man die LOS-Bewegung mit dem Sinus des Einfallwinkels von Envisat (23°) so liegen die abgeleiteten Horizontalbewegungen zwischen 1,6 und 1,8 mm/a und sind somit etwas kleiner als die publizierten Werte von DIXON et al. (2002). Es zeigt sich, dass es mit einer ausreichend großen Datenmenge und Zeitbasis möglich ist mit radarinterferometrischen Methoden die Bewegung einer langsamen Seitenverschiebung (strike- slip fault) zu bestimmen.

Im Rahmen des laufenden Forschungsprojektes werden die Sentineldaten weiter erfasst und prozessiert. Hierbei wird auch eine gemeinsame Lösung der auf- und absteigenden Bahnbögen mit Berücksichtigung des Aufnahmewinkels in LOS angestrebt, um die horizontale Bewegung entlang der Agua Blanca Störung signifikant zu bestimmen.

5 Literaturverzeichnis

ALBERTZ, J., 2007: Einführung in die Fernerkundung- Grundlagen der Interpretation von Luft- und Satellitenbildern. Wissenschaftliche Buchgesellschaft, Darmstadt, 3. Auflage, 56-68.

BERARDINO, P., FORNARO, G., LANARI, R. & SANSOSTI, E., 2002: A New Algorithm for Surface Deformation Monitoring Based on Small Baseline Differential SAR Interferograms, IEEE Transactions on Geoscience and Remote Sensing, **40**(11), 2375-2383.

CONAGUA (COMISIÓN NACIONAL DEL AGUA), 2017: Banco nacional de datos de Aguas superficiales BANDAS, Ensendada.

DAESSLÈ, L.W., PEREZ-FLORES, M.A., SERRANO-ORTIZ, J., MENDOSA-ESPINOSA, L., MANJARREZ-MASUDA, E., LUGO-IBARRA, K.C. & GOMEZ-TREVINO, E., 2014: A geochemical and 3D-geometry geophysical survey to assess artificial groundwater recharge potential in the Pacific coast of Baja California, Mexico. Environmental Earth Science, **71**(8), 3477-3490.

DE LUNA, R.M.R., GARNÈS, S.J.A. & CABRAL, J.J.S.P., 2017: Groundwater overexploitation and soil subsidence monitoring on Recife plain (Brazil). Natural Hazards, **86**(3), 1363-1376.

DIXON, T., DECAIX, J., FARINA, F., FURLONG, K., MALSERVICI, R., BENNETT, R., SUAREZ-VIDAL, F., FLETCHER, J. & LEE, J., 2002: Seismic cycle and rheological effects on estimation of present-day slip rates for the Agua Blanca and San Miguel-Vallecitos faults, northern Baja California, Mexico. Journal of Geophysical Research, **107**(B10), 2226.

FERRETTI, A., PRATI, C. & ROCCA, F., 2001: Permanent Scatterers in SAR Interferometry, IEEE Transactions on Geoscience and Remote Sensing, **39**(1), 8-20.

GONG, W., THIELE, A., HINZ, S., MEYER, F.J., HOOPER, A. & AGRAM, P.S., 2016: Comparison of Small Baseline Interferometric SAR Processors for Estimating Ground Deformation. Remote Sensing, **8**(330), 1-26.

HANSSEN, R.F., 2001: Radar Interferometry: Data Interpretation and Error Analysis. Kluwer Academic Publisher, Dordrecht, 2, 148.

HENRY, C.D. & ARANDA-GOMEZ, J., 2000: Plate interactions control middle–late Miocene, proto-Gulf and Basin and Range extension in the southern Basin and Range. Tectonophysics, **318**, 1-26.

HIRABAYASHI, K., ROCKWELL, T.K., WESNOUSKY, S.G., STIFLING, M.W. & SUAREZ-VIDAL, F., 1996: A neotectonic study of the San Miguel-Vallecitos Fault, Baja California, Mexico. Bulletin of the Seismological Society of America, **86**, 1770-1783.

IMIP (INSTITUTO MUNICIPAL DE INVESTIGATIÓN Y PLANEACIÓN DE ENSENADA) & CODEEN (CONSEJO DE DESAROLLO ECONÓMICO DE ENSENADA), 2011: Plan estratégo des desarrollo económico del municipio des Ensenada (PEDEME). Ensenada, 56.

ISYA, L., RIEDEL, A., RIEDEL, B. & NIEMEIER, W., 2018: Comparison of Power Law Tropospheric Correction for Time Series InSAR Application. 38. Wissenschaftlich-Technische Jahrestagung der DGPF und PFGK18 Tagung in München – Publikationen der DGPF, Band 27

MOTAGH, M., 2016: InSAR Time-series Analysis. SAR-EDU Sommerschule 2016, Jena, 11-13.

NIEMEIER, W., CASARÌN, R.S. & ESPEJEL CARBAJAL, M.I., 2015: Bio-Ökonomische Folgen der Ausbeutung von Grundwassersystemen in Ariden Küstengebieten - BMBF-

Projektbeschreibung ECOAQUA. Institut für Geodäsie und Photogrammetrie, Braunschweig, 1-3.

ROCKWELL, T.K., SCHUG, D.L. & HATCH, M.E., 1993: Late quaternary slip rate along the Agua Blanca fault Baja California, Mexico. Geological investigations in Baja California, Publikationen der South coast geological Society, Abbott, P. L., Sangines E. M. & Rendina M. A. (Hrsg.), Santa Ana, 1-40.

SCHMIDT, K., WETMORE, P. & PATERSON, S., 2002: Controls on orogenesis along an ocean-continent margin transition in the Jura-Cretaceous Peninsular Ranges batholiths. Geological Society of America, Special Paper 365, 49-71.

SGM (SERVICIO GEOLÓGICO MEXICANO), 2003: Carta Geológico- Minera 1:250.000 Ensenada H11-2, Baja California. Servicio Geológico Mexicano, 1.

SPRINGER, A., 2010: Constraining basin geometry and fault kinematics on the Santo Tomás segment of the Agua Blanca Fault through a combined geophysical and structural study. Masterarbeit der University of South Florida, 11.

Validierung einer UAV-basierten Multispektralkamera zur Erfassung von Bestandesmerkmalen in Weizen

LUKAS PREY[1], MICHAEL WEIGEL[1] & URS SCHMIDHALTER[1]

Zusammenfassung: Drohnenbasierte Systeme (UAV) werden zunehmend im Precision Farming zur schnellen und günstigen Erfassung kleinskalig unterschiedlicher Bestandesmerkmale sowie in der Phänotypisierung von Feldversuchen zum Erkennen genotypischer Unterschiede eingesetzt. Ziel der Untersuchung war der Vergleich einer an einer Drohne angebrachten Multispektralkamera mit einem Hyperspektralsensor eines bodenbasierten Multisensorsystems, um Biomasse und Stickstoffstatus in Winterweizenparzellenversuchen zu detektieren. Mit beiden System wurden nur mäßig gute Bezüge mit den Referenzwerten in der vegetativen Phase gefunden, jedoch wurden Kornertrag und Kornstickstoffstatus in einem Stickstoffdüngeversuch mit beiden Sensoren gut erfasst (R^2>0,6). In einem Sortenversuch ergaben sich R^2-Werte bis 0.45 mit Red Edge-basierten UAV-Indizes und bis 0.55 mit einem Wasserbandenindex des Hyperspektralsensors, wobei die Bezüge durch Messen in der Kornfüllungsphase und durch die Wahl geeigneter Banden optimiert werden können. Die Ergebnisse legen die grundsätzliche Eignung beider Systeme nahe.

1 Einleitung

Seit jeher ist die optische Beurteilung der Gesundheit und des Ernährungszustandes von Pflanzenbeständen entscheidend für eine angemessene Kulturführung. Mit der zunehmend strenger werdenden Regulierung im Düngerecht werden eine bedarfsorientierte Nährstoffzufuhr zur Vermeidung von Überschüssen sowie die Dokumentation der Entscheidungsgrundlagen für die Bemessung der Düngung immer wichtiger. Eine besondere Bedeutung hat neben der Bemessung des Anwendungsniveaus von Dünge- und Pflanzenschutzmitteln die teilflächenspezifische Bewirtschaftung gewonnen (ALBERT et al. 2006).

Zwar gibt es seit etwa 20 Jahren entsprechende Sensoren auf dem Markt, die zur Biomasse- und Stickstoffabschätzung online im Feld angewendet werden (SCHMIDHALTER 2015), jedoch ist deren Einsatz bedingt durch die hohen Kosten bisher besonders in kleinen Betrieben kaum verbreitet (SCHARF et al. 2011).

Die Etablierung von Fernerkundungssystemen auf Drohnenplattformen (engl. unmanned aerial vehicle, UAV), wie sie von technischer Seite in den letzten Jahren fortgeschritten ist, könnte ein bedeutender Schritt sein, um die teils erprobten Technologien bodenbasierter Sensoren einer breiteren Anwendung zuzuführen (WATTS et al. 2012). Hier erscheint eine Trennung hin zur Kartierung des Schlages in Ertragspotenzialzonen und mit bei Bedarf zusätzlicher Befliegung zur Erfassung des aktuellen Bestandeszustandes und der eigentlichen Anwendung zielführend. Dabei könnten Drohnensysteme zur Ertragskartierung eine günstige indirekte Alternative zu Systemen in Erntemaschinen sein.

[1] TU München, Lehrstuhl für Pflanzenernährung, Emil-Ramann-Straße 2, D-85354 Freising, E-Mail: prey@wzw.tum.de

Hierzu erscheinen folgende Punkte von Bedeutung zu sein: (i) Überschaubare Kosten für den Endanwender, (ii) leichte Bedienbarkeit oder alternativ die Durchführung durch einen Dienstleister, (iii) technisch bedingt ausreichende Qualität der aufgenommenen Daten, (iv) zugängliche Interpretierbarkeit der Daten verbunden mit der Überführbarkeit in anwendbare Informationen sowie (v) eine verlässliche Validität der Informationen hinsichtlich abzuleitender Zielmerkmale.
Mapping-Ansätze, die sich von der zeitnahen Messung zur Anwendung unterscheiden, haben den Vorteil, Bodeneffekte stabiler erfassen zu können, und bieten damit potentiell den Vorteil, langfristig als ertragslimitierend gefundene Faktoren zu berücksichtigen (HEIL & SCHMIDHALTER 2017; SCHMIDHALTER et al. 2008), wohingegen Düngesensoren den Vorteil zeitnaher Messung und damit der Erfassung jahresspezifischer Effekte mit sich bringen. Grundsätzlich können beide Strategien zielführend sein (EBERTSEDER et al. 2005).
Meist haben UAV-Systeme den Vorteil schneller Messungen, bei denen eine Befahrung des Bestandes und damit verbundene Fahrschäden vermieden werden kann. Dazu kommen potentiell niedrigere Kosten, wenn auf teure Trägerfahrzeuge verzichtet werden kann, die Möglichkeit, vom Boden aus schwer zugänglicher Stellen zu erfassen, die Vorteile des autonomen Fluges sowie meist die Verwendung bildgebender Sensoren, die per se eine bessere optische Bestandesbeurteilung ermöglichen und deren Information für den Endanwender zugänglicher ist.
Neben der Anwendung von Sensorsystemen für Praxisschläge kommt auch der Hochdurchsatz-Phänotypisierung im Feld eine zunehmende Bedeutung zu (HAGHIGHATTALAB et al. 2016; KEFAUVER et al. 2017; ZAMAN-ALLAH et al. 2015). Ziel dabei ist es dabei, die Leistungsfähigkeit verschiedener Genotypen zu beurteilen. Dies käme einer bedeutenden Zeit- und Kosteneinsparung für die Pflanzenzüchter gleich.
Verschiedene Pflanzenmerkmale sind bezüglich der Erfassung mit Boden- und UAV-basierten Systemen von Interesse. Hierzu zählen die Bodenbedeckung als Maß für die Frühentwicklung , die frühzeitige Erfassung der Pflanzenzahl (GNÄDINGER & SCHMIDHALTER 2017), die Bestandestemperatur um Wasserstress zu detektieren (KHOT et al. 2016), die Bonitur von Pflanzenhöhe und Lager mittels Bildsegementierung oder aus Bildern erstellten 3D-Modellen (CHAPMAN et al. 2014; WATANABE et al. 2017), die Erkennung des Unkrautbesatzes (LÓPEZ-GRANADOS et al. 2016; RASMUSSEN et al. 2013), die Frühabschätzung des Ertragspotentials (Tattaris et al. 2016), die Erfassung des Nährstoffbedarfs und des Krankheitsbefalls (PSIROFONIA et al. 2017), die Identifikation von Kulturarten (REBETEZ et al. 2016), sowie die Blühdetektion und die zeitlich aufgelöste Detektion des Pflanzenwachstums (BENDIG et al. 2014) und der Seneszenz (SANKARAN et al. 2015).
Manche der Merkmale können dabei nur sinnvollerweise durch die mit UAVs mögliche hohe zeitliche Auflösung erfasst werden, da eine gesonderte Befahrung des Bestandes oft nicht möglich sein dürfte.
Weiterhin muss die Wahl des in Abhängigkeit von Plattform und Zielmerkmal passenden Sensors erfolgen. Die meisten Sensoren von einfachen RGB-Kameras bis hin zu hyperspektralen Spektrometern sind grundsätzlich auch mit UAV verwendbar (BARETH et al. 2014). Da bei der Sensorwahl aber ein Kompromiss zwischen Gewicht, Messgeschwindigkeit, räumlicher und zeitlicher Auflösung zu treffen ist, erscheint für die meisten Anwendungen die Verwendung von Multispektralkameras (MSK) ein sinnvoller Kompromiss zu sein. Bild- und Spektraldaten können dabei zu Vegetationsindizes verrechnet werden oder indirekt etwa über die Ableitung der

Bestandeshöhe mit Zielmerkmalen wie Biomasse in Bezug gesetzt werden (BENDIG et al. 2014; LIEBISCH et al. 2015; TATTARIS et al. 2016).
Die Verwendung adäquater Filter vorausgesetzt kann ein guter Bezug von multi- zu hyperspektralen Messungen erreicht werden (VON BUEREN et al. 2015).
Ziel dieser Arbeit war es deshalb, die grundsätzliche Einsetzbarkeit einer drohnenbasierten MSK für die Erfassung von Biomassebildung und N-Aufnahme während des Schossens und der frühen Kornfüllungsphase im Vergleich zu hyperspektralen Referenzmessungen zu evaluieren.

2 Material und Methoden

2.1 Feldversuche

Der Sensorvergleich wurde an drei benachbarten Winterweizenversuchen am Standort Dürnast/Freising im Frühjahr 2017 durchgeführt. Dabei wurde ein Versuch mit Doppelhaploid-Zuchtlinien (V1), ein Sortenversuch mit zwei Stickstoffdüngestufen (60 und 120 kg N/ha) (V2), ein Stickstoffdüngeversuch mit 8 N-Stufen und 2 Sorten (V3) sowie ein Saatdichteversuch mit 4 Sorten und 3 Saatdichten (150, 250 und 350 Körner/m²) (V4) berücksichtigt. V1 bestand aus 36 Zuchtlinien und 3 Referenzsorten. In V2 waren zusätzlich zur Kontrollvariante ein Block mit einer Frühsaatvariante (Saat am 28.09.) und ein Block mit einer reduzierten Fungizidvariante angelegt, in dem außer einer Beizbehandlung kein Fungizid angewendet wurde. In beiden Blöcken waren 9 Sorten in beiden N-Stufen angelegt wohingegen die Kontrollvariante aus 26 Sorten, angelegt in denselben N-Stufen bestand. Der normale Saattermin war der 24.10.2016. In V2 wurde die N-Düngung in zwei Gaben appliziert (27.03 und 18.05.), in V1 zusätzlich am 08.06. Die Gaben betrugen hier 50, 50 und 30 kg N/ha in der 1., 2. bzw. der 3. Gabe in V1. In V3 waren 8 N-Stufen angelegt, die in Abständen von 60 kg N/ha von 0 bis 420 kg N/ha reichten. Fungizidbehandlungen fanden am 18.05. und am 30.05. statt. Die Versuche waren 4-fach, V4 3-fach wiederholt angelegt. Die Vorfrucht aller Versuche war Kleegras.

2.2 Sensoren

Seit ca. 20 Jahren kommt am Lehrstuhl für Pflanzenernährung (TUM) die mobile Sensorplattform Phenotrac IV zum Einsatz, die sich aus der Erprobung in Feldversuchen heraus in verschiedenen Bauarten kontinuierlich weiterentwickelt hat (Abb. 1). Anfangs kamen überwiegend an Traktoren montierte gängige Düngesensoren zum Einsatz, die durch weitere Sensoren erweitert wurden. Zur Verfügung stehen die aktiv messenden Sensoren GreenSeeker RT100, CropCircle ACS 470 und Yara-Active Light Sensor (ALS), die jeweils nur über 2-4 spektrale Banden verfügen und sich in der Stärke der Lichtquelle, der spektralen Sensitivität und den Ansprüchen an die Messgeometrie unterscheiden (KIPP et al. 2014). Aktive Sensoren haben den Vorteil, unabhängiger von der einfallenden Strahlung zu sein. Der gleichzeitig eingesetzte passive hyperspektrale Sensor (tec5, Oberursel, Deutschland) misst bidirektional hochaufgelöst von 300-1700 nm in nominaler Auflösung von ca. 3.3 nm. Zusätzlich kann die Bestandestemperatur mit zwei Punktsensoren sowie mittels Ultraschall- und Lasersensoren die Pflanzenhöhe erfasst werden. Die Sensordaten werden integriert in 5 Hz-Taktung mittels RTK zentimetergenau erfasst. Je nach Wachstumsstadium zeigten sich dabei bisher Unterschiede zwischen den Sensoren, wobei bei richtiger Auswahl von Vegetationsindizes der passive Hyperspektralsensor das beste Potential

zeigte (ERDLE et al. 2011). Mittels der mobilen Sensorplattform werden die Versuchsparzellen mittig befahren.

Abb. 1: Mobile Sensorplattform Phenotrac IV (Lehrstuhl für Pflanzenernährung, TUM), ausgestattet mit Sensorarm und RTK-Antenne

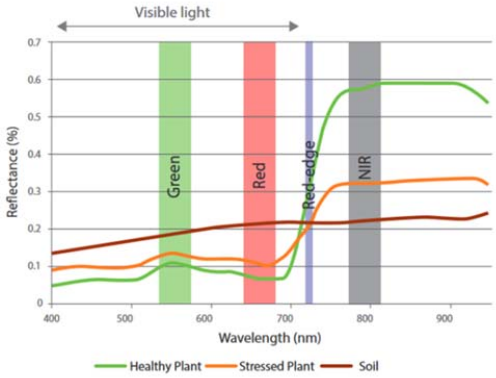


Abb. 2: Spektrale Signatur der verwendetet Sequoia-Multispektralkamera und übliche Spektren gesunder und gestresster Vegetation sowie des Bodens. Quelle: senseFly

Für die UAV-Befliegungen wurde die Starrflügler-Drohne eBee (senseFly, Cheseaux-Lausanne, Schweiz) eingesetzt. Daran war eine Sequoia-Multispektralkamera (Parrot, Paris, Frankreich) mit einer Auflösung von 1,2 MP montiert. Mit der Multispektralkamera (MSK) wurden Bilder in vier Banden aufgenommen, die mittels Referenzsensoren zur Erfassung der einfallenden Strahlung zu Reflektion verrechnet werden: Grün (530-570 nm), Rot (640-680 nm), Red Edge (730-740 nm) und NIR (770-810 nm) (Abb. 2). Die Drohne zeichnet sich durch die hohe Fluggeschwindigkeit bei niedrigem Energieverbrauch aus, so dass in Abhängigkeit der gewählten Überlappung und Auflösung die zusammenhängende Befliegung von etwa 20 ha möglich ist.

2.3 Beerntung, Sensormessungen und Datenauswertung

Über die Vegetationsperiode hinweg wurden von der Bestockung bis zur Gelbreife an 11 Terminen Messungen durchgeführt, wobei an vier Terminen in den BBCH-Stadien 32-73 um maximal einen Tag versetzt zusätzlich UAV-Befliegungen stattfanden. Dafür wurde die maximale Auflösung von 5 cm/Px bei einer Flughöhe von ca. 50 m und einer Überlappung von 80% in längs- und 60% in seitlicher Richtung gewählt. Die Flugroute wurde mit gesetzten Eckpunkten, gesteuert mit der Software eMotion 3.1 autonom beflogen. Begleitend zu den Befliegungen wurden Beerntungen der oberirdischen Biomasse in ausgewählten Parzellen in V2, V3 und V4 durchgeführt. Hierzu wurde in einer der inneren Saatreihen auf der Länge von 1 m Biomasse geschnitten, bei T3 auf 2 m. Daraus wurde Trockenmasse (TM), sowie in V2 und V4 Stickstoffgehalt (Nconc) mittels NIRS-Analytik und die resultierende Stickstoffaufnahme (NA) ermittelt.

Tab. 1: Termine von Biomassebeerntung sowie UAV- und Spektralmessungen nach Wachstumsgradtagen. ([1]mit Basiswert 5°C, bezogen auf normale Saatzeit)

Termin	Wachsstumsgradtage[1]	UAV	Spektral
T1	231	10.05.17	11.05.17
T2	286	16.05.17	17.05.17
T3	465	01.06.17	01.06.17
T4	533	08.06.17	08.06.17

Für die Spektralwerte wurde eine Versatzkorrektur durchgeführt, um die Daten dem RTK-Signal räumlich zuordnen zu können. Die UAV-Bilder wurden mittels der Software Pix4Dmapper Pro 3.2.23 zu einem Orthomosaik so zusammengefügt, dass die Szenen möglichst Nadir und mit minimaler Winkelverzerrung nutzbar waren. Die Spektral- und UAV-Daten wurden auf Parzellenebene in ArcGIS (ESRI) aggregiert. Hierzu wurden jeweils die äußeren beiden Saatreihen in Längsrichtung und die Stirnseiten zur Vermeidung von Randeffekten sowie die eigentliche Beerntungsfläche entfernt. Daten beider Systeme wurden zu verschiedenen Vegetationsindizes verrechnet. Die Reflexion der UAV-Banden wurde paarweise als „Simple-Ratio“ (SR) verrechnet, das heißt die Reflexion jeder Bande durch die einer zweiten anderen Bande geteilt. Für T3 und T4 wurden die UAV-Daten zweier SR gegen die analogen Indizes des Hyperspektralsensors validiert um die grundsätzliche Übereinstimmung der Daten zu testen.

Tab. 2. Tabelle getesteter Spektralindizes berechnet aus den bodenbasierten Phenotrac-Messungen. ‚Tec5‘ bezeichnet den Hyperspektralsensor

Abkürzung	Sensor	Formel
CC730_670	CropCircle	R730/R670
CC760_730	CropCircle	R760/R730
CC760_670	CropCircle	R760/R670
SR_Greenseeker	Greenseeker	NIR/Red
NNI	ALS	R760/R730
Biomasse	ALS	R970/R900
R780/R550	tec5	R780/R550
R780/R670	tec5	R780/R670
R780/R700	tec5	R780/R700
R760/R730	tec5	R760/R730
R780/R740	tec5	R780/R740
R900/R970	tec5	R900/R970
Reip	tec5	R700+40(((R670+R780)/2−R700)/ (R740−R700))
R1100/R1200	tec5	R1100/R1200

Für alle Termine wurden jeweils lineare Korrelationen zwischen Sensorindizes und den Zielmerkmalen TM, Nconc und NA zusammen für die Versuche V2 und V4 gerechnet. Dabei flossen für die UAV-Daten alle vier Wellenlängen, alle 6 daraus berechneten paarweisen Simple Ratio-Indizes sowie die normalisierten Indizes mit der NIR-Bande und der Grün- bzw. Rot-Bande ein. Für die bodenbasierten Sensoren wurden verschiedene Indizes verwendet, die sich

aus anderen Arbeiten mit dem Sensorsystem als geeignet für agronomische Merkmale gezeigt hatten (Tab. 2).
Zusätzlich wurden Beziehungen zum Kornertrag (V2 und V3) sowie zum Rohproteingehalt und Kornstickstoffertrag (V3) getestet.

3 Ergebnisse und Diskussion

3.1 Sensor- und Referenzdaten

Bodenbasierte Sensormessungen konnten von Ende März bis 11. Juli durchgeführt werden, während UAV-Befliegungen nur während des Schossens bis in die frühe Milchreife (11. Mai bis 08. Juni) möglich waren, bevor es zu einem technischen Ausfall des Systems kam. Durch schematisch dargestellte Interpolation der Daten über die Messzeitpunkte zeigt sich in den Hyperspektraldaten ein typischer Verlauf mit starker Absorption im sichtbaren Spektrum und dem starken Reflexionsanstieg am Red Edge bis zum NIR-Bereich (Abb. 3 links). Abgeleitet vom NDVI (rechts) bestätigt sich ein rasches Entwickeln des Bestandes im April bis zum Erreichen eines Sättigungsniveaus etwa Mitte Mai, wobei sich der Vorsprung der Frühsaatvariante deutlich zeigt. Nach der Blüte ca. am 11. Juli trat eine stetige Abnahme des Indexwertes als Indikator für zunehmende Seneszenz auf. Hier kam es zu zunehmender Differenzierung mit niedrigeren Werten der Variante mit reduzierterer Fungizidapplikation.
Grundsätzlich konnten mit beiden Sensorsystemen die Hauptbehandlungseffekte erkannt werden, wie für die Düngestufen in V3 erkennbar ist (Abb. 4).

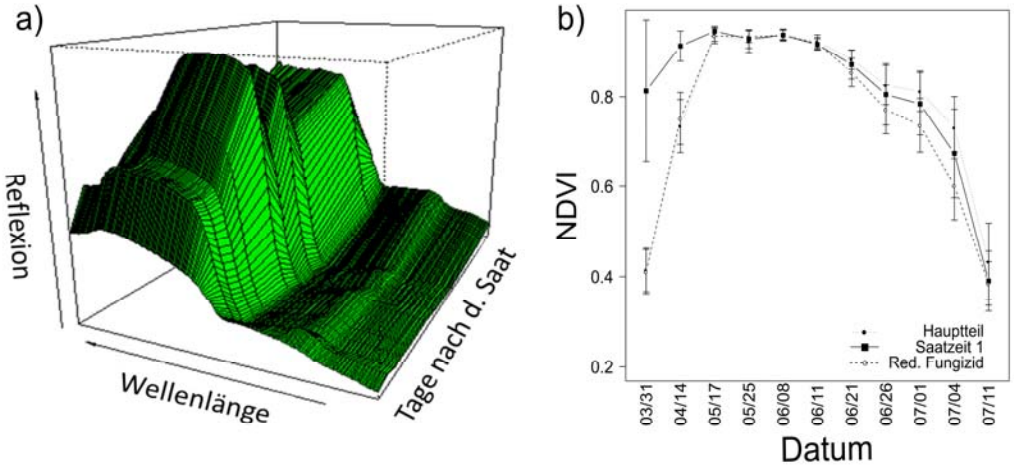


Abb. 3: Über die Zeitachse aus 11 Bestandesmessungen integrierte Pflanzenreflektion, schematisch als Mittelwert der zweiten N-Stufe in V2 dargestellt (a). NDVI dargestellt für verschiedene Messtage (Monat/Tag), aggregiert nach Versuchs-Hauptteilen (b)

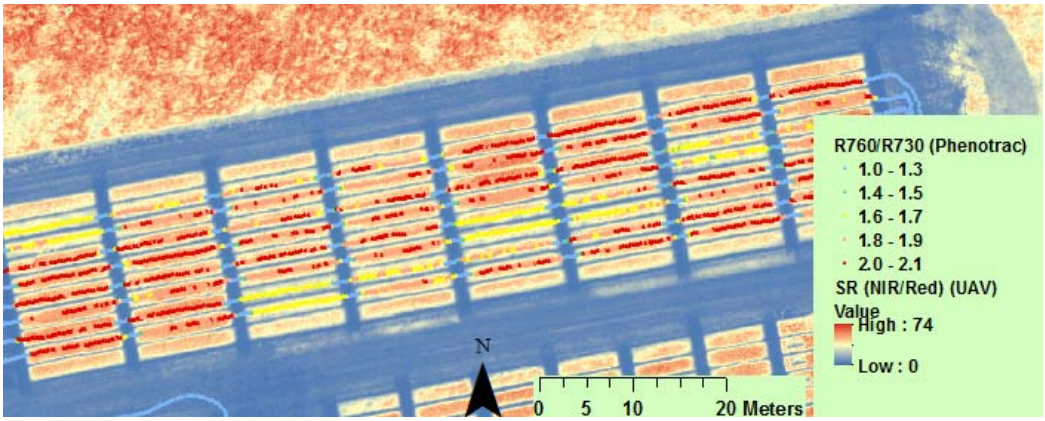


Abb. 4: UAV-SR (NIR/Rot) (Karte) und NIR/Rot-Verhältnis des Hyperspektralsensors (unten) am 08.06.2017 in V3. Die blockförmig angelegten Düngestufen sind gut erkennbar

Für die Referenzdaten war ein gewisser Versatz der TM zwischen V3 einerseits und V2 und V4 andererseits feststellbar (Abb. 5a), was zum Teil durch den Abfall in den niedrigsten N-Stufen in V3 erklärt werden kann. An allen Terminen wurden in den Parzellen in V4 leicht höhere Nconc Werte als in V2 gefunden. Dies könnte durch die meist höhere Bestandesdichte und einige wüchsige Sorten in V2 erklärt werden, so dass es hier vermutlich zu einem Verdünnungseffekt gekommen war (Lemaire et al. 2008). Dominiert durch höhere Nconc-Werte (Abb. 5b), zeigte aber V4 eine im Zeitverlauf zunehmend höhere N-Aufnahme als V2 (Abb. 5c). Dabei wurden zum Teil sehr hohe Werte erreicht, was auf die Vorfrucht und einen gewissen Parzelleneffekt zurückzuführen ist.

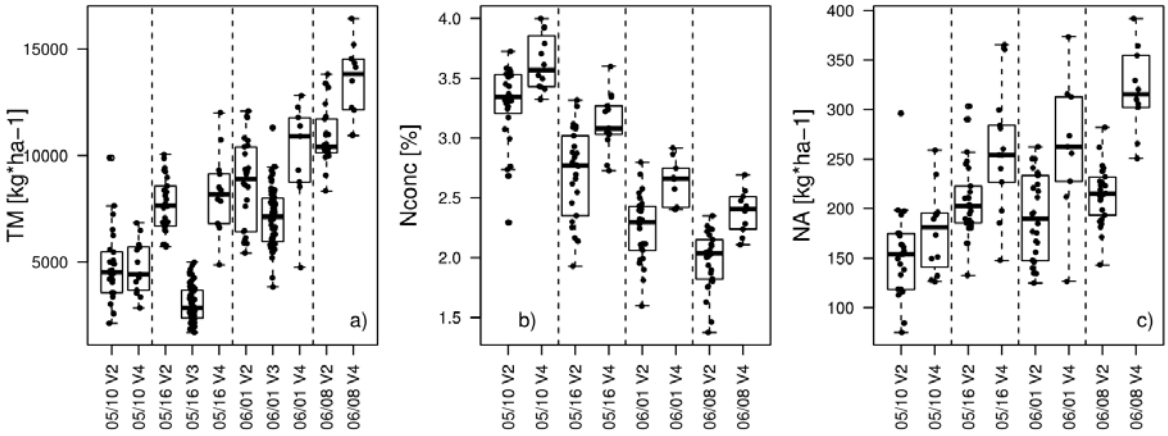


Abb. 5: Zeitliche Entwicklung von TM (a), Nconc (b) und NA (c) nach Versuchen und Beerntungsdatum

Die SR-Werte der UAV-MSK zeigten je nach Messzeitpunkt unterschiedliche Entwicklungen (Abb. 6). Das Grün/Rot-Verhältnis (G/R) sättigte bis zum 06. Juni mit zunehmend niedrigeren Werten ab. Das Grün/Red Edge-Verhältnis zeigte hingegen keine konsistente Entwicklung, was auf die gleichzeitige Abnahme der Reflexion in beiden Banden zurückzuführen ist. Ähnlich wie G/R nahm auch G/NIR ab, stieg jedoch vom 01. bis 06. Juni wieder leicht an. Das Rot/Red Edge (R/RE)-Verhältnis zeigte einen inversen Verlauf zu G/R, jedoch mit einer stärkeren Differenzierung zwischen und innerhalb der Versuche in den beiden späteren Terminen, was auf einen

leichten Anstieg bzw. Abfall der Reflexion von R und RE zurückzuführen ist (nicht dargestellt). Einen ähnlichen Verlauf zeigt dabei R/NIR, jedoch mit schwächerer Differenzierung, da sowohl die R- als auch NIR- Reflexion tendenziell anstiegen. Eine stets starke Differenzierung innerhalb der Versuche ist für RE/NIR erkennbar, wobei sich zwischen 01. und 08. Juni kaum mehr Veränderungen zeigten.

Der Vergleich analoger Banden lässt grundsätzliche eine gute Übereinstimmung zwischen den multi- und hyperspektralen Sensoren erkennen (Abb. 7 für T4). Höhere NIR/Rot-Werte des Hyperspektralsensors wurden jedoch weniger gut vom UAV-Index abgedeckt. Unterschiede könnten durch die Breite der Banden zustande kommen. Zudem ist ein direkter Vergleich durch Unterschiede im Messfeld beeinflusst, da das Messfeld punktueller Messungen des Phenotracs nicht exakt lokalisierbar ist.

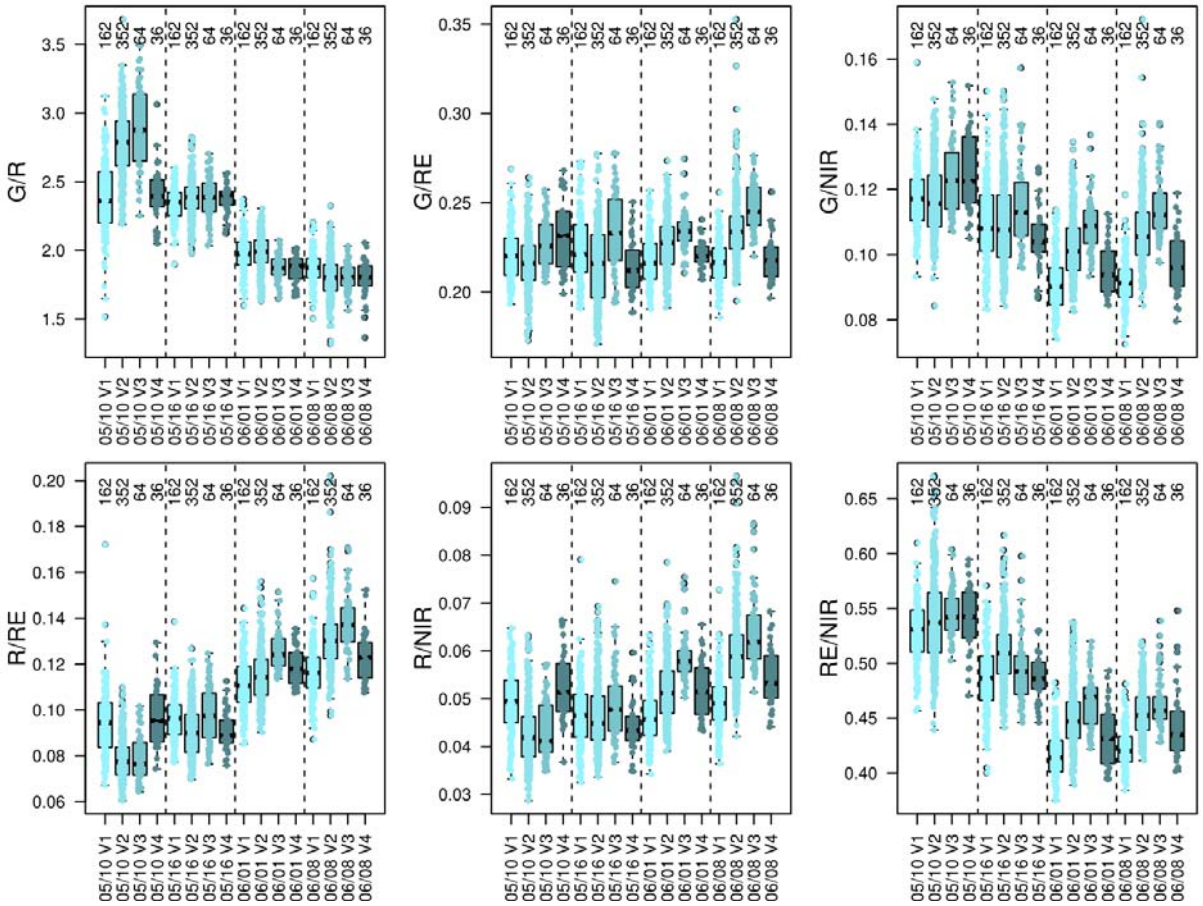


Abb. 6: Zeitliche Entwicklung der paarweise als ‚Simple Ratio' verrechneten UAV-Banden nach Versuchen und Beerntungsdatum: G: Grün, R: Rot, RE: Red Edge. Die Zahlen über den Boxplots beschreiben die Anzahl der verfügbaren Datenpunkte (Versuchsparzellen)

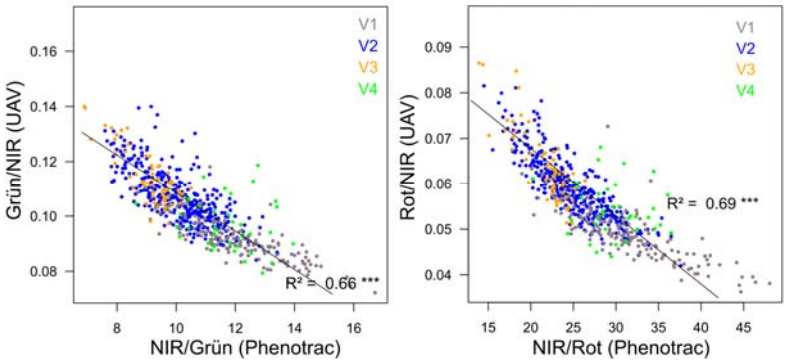


Abb. 7: Simple Ratios Grün/Rot und Rot/NIR aller Versuche dargestellt im Vergleich zu analogen Phenotrac-Indices zu T4 (08.06.)

3.2 Korrelationen mit destruktiv erfassten Bestandesmerkmalen

Für die destruktiv beernteten Merkmale TM, Nconc und NA wurden an V2 und V4 durchwegs maximal mäßig gute Bezüge mit den Sensordaten gefunden (Tab. 3,

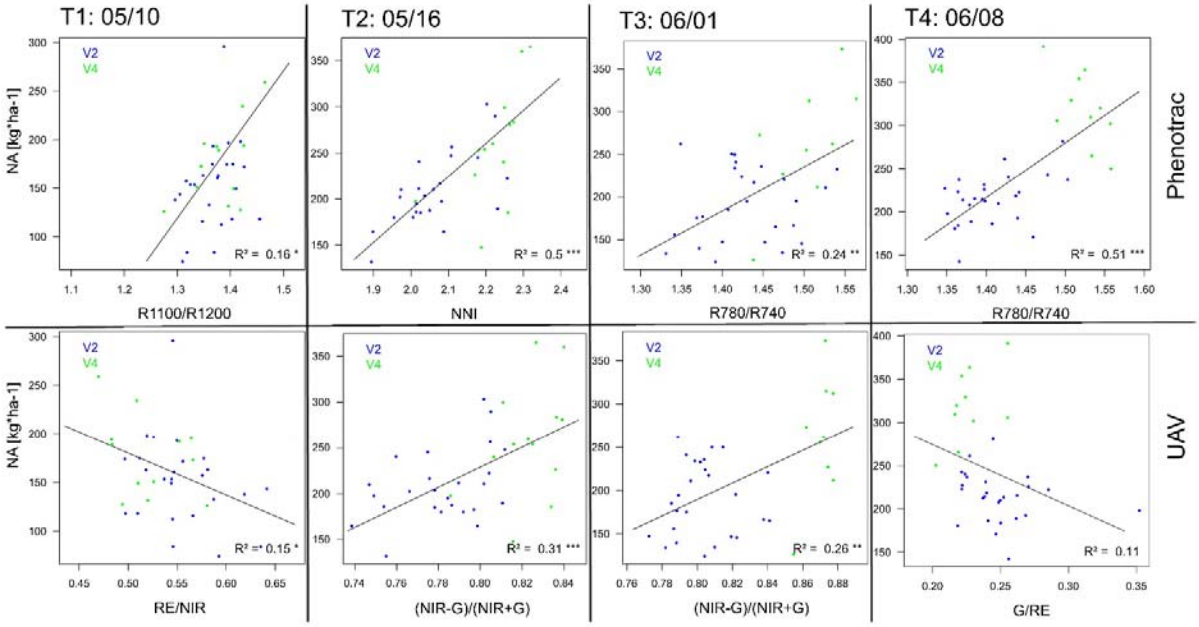


Abb. 8). Vor allem TM war mit beiden Sensoren mit einem maximalen R^2 von 0.39 (T4, R1100/R1200) schwer erfassbar. Mit Ausnahme von NA an T3 waren die besten UAV-Indizes den besten Phenotrac-Indizes stets unterlegen. Zum einen standen verschiedene Sensoren als Referenz zur Verfügung, wobei für die Nconc stets einer der aktiv messenden Düngesensoren die besten Bezüge zeigte. Zum anderen bietet der Hyperspektralsensor deutlich mehr Möglichkeiten in der Indexberechnung. Hierbei hob sich vor allem in drei Testfällen R1100/R1200 ab. Daran bestätigen sich häufig gefundene gute Bezüge von Wasserbanden mit biomassebezogenen Merkmalen (PRASAD et al. 2007). Für die N-Aufnahme zeigte an T3 und T4 R780/R740 die besten Bezüge (R^2 von 0.24 und 0.51), ein dem Red Edge Inflection Point (REIP) verwandter Index (MISTELE & SCHMIDHALTER 2006). Obwohl die getesteten UAV-Banden ähnliche Wellenlängen umfassten, zeigten sich an T4 deutlich schlechtere Bezüge. Mit Ausnahme der nur schwach sig-

nifikanten Korrelationen der UAV-Bande Rot für TM und anderen Korrelationen unter 0.2 wurden meist mit dem normalisierten Grün/NIR-Verhältnis die besten Bezüge unter den UAV-Indizes gefunden.

Tab. 3: Korrelationen für die jeweils besten gefundenen Indizes für TM, Nconc und NA nach Messtag für bodengestützte Phenotrac-Sensoren (PT) und UAV

Merkmal	Termin	R^2_{max} PT	bester Index PT	R^2_{max} UAV	bester Index UAV
TM	05/10	0.17*	R1100/R1200	0.14*	G/NIR
	05/16	0.24**	R780/R740	0.15*	Red
	06/01	0.25*	CC760/730	n.s.	RE/NIR
	06/08	0.39**	R1100/R1200	0.11*	Red
Nconc	05/10	0.22**	CC760/730	0.19**	G/R
	05/16	0.41***	NNI	0.18**	(NIR-G)/(NIR+G)
	06/01	0.40***	SR_Greenseeker	0.36***	(NIR-G)/(NIR+G)
	06/08	0.57***	SR_Greenseeker	n.s.	RE
NA	05/10	0.16*	R1100/R1200	0.15*	RE/NIR
	05/16	0.50*	NNI	0.31***	(NIR-G)/(NIR+G)
	06/01	0.24**	R780/R740	0.26**	(NIR-G)/(NIR+G)
	06/08	0.51***	R780/R740	n.s.	G/RE

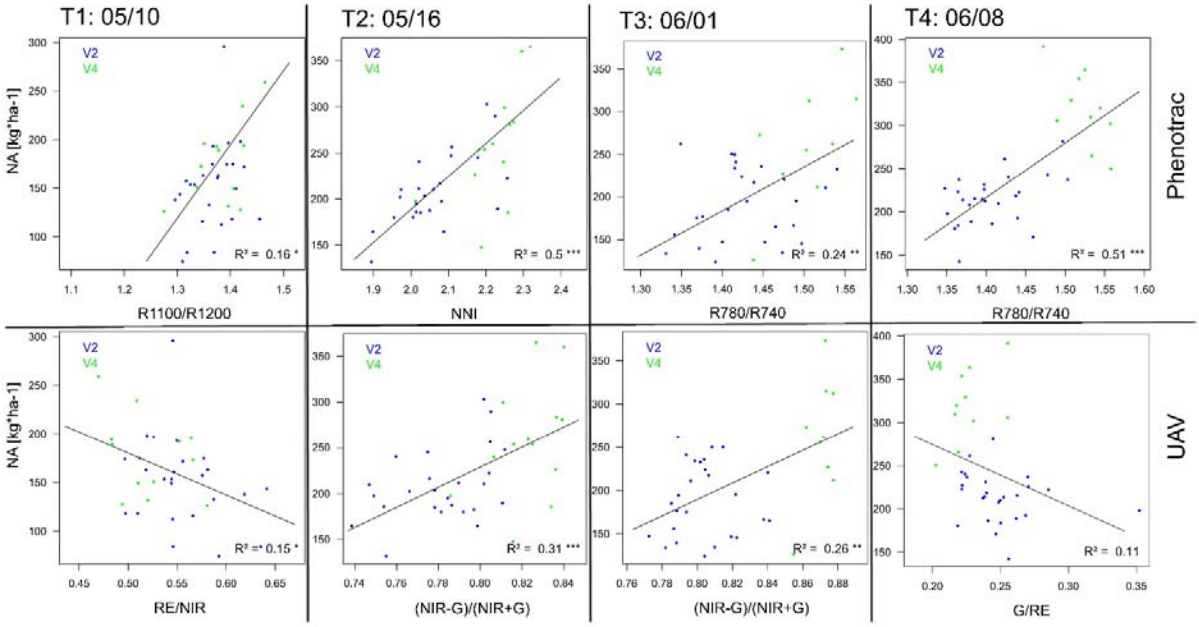


Abb. 8: NA: Darstellung im Bezug zu den jeweils besten Indizes analog zu Tab. 3 für Phenotrac-Sensoren (oben) und UAV (unten)

3.3 Korrelationen mit Kornertrag und Korn-Stickstoff-Merkmalen

Zu den wichtigsten Merkmalen für Phänotypisierungsanwendungen in Feldversuchen zählen Kornertrag und Rohproteingehalt, die den Korn-Stickstoffertrag bestimmen. Gefördert von der Differenzierung durch die N-Stufen zeigten sich vor allem in V3 gute Bezüge zischen UAV-Indizes und diesen drei Merkmalen (Tab. 4), die sich durch die Phenotrac-Indizes auf ähnlichem, aber kaum höherem Niveau bestätigen ließen (nicht dargestellt). Mit Ausnahme von T3 (01.06.),

an dem alle UAV-Indizes deutlich schlechtere R^2-Werte lieferten, war keine deutliche Abhängigkeit nach Messtagen erkennbar, jedoch nach Indizes: Das Grün-Rot-Verhältnis zeigte erneut kaum Bezüge, was die Notwendigkeit einer Red Edge- oder NIR-Bande bestätigt. Somit erscheint die Verwendung nur einer RGB-Kamera in geschlossenen Beständen nicht zielführend. Wie auch schon bei den destruktiv beernteten Merkmalen in V2, ergaben meist Kombinationen mit der NIR-Bande die besten Bezüge. Vor allem an den beiden frühen Messterminen war dabei das Verhältnis Rot zu NIR dem von Grün zu NIR überlegen. Dies widerspricht Literaturangaben zu Vorteilen der grünen Bande, bedingt durch geringere Absättigung in dichten Pflanzenbeständen (CRISTIANO et al. 2010), die sich aber auch bei Validierung der Sensoren an Mais meist nicht bestätigten (MISTELE & SCHMIDHALTER 2008). Trotz der zu erwartenden Sensitivität des Red Edge-Bereiches für den N-Status (GITELSON et al. 1996), war vor allem für Rohproteingehalt aber auch für Korn-N-Ertrag jede Kombination mit der Red Edge-Bande analogen Indizes mit NIR deutlich unterlegen. Nur für den Kornertrag ergab sich für Rot/Red Edge ein vergleichbares Niveau (Abb. 10). Hyperspektralindizes verglichen für die Abschätzung der absorbierten Strahlung hatten jedoch Vorteile einer NIR-Bande nahe über dem Red Edge bei ca. 750 nm für die Berechnung des NDVI gezeigt (VIÑA & GITELSON 2005), wofür die Red Edge-Bande der getesteten MSK offenbar aber in einem zu niedrigen Wellenlängenbereich liegt. Frühe Abschätzungen von Merkmalen zur Endreife kommen stets durch indirekte Bezüge über Biomasse, N-Aufnahme und Seneszenzeffekte zustande. Für V3 wurden nur zwei Sorten aber acht N-Stufen angebaut, wobei die N-Düngung in der gesamten Wachstumsperiode gleichgerichtet das Reflexionssignal beeinflusst haben dürfte. Im Gegensatz hierzu waren in V2 die beiden N-Stufen aufgrund der Kleegras-Vorfrucht nur schwach wirksam, so dass Ertragsunterschiede stärker durch Effekte der Sorten sowie den Saatzeitpunkt und die Fungizidvarianten zustande kamen. Diese Einflüsse ändern sich jedoch stark im Wachstumsverlauf. Für die Hyperspektraldaten wurde in V2 mit ca. 60 Vegetationsindizes für alle 11 Messtage die Korrelation mit Kornertrag berechnet.

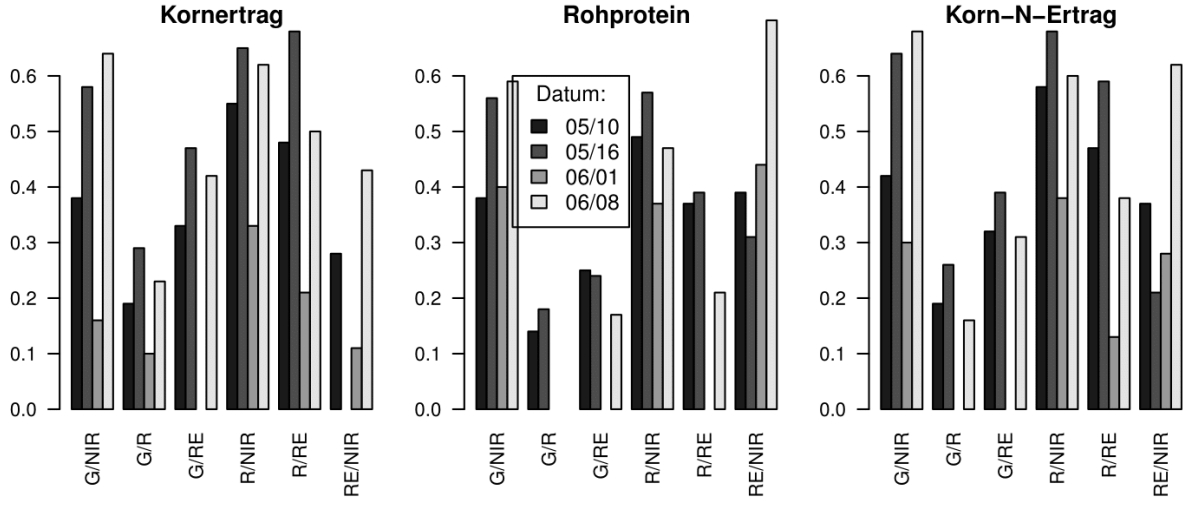


Abb. 9: Bestimmtheitsmaß (R^2) der UAV-SR-Indizes nach Messterminen für Kornertrag, Rohproteingehalt und Korn-N-Ertrag in V3.

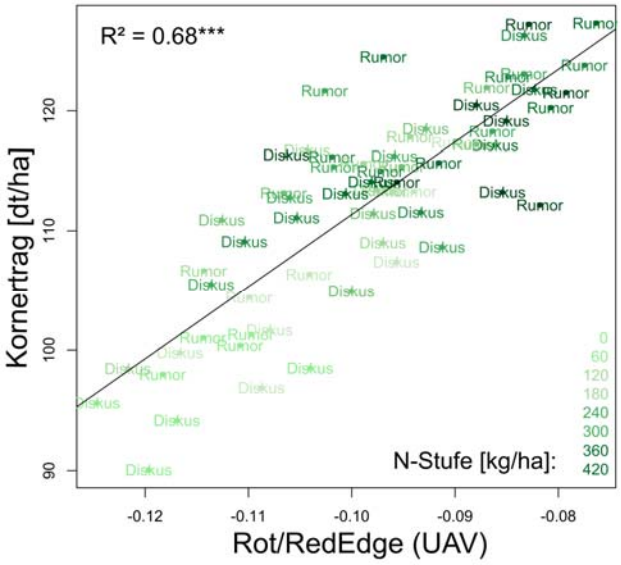


Abb. 10: Beziehung von Kornertrag und UAV-Red/RedEdge-Index am 16. Mai in V3, für beide Weizensorten kombiniert. Die Farbabstufung stellt die N-Stufen an

Tab. 4: Bestimmtheitsmaß (R^2, $p<0.05$) der UAV-Indizes (links) und der hyperspektralen Phenotrac-Indizes mit Kornertrag in V2.[1]

Datum	G/R	G/RE	G/NIR	R/RE	R/NIR	RE/NIR	Mult. Regression	R780/ R550	R780/ R670	R760/ R730	R780/ R740	R900/ R970	Reip	R1100/ R1200
05/10	0.14	0.19	0.2	0.45	0.45		0.47				0.06	0.07	0.21	0.27
05/16	0.07		0.07	0.06	0.1	0.11	0.16	0.15	0.09	0.18	0.23	0.19	0.22	0.07
06/01			0.06			0.2	0.23	0.05	0.07	0.08	0.1	0.16	0.15	0.13
06/08	0.05		0.06			0.11	0.17	0.09		0.19	0.24	0.23	0.24	0.26

[1]Die Werte der aktiven Sensoren sind auf Grund niedriger Bezüge nicht dargestellt.

Hierbei zeigten sich insgesamt für den Wasserbandenindex (R970/R900) vor allem in der Kornfüllungsphase die besten Bezüge, wobei es bedingt durch die unterschiedliche Wirksamkeit der Behandlungen zu starken Unterschieden zwischen den Messtagen kam (Abb. 11). Wies WBI ebenso wie die anderen Indizes im Schossen kaum Bezüge auf (Abb. 11), so erreichten die die UAV-SR-Indizes R/RE und R/NIR hier am 11. Mai mittelgute Bezüge (0.45), fielen bei den folgenden Messtagen aber wieder ab. Auch hier zeigten sich Kombinationen mit der Grün-Bande unterlegen (Abb. 11). Die Verrechnung aller vier Kanäle in einer multiplen Regression ergab ebenfalls keine brauchbare Verbesserung. Aufgrund des Ausfalls der Drohne war es nicht möglich, die Bezüge in der Kornfüllungsphase zu testen. Trotz schwacher Ergebnisse bis zur Blüte zeigt sich somit der Kornertrag auch für verschiedene Sorten erfassbar, wobei meist aber erst nach der Blüte brauchbare Bezüge gefunden werden (Abb. 11). Da jedoch trotz niedrigerer spektraler Auflösung bessere Bezüge aus den UAV-Daten an manchen Tagen gefunden wurden, sollte der Drohneneinsatz auch in der späten Phase zielführend sein. Eine weitere Bande im Bereich 900-1000 nm könnte aber die Bezüge verbessern.

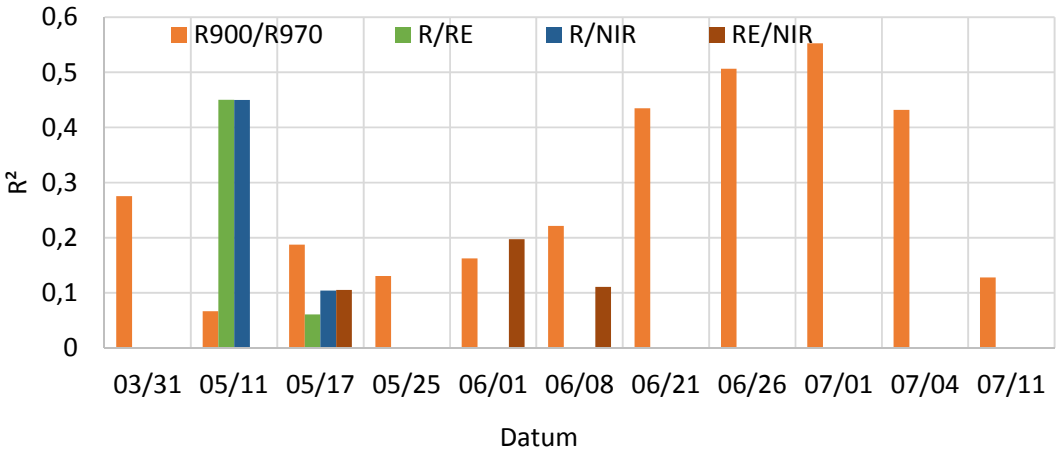


Abb. 11: V2: Bestimmtheitsmaße (R^2) ausgewählter UAV-Indizes (vgl. Tab. 4) an vier Terminen im Vergleich zu dem im Durchschnitt besten Hyperspektral-Index R900/R970 (11 Termine).

4 Schlussfolgerungen

Ziel der Untersuchung war es, die mit bodenbasierten Sensoren besser erprobten Methoden der Bestandeserfassung, wie sie im Precision Farming im Einsatz sind und für die Phänotypisierung zunehmend optimiert werden, mit einer einfachen, an einer Drohne montierten Multispektralkamera (MSK) zu überprüfen. Dabei wurden (i) destruktiv beerntete Merkmale einiger Parzellen bis kurz nach der Blüte sowie (ii) Kornertrag und Korn-N-Ertrag und Rohproteingehalt erfasst. Trotz der stark differenzierten N-Düngung war deren Wirksamkeit in im N-SteigerungsversuchV3 wegen der Vorfrucht schwach, so dass Unterschiede im Kornertrag von maximal ca. 40 dt/ha bei hohem Ertragsniveau durchaus Heterogenitäten in Praxisschlägen abbilden. Somit zeigen die guten Korrelationen in (ii), dass das aktuelle UAV-System durchaus zur Erfassung von Praxisschlägen zum Einsatz kommen kann. Zwar wurden die Merkmale bei (i) allenfalls mittelgut erfasst, jedoch konnten sich nur manche der Vergleichssensoren an einigen Messtagen davon absetzen. Beides mag an der geringen Anzahl der Datenpunkte und der kleinen Beerntungsfläche liegen, wodurch ein räumlicher Versatz zu der von den Sensoren erfassten Fläche nicht vermeidbar war, aber durch Heterogenitäten innerhalb der Parzellen überlagert war. Dennoch müssen gefundene Bezüge in (ii) durch die Erfassbarkeit von Biomasse und N-Aufnahme zum Messzeitpunkt zu erklären sein. Die oft relativ guten Korrelationen von Wasserbandenindizes des Hyperspektralsensors lassen darauf schließen, dass auch die MSK von einer ähnlichen Bande profitieren könnte. Die Rot-Bande war der Grün-Bande meist überlegen, welche aber in früheren Stadien zur Erfassung des Bedeckungsgrades ihre Berechtigung haben dürfte. Deutlich bessere Bezüge der MSK-Daten zu einem Zeitpunkt lassen auf den Vorteil zeitlich fast simultaner Messung innerhalb von ca. 10 Minuten schließen, wohingegen die bodenbasierte Messung der vier Versuche ca. 2-3 Stunden in Anspruch nahm und deshalb durch Veränderung der Einstrahlung, der Umgebungstemperatur und des Wasserstatus negativ beeinflusst werden kann. Zudem lagen an den Terminen T1 und T2 die Sensormessungen um einen Tag auseinander, was die direkte Vergleichbarkeit etwas einschränkt. Weitere Sensorvergleiche in früheren

als auch in späteren Wachstumsphasen, mit weiteren Kulturen und mehr Referenzdaten in der vegetativen Phase sind wünschenswert. Eine weitere Auswertung mit in V2 erfassten Seneszenzdaten schließt sich an. Für die Implementierung derartiger Systeme wären On-Farm-Versuche sowie Befliegungen von Zuchtgärten zielführend, wobei sich zum Teil ohnehin vorhandene Referenzdaten nutzen ließen.

5 Dank

Die Autoren danken der Firma SenseFly für die Bereitstellung des Drohnensystems. Weiterer Dank gilt Jürgen Plass für die Durchführung der Befliegung und die Vorprozessierung der UAV-Daten.

6 Literaturverzeichnis

ALBERT, E., KOWALEWSKY, H. H., LORENZ, F., ORTSEIFEN, U. & SCHINTLING-HORNY, L. VON, 2006: N-Düngung effizient gestalten. DLG-Merkblatt, **350**, DLG Verlag Frankfurt.

BARETH, G., AASEN, H., BENDIG, J., GNYP, M.L., BOLTEN, A., JUNG, A., MICHELS, R. & SOUKKAMÄKI, J., 2014: Spectral comparison of low-weight and UAV-based hyperspectral frame cameras with portable spectroradiometer measurements. Proc. Work. UAV-basaed Remote Sens. Methods Monit. Veg., **94**, 1-6.

BENDIG, J., BOLTEN, A., BENNERTZ, S., BROSCHEIT, J., EICHFUSS, S. & BARETH, G., 2014: Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. Remote Sens., **6**, 10395-10412.

CHAPMAN, S., MERZ, T., CHAN, A., JACKWAY, P., HRABAR, S., DRECCER, M., H., HOLLAND, E., ZHENG, B., LING, T. & JIMENEZ-BERNI, J., 2014: Pheno-Copter: A Low-Altitude, Autonomous Remote-Sensing Robotic Helicopter for High-Throughput Field-Based Phenotyping. Agronomy, **4**, 279-301.

CRISTIANO, P.M., POSSE, G., DI BELLA, C.M. & JAIMES, F.R., 2010: Uncertainties in fPAR estimation of grass canopies under different stress situations and differences in architecture. Int. J. Remote Sens., **31**, 4095-4109.

EBERTSEDER, T., SCHMIDHALTER, U., GU TSER, R., HEGE, U., JUNGERT, S. & STAFFORD, J. V, 2005: Evaluation of mapping and on-line nitrogen fertilizer application strategies in multi-year multi-location static field trials for increasing nitrogen use efficiency of cereals. Proc. Fifth Eur. Conf. Precis. Agric. 327-335.

ERDLE, K., MISTELE, B. &SCHMIDHALTER, U., 2011: Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. F. Crop. Res., **124**, 74-84.

GITELSON, A.A., MERZLYAK, M.N.& LICHTENTHALER, H.K., 1996: Detection of Red Edge Position and Chlorophyll Content by Reflectance Measurements Near 700 nm. J. Plant Physiol., **148**, 501-508.

GNÄDINGER, F. & SCHMIDHALTER, U., 2017: Digital Counts of Maize Plants by Unmanned Aerial Vehicles (UAVs). Remote Sens., **9**, 544.

HAGHIGHATTALAB, A., GONZÁLEZ PÉREZ, L., MONDAL, S., SINGH, D., SCHINSTOCK, D.,

RUTKOSKI, J., ORTIZ-MONASTERIO, I., SINGH, R.P., GOODIN, D. & POLAND, J., 2016: Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. Plant Methods, **12**, 35.

HEIL, K. & SCHMIDHALTER, U., 2017: Improved evaluation of field experiments by accounting for inherent soil variability. Eur. J. Agron., **89**, 1-15.

KEFAUVER, S.C., VICENTE, R., VERGARA-DÍAZ, O. & FERNANDEZ-GALLEGO, J.A., 2017: Comparative UAV and Field Phenotyping to Assess Yield and Nitrogen Use Efficiency in Hybrid and Conventional Barley **8**, 1-15.

KHOT, L.R., SANKARAN, &S., CARTER, A.H., JOHNSON, D.A. & CUMMINGS, T.F., 2016: UAS imaging-based decision tools for arid winter wheat and irrigated potato production management. Int. J. Remote Sens., **37**, 125-137.

KIPP, S., MISTELE, B. & SCHMIDHALTER, U., 2014: The performance of active spectral reflectance sensors as influenced by measuring distance, device temperature and light intensity. Comput. Electron. Agric., **100**, 24-33.

Lemaire, G., Jeuffroy, M.-H. & Gastal, F., 2008: Diagnosis tool for plant and crop N status in vegetative stage. Eur. J. Agron., **28**, 614-624.

Liebisch, F., Kirchgessner, N., Schneider, D., Walter, A. & Hund, A., 2015: Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach. Plant Methods, **11**, 9.

López-Granados, F., Torres-Sánchez, J., De Castro, A.-I., Serrano-Pérez, A., Mesas-Carrascosa & F.-J. Peña, J.-M., 2016: Object-based early monitoring of a grass weed in a grass crop using high resolution UAV imagery. Agron. Sustain. Dev., **36**, 67.

MISTELE, B. & SCHMIDHALTER, U., 2008: Spectral measurements of the total aerial N and biomass dry weight in maize using a quadrilateral-view optic. F. Crop. Res., **106**, 94-103.

MISTELE, B. & SCHMIDHALTER, U., 2006: Tractor-Based Quadrilateral Spectral Reflectance Measurements to Detect Biomass and Total Aerial Nitrogen in Winter Wheat. Agron. J., **102**, 499-506.

PRASAD, B., CARVER, B.F., STONE, M.L., BABAR, M.A., RAUN, W.R. & KLATT, A.R., 2007: Potential use of spectral reflectance indices as a selection tool for grain yield in winter wheat under great plains conditions. Crop Sci., **47**, 1426-1440.

PSIROFONIA, P., SAMARITAKIS, V., ELIOPOULOS, P. & POTAMITIS, I., 2017: Use of Unmanned Aerial Vehicles for Agricultural Applications with Emphasis on Crop Protection: Three Novel Case-studies. Int. J. Agric. Sci. Technol., 30-39.

RASMUSSEN, J., NIELSEN, J., GARCIA-RUIZ, F., CHRISTENSEN S. & STREIBIG, J.C., 2013: Potential uses of small unmanned aircraft systems (UAS) in weed research. Weed Res., 242-248.

REBETEZ, J., SATIZÁBAL, H.F., MOTA, M., NOLL, D., BÜCHI, L., WENDLING, M., CANNELLE, B., PEREZ-URIBE, A. & BURGOS, S., 2016: Augmenting a convolutional neural network with local histograms - A case study in crop classification from high-resolution UAV imagery. ESANN 2016 - 24th Eur. Symp. Artif. Neural Networks, 27-29.

SANKARAN, S., KHOT, L.R., ESPINOZA, C.Z., JAROLMASJED, S., SATHUVALLI, V.R., VANDEMARK, G.J., MIKLAS, P.N., CARTER, A.H., PUMPHREY, M.O., KNOWLES & N.R., PAVEK, M.J., 2015: Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. Eur. J. Agron., **70**, 112-123.

SCHARF, P.C., SHANNON, D.K., PALM, H.L., SUDDUTH, K.A., DRUMMOND, S.T., KITCHEN, N.R.,

MUELLER, L.J., HUBBARD, V.C. & OLIVEIRA, L.F., 2011: Sensor-Based Nitrogen Applications Out-Performed Producer-Chosen Rates for Corn in On-Farm Demonstrations. Agron. J., **103**, 1683.

SCHMIDHALTER, U., MAIDL, F.-X., HEUWINKEL, H., DEMMEL, M., AUERNHAMMER, H., NOACK & P.O. ROTHMUND, M., 2008.:Precision Farming – Adaptation of Land Use Management to Small Scale Heterogeneity. Perspectives for Agroecosystem Management, 121-199.

TATTARIS, M., REYNOLDS, M.P. & CHAPMAN, S.C., 2016: A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding. Front. Plant Sci., **7**, 1-9.

VIÑA, A. & GITELSON, A.A., 2005: New developments in the remote estimation of the fraction of absorbed photosynthetically active radiation in crops. Geophys. Res. Lett., **32**, 1-4.

VON BUEREN, S.K., BURKART, A., HUEN6I, A., RASCHER, U., TUOHY, M.P. & YULE, I.J., 2015: Deploying four optical UAV-based sensors over grassland: Challenges and limitations. Biogeosciences, **12**, 163-175.

WATANABE, K., GUO, W., ARAI, K., TAKANASHI, H., KAJIYA-KANEGAE, H., KOBAYASHI, M., YANO, K., TOKUNAGA,7 T., FUJIWARA, T., TSUTSUMI, N. & IWATA, H., 2017: High-Throughput Phenotyping of Sorghum Plant Height Using an Unmanned Aerial Vehicle and Its Application to Genomic Prediction Modeling. Front. Plant Sci. **8**, 1-11.

WATTS, A.C., AMBROSIA, V.G. & HINKLEY, E.A., 2012: Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use. Remote Sensing, **4**(6), 1671-1692.

ZAMAN-ALLAH, M., VERGARA, O., ARAUS, J.L., TAREKEGNE, A., MAGOROKOSHO, C., ZARCO-TEJADA, P.J., HORNERO, A., ALBÀ, A.H., DAS, B., CRAUFURD, P., OLSEN, M., PRASANNA, B.M. & CAIRNS, J., 2015: Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. Plant Methods, **11**, 35.

Quantifizierung der Ährenanzahl mittels UAV gestützter Bildaufnahmen

NORMAN WILKE[1], BASTIAN SIEGMANN [1], ANDREAS BURKART[2], LASSE KLINGBEIL[3] & UWE RASCHER[1]

Zusammenfassung: Die Erfassung von Vegetationsparametern in landwirtschaftlichen Beständen gewinnt zunehmend an Bedeutung. Mit Hilfe von Unmanned Aerial Vehicles (UAVs) können Kulturpflanzen mit einer hohen räumlichen und zeitlichen Auflösung phänotypisiert werden. In der vorliegenden Studie wird die Ährenanzahl von Gerste mit Hilfe eines Multikopters und einer RGB Kamera multitemporal erfasst. Zur Validierung des UAV-Ergebnisses wurden die Ähren im Feld innerhalb eines definierten Quadratmeters gezählt. Um den optimalen Aufnahmezeitpunkt für die Erfassung der Ähren mittels UAV-Aufnahmen identifizieren zu können, wurde im wöchentlichen Abstand die Ährenanzahl an einer Versuchsparzelle im UAV-Einzelbild bestimmt. Verglichen mit der Referenzmessung im Feld konnten in den UAV-Einzelbildern etwa jede zweite Ähre innerhalb des vordefinierten Quadratmeters erfasst werden. Bei der Analyse der Diskrepanz zwischen Referenzmessung und Einzelbild war auffällig, dass die Werte hinsichtlich der Genotypen und Saatdichten konstant zwischen 50% und 56% abweichen und die Standardabweichung (± 3%) gering ist. Wird dieser Wert von zukünftigen Untersuchungen bestätigt, könnte die aus den UAV-Bildern ermittelte Ährenanzahl mit einem entsprechenden Faktor multipliziert werden, um eine genauere Erfassung der Ährenanzahl zu ermöglichen. Weiterhin zeigen die Ergebnisse, dass der Termin direkt zu Beginn der Ährenentwicklung (BBCH 61) der beste Zeitpunkt für die Bonitur mittels UAV-Bildaufnahmen ist. Beeinflusst wurde dies durch den hohen Farbkontrast in diesem Entwicklungsstadium, in dem sich die Ähre deutlich vom Rest der Pflanze abhebt.

1 Einleitung

Die Einsatzgebiete von Unmanned Aerial Vehicles (UAVs) wie Multikopter, Modellflugzeuge oder ähnlichen Flugobjekten mit Autopiloten sind vielseitig und breit gefächert. Abhängig von der angestrebten Datenaufnahme bieten unterschiedliche Systeme verschiedene Vor- und Nachteile (KHOT et al. 2015). Neben der in der Studie wesentlichen Domäne der Landwirtschaft (Burkart et al. 2015), finden sie ebenfalls Verwendung in der Forstwirtschaft (GRENZDÖRFFER et al. 2008), in der Geowissenschaft (WESTOBY et al. 2012) sowie in der Geodäsie (ELING et al. 2015) oder beim Militär (BENJAMIN & EHRENREICH 2013). Gründe für die vielseitige Verwendung sind die niedrigen Kosten, die leichte Bedienbarkeit und die deutlich höhere Flexibilität beispielsweise im Vergleich zu klassischen Flugzeugen. Die notwendige Vorbereitungszeit mit dem UAV ist gering und die Bodenauflösung kann, verglichen mit Satellitenbildern, erheblich gesteigert werden (CANDIAGO et al. 2015).

[1] Forschungszentrum Jülich GmbH, IBG-2: Plant Sciences, Wilhelm-Johnen-Straße, D-52428 Jülich, E-Mail: [n.wilke, b.siegmann, u.rascher]@fz-juelich.de
[2] JB Hyperspectral Devices UG, Am Botanischen Garten 33, D-40225 Düsseldorf, E-Mail: andreas@jb-hyperspectral.com
[3] Universität Bonn, Institut für Geodäsie und Geoinformation, Nußallee 17, D-53115 Bonn, E-Mail: klingbeil@igg.uni-bonn.de

Die Erfassung von Vegetationsparametern in landwirtschaftlichen Beständen gewinnt dabei zunehmend an Bedeutung (SANKARAN et al. 2015; TORRES-SÁNCHEZ et al. 2014; BENDIG et al. 2013; LIU et al. 2014). Moderne Verfahren der Bildverarbeitung in Kombination mit neuen Sensoren und autonom agierenden Kleinflugzeugen werden in den kommenden Jahren eine Revolution in der landwirtschaftlichen Pflanzenzüchtung und in der Nutzpflanzenproduktion nach sich ziehen.
Als Bestandteil dieser Entwicklung wird in der vorliegenden Studie die Ährenanzahl von Gerste mit Hilfe einer RGB Kamera und eines Multikopters unter Verwendung verschiedener Bildverarbeitungsschritte (z. B. Structure from Motion) erfasst. Die Ährenanzahl kann zur Phänotypisierung, zur Pflanzenzüchtung oder ebenfalls zur Ertragsabschätzung verwendet werden. Die frühzeitige Abschätzung des Ertrags wird unter anderem für die Organisation der Ernte- und Lagerlogistik benötigt. Im Bereich Precision Agriculture (PA) kann sie zusätzlich dazu genutzt werden, den späteren Einsatz von Düngemitteln zu optimieren oder dessen Wirksamkeit zu überprüfen (COINTAULT et al. 2008). Der Begriff PA umfasst Techniken und Methoden, die Heterogenität eines Bestandes ortspezifisch zu erfassen. So wird zum einen umweltschonende Landwirtschaft ermöglicht (LELONG et al. 2008), zum anderen kann die Farmprofitabilität und Pflanzenproduktivität optimiert werden. In Anbetracht der stetig wachsenden Bevölkerung und der abnehmenden Flächenverfügbarkeit pro Kopf stellt dies ein wichtiges Unterfangen dar (ATZBERGER 2013).
Durchgeführt wurde die Ährenerkennung bereits unter Verwendung von Lichtschrankensensoren durch KOVACHEVA et al. (2014). Mit Hilfe der Lichtschranke wird ein Schattenbild von den Pflanzen erzeugt. Anhand von Bildbearbeitungsalgorithmen kann die Ährenanzahl automatisiert bestimmt werden. Im Gegensatz zur UAV-Aufnahme können durch die seitliche Erfassung mit einer Lichtschranke Informationen gewonnen werden, die durch den Blick von oben (Nadir) nur eingeschränkt erfasst werden können. Der Abstand zwischen den Lichtgittern ist mit 30 Zentimetern sehr gering, um jede Einzelpflanze und Ähre identifizieren zu können. Demenentsprechend wäre die Erfassung eines Bestandes mit dieser Methode allerdings sehr zeit- und kostenintensiv. Des Weiteren wurde von COINTAULT et al. (2008) bereits versucht die Ähren mit einem Stativ und einer RGB Kamera automatisch zu identifizieren. Der Abstand zwischen Kamera und Bestand betrug dabei etwa 0,85 Meter, in dem ein Bereich von 0,25 Quadratmetern untersucht werden konnte. Verglichen mit einer manuellen Referenzmessung im Bild konnte durch eine automatische Erkennung eine Genauigkeit in einem Bereich von 73% bis 85% erzielt werden. Eine Referenzmessung im Feld wurde nicht durchgeführt, da diese durch den geringen Abstand nicht zwingend erforderlich ist.
Das UAV-Verfahren liefert im Vergleich zu den genannten Praktiken die Möglichkeit, die Ährenanzahl auf einer deutlich größeren Fläche zu quantifizieren. Um die Aussagekraft des UAV-Ergebnisses abzuschätzen, wurde eine manuelle Referenzmessung im Feld durchgeführt.

2 Material und Methoden

2.1 Versuchsfläche

Sommergerste (Hordeum vulgare) wurde am 09.04.2016 mit einer niedrigen und normalen Saatdichte in der Nähe von Bonn (Campus Klein-Altendorf, GPS: 50.617285, 6.990364, WGS84)

ausgesät. Die normale Saatdichte mit 300 Körnern pro Quadratmeter beschreibt die in der Regel übliche Praxis in der Landwirtschaft. Die niedrige Saatdichte betrug 150 Körner pro Quadratmeter. In der Studie wurden zwei zweizeilige Sorten (Buloke, Commander) mit drei Wiederholung je Saatdichte untersucht. Eine Versuchsparzelle hatte eine Größe von 7,86 m². Für die Georeferenzierung der Bilddaten wurden acht Passpunkte gleichmäßig auf der Versuchsfläche verteilt.

2.2 Datenerfassung

Für die Datenaufnahme wurde der Falcon-8 Oktokopter von der Firma AscTec (Ascending Technologies GmbH, Krailing, Deutschland) verwendet. Dieser ist mit einer Sony Alpha NEX-5N RGB Weitwinkelkamera (16 Megapixel, Bildauflösung 4.912 × 3.264) ausgestattet. Die RGB Kamera ist in einen Gimbal integriert, der die UAV-Bewegungen ausgleicht und die zuvor programmierte Kameraausrichtung beim Flug hält.
Bei einer Flughöhe von 10 Metern (Bodenauflösung: 2,3 mm/Pixel) wurde die UAV-Aufnahmen 89 Tage nach Aussaat (TnA) bei einer Kameraausrichtung von 90° (Nadir) aufgenommen.
Um das UAV-Ergebnis validieren zu können, wurden vor dem Überflug auf horizontaler Ebene Schnüre gespannt, die in den Versuchsparzellen einen Quadratmeter definieren (Abb. 2). Für die Referenzmessung im Feld wurden die Ähren zum gleichen Zeitpunkt 89 TnA innerhalb des Quadratmeters mit Blickrichtung in die Saatreihe gezählt, sodass auch tiefer liegende Ähren in die Zählung einflossen (Abb. 1).

Abb. 1: Zählung der Ähren für die Referenzmessung im Feld mit Blickrichtung in die Saatreihe

Die Entwicklung von der Blüte (BBCH 61) bis hin zur Frucht- und Samenreife (BBCH 89) fand in einem Zeitfenster über sechs Wochen (68 TnA bis 102 TnA) statt. Um den optimalen Aufnahmezeitpunkt für die Erfassung der Ähren mittels UAV-Bilddaten identifizieren zu können, wurde im wöchentlichen Abstand die Ährenanzahl an einer Versuchsparzelle im UAV-Einzelbild bestimmt.

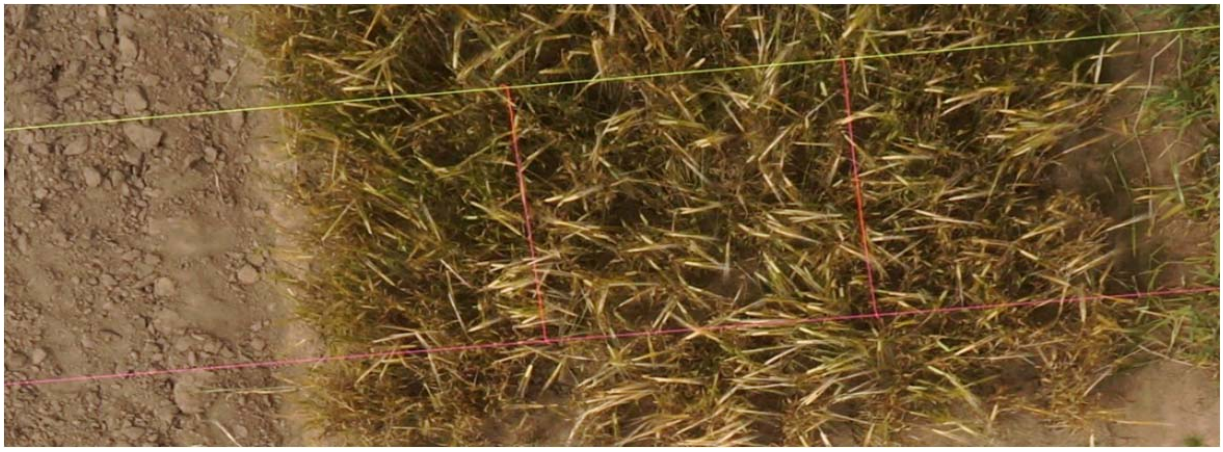


Abb. 2: Definition eines Quadratmeters unter Verwendung von Schnüren zur Validierung der UAV-Ergebnisse mit der Referenzmessung im Feld

2.3 Bildprozessierung und Analyse

Die UAV-Einzelbilder wurden unter Verwendung der Structure from Motion (SfM) Technik in Agisoft Photoscan (Version 1.2.4) prozessiert. Der Begriff SfM stammt aus dem Bereich der Computer Vision. Er bezeichnet einen Prozess mit verschiedenen Algorithmen, indem korrespondierende Merkmale in Bildern erkannt und durch eine gewisse Anzahl überlappender Bilder in einer räumlichen dreidimensionalen (3D) Struktur dargestellt werden können (WESTOBY et al. 2012). Details zu den Algorithmen der SfM Technik werden in LOWE (2004) oder KERSTEN (2016) behandelt. Die durchgeführten Schritte für die Datenprozessierung und die dabei verwendeten Einstellungen werden in Abb. 3 zusammengefasst. Auf der Grundlange einer erstellten 3D Punktwolken (2. Prozess) sowie durch die Generierung von Oberflächenelemente (3. Prozess), kann ein georeferenziertes Orthomosaik berechnet werden.

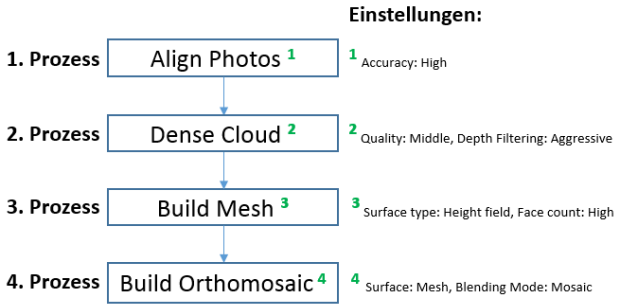


Abb. 3: Schritte der Prozessierung in Agisoft PhotoScan mit Einstellungen für den jeweiligen Prozess

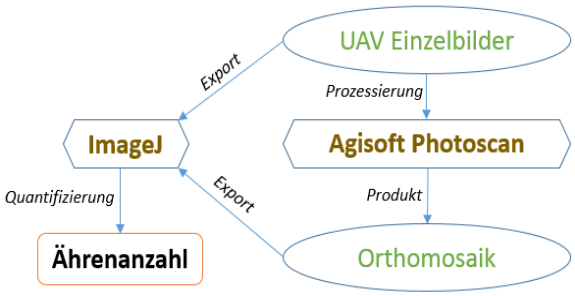


Abb. 4: Schematische Übersicht der Datenverarbeitung zur Erhebung der Ähren-anzahl mit Einstellungen für den jeweiligen Prozess

Mit Hilfe des Orthomosaiks können Bestandsparameter wie die Ährenanzahl unabhängig von Flughöhe und Einzelbildern auf einer beliebig großen Fläche quantifiziert und lokalisiert werden (TURNER et al. 2012).

Für die Analyse wurden die Ähren im vordefinierten Quadratmeter für die UAV-Bilddaten (Einzelbild, Orthomosaik) in ImageJ manuell markiert und gezählt (Abb. 4). Ausschlaggebendes Kriterium der Ährenerkennung ist die farbliche Absetzung der Ähren vom restlichen Teil der Pflanze.

3 Ergebnisse und Diskussion

3.1 Optimaler UAV-Aufnahmezeitpunkt

Zur Ermittlung des optimalen UAV-Aufnahmeszeitpunktes wurde im wöchentlichen Abstand die Ährenanzahl exemplarisch an einer Versuchsparzelle im UAV-Einzelbild gezählt. Zusammengefasst ist das Ergebnis in Abb. 5. Im Vorhinein wurde von der Vermutung ausgegangen, dass sich das Entwicklungsstadium der Frucht und Samenreife (BBCH Makrostadium 8) am idealsten für die Bonitur mittels UAV-Bildaufnahme eignet. Die oberen Ähren befinden sich kurz vor dem abreifen, die restliche Pflanze ist zum Teil noch grün und die Ähre setzt sich farblich vermeintlich am besten in den Aufnahmen ab (Abb. 1).

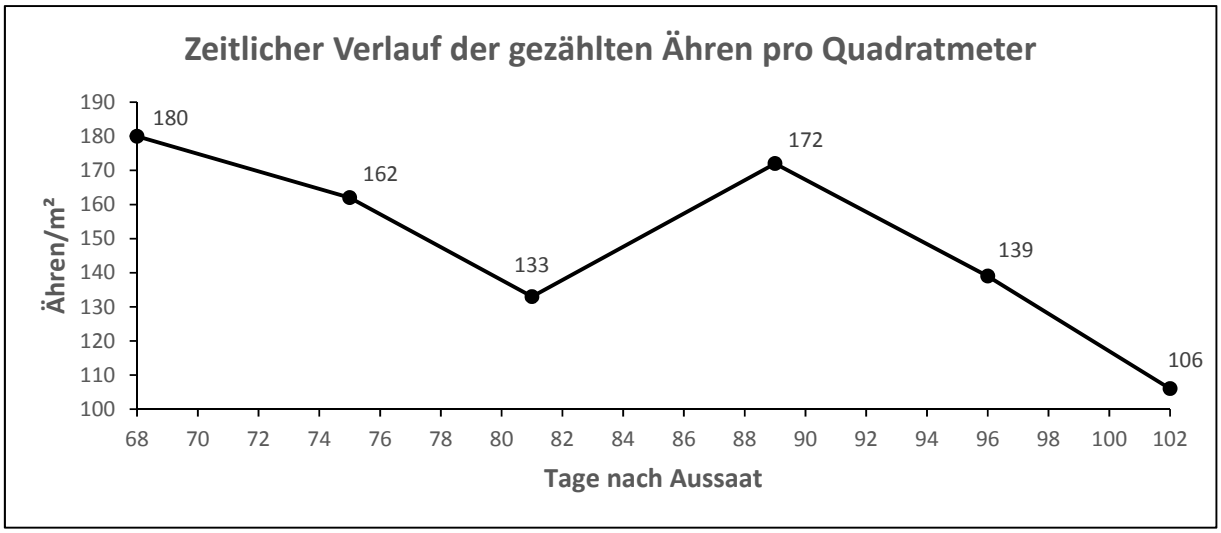


Abb. 5: UAV Ährenzählung im definierten Quadratmeter einer exemplarischen Versuchsparzelle in einem zeitlichen Verlauf von 68 bis 102 TnA zur Ermittlung des optimalen Aufnahmezeitpunktes

Die Auswertung zu Beginn des Entwicklungsstadiums 68 und 75 TnA ist stark von der Ährenposition abhängig. Für eine quantitative Auswertung müssen Bilder verwendet werden, auf denen die Ähren in einer Windrichtung positioniert sind (Abb. 6 A).

Bei einer Aufnahme ohne Neigung der Ähren, wie in Abb. 6 (B) dargestellt, können deutlich weniger Ähren zum gleichen Zeitpunkt identifiziert werden. Mit den richtigen Voraussetzungen kann beim ersten Termin 68 TnA (BBCH 61) mit 180 Ähren/m² die höchste Ährenanzahl bonitiert werden (Abb. 5). Im BBCH Stadium 61 sind die Ähren erstmalig vollständig sichtbar und die Grannen noch nicht sehr weit entwickelt.

Die Ähren setzen sich farblich deutlich von der Gesamtpflanze ab und können gut identifiziert werden (Abb. 6 A). An den beiden fortlaufenden Zählterminen (75 und 81 TnA) überdecken die Grannen zunehmend den entscheidenden hellen Bereich der Ähre, sodass eine Identifizierung erschwert wird. Zusätzlich verringert sich der Farbkontrast zwischen Ähre und Pflanze. Als Konsequenz können weniger Ähren im Einzelbild identifiziert werden (Abb. 5).

Abb. 6: Ergebnis der UAV-Ährenzählung einer Versuchsparzelle 68 TnA abhängig von der korrekten Aufnahme- und Ährenposition

Entgegen dieses Trends wird 89 TnA (BBCH 85) mit 172 Ähren/m² die zweithöchste Anzahl von Ähren erfasst. Die oberen Ähren sind überwiegend abgereift und beginnen sich zu senken (Abb. 7). Nachzügler, die sich in den unteren Schichten des Bestandes befanden und noch grün waren, wurden sichtbar und konnten ebenfalls erfasst werden. Mit fortlaufender Entwicklung (Abreifung der Pflanze inklusive Ähre) vermindert sich dieser farbliche Kontrast und die Identifizierung der Ähren wird erneut deutlich schwieriger (Abb. 7 B).
Die Ährenanzahl nimmt in den beiden nachfolgenden Messungen folglich wieder ab (Abb. 4).

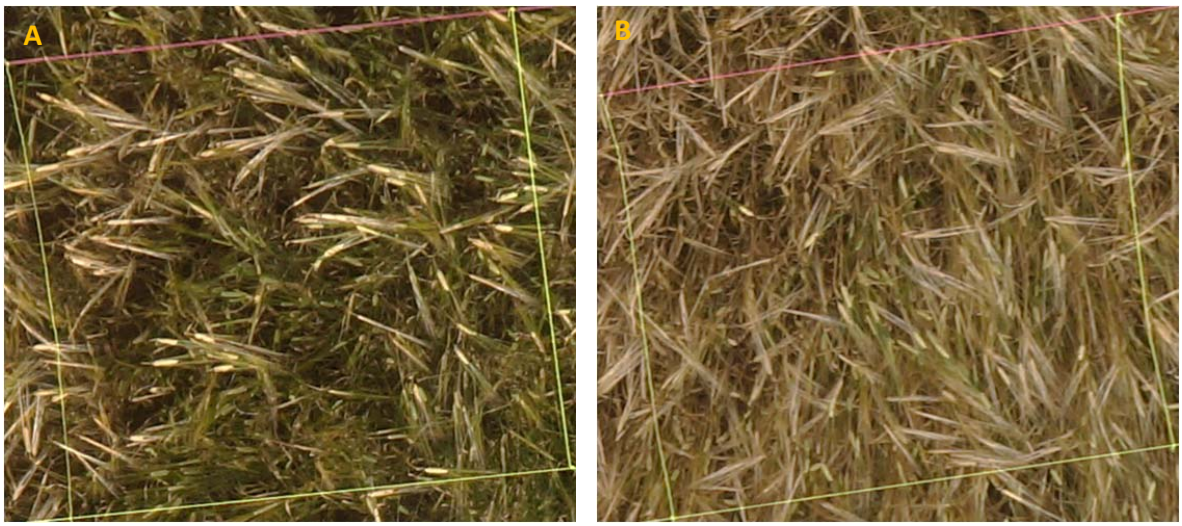


Abb. 7: Erfassung der Ährenzahl 89 TnA (A) und 96 TnA (B) im definierten Bereich des Quadratmeters über die UAV-Aufnahmen aus 10 Metern Flughöhe

Die Ergebnisse zeigen, dass der Termin direkt zu Beginn des Entwicklungsstadiums 68 TnA mit 180 Ähren/m² der vermeintlich beste Zeitpunkt für die Erfassung der Ähren mittels UAV-Bildaufnahmen ist. Beeinflusst wird dies durch den guten Farbkontrast, durch den sich die Ähre sehr deutlich vom Rest der Pflanze abhebt (Abb. 6 A). Die Ähre kann mit hoher Genauigkeit erfasst werden. Das Ergebnis wird jedoch von der Neigung der Ähren durch externe Faktoren wie Wind beeinflusst, da die Ähren zur Blüte (BBCH 61) nach oben orientiert sind und durch Fernerkundungsdaten aus der Nadir Perspektive (90° Kameraausrichtung) erschwert erfasst werden können.
Zum Zeitpunkt der Datenerfasssung in der Studie (89 TnA) wird lediglich die zweithöchste Anzahl von Ähren bestimmt. Durch die verringerten Farbkontraste können die oberen Ähren im Vergleich zum Zeitpunkt 68 TnA schwieriger erfasst werden. Verstärkt wird die Problematik durch die entwickelten Grannen und die Senkung der Ähren, durch die weitere Ähren durch auftretende Überlappungen schwieriger identifiziert werden können. Wie bereits zuvor vermutet können zum Zeitpunkt 89 TnA jedoch Nachzügler oder tiefere Ähren in der Zählung berücksichtigt werden (Abb. 7 A). Durch die natürliche Neigung der Ähren im fortgeschrittenen Entwicklungsstadium ist die Quantifizierung der Ähren ebenfalls unabhängiger von externen Faktoren. Schlussfolgernd ist der Aufnahmezeitpunkt 89 TnA mit einer marginalen Abweichung von fünf Prozent im Vergleich zum maximal Wert ebenfalls gut geeignet um die Ähren zu quantifizieren.

3.2 UAV Ährenanzahl und Validierung

Mit den UAV-Aufnahmen können größtenteils nur die Ähren im oberen Kronendach bonitiert werden und nur eingeschränkt Ähren, die sich in weiter untenliegenden Schichten des Bestandes befinden. Verglichen mit der Referenzmessung im Feld wird somit im Einzelbild unabhängig von der Saatdichte mindestens jede zweite Ähre im definierten Bereich des Quadratmeters erfasst (Abb. 8). Verglichen mit der niedrigen Saatdichte kann in der normalen Saatdichte eine erhöhte Ährenanzahl im UAV-Einzelbild identifiziert werden.

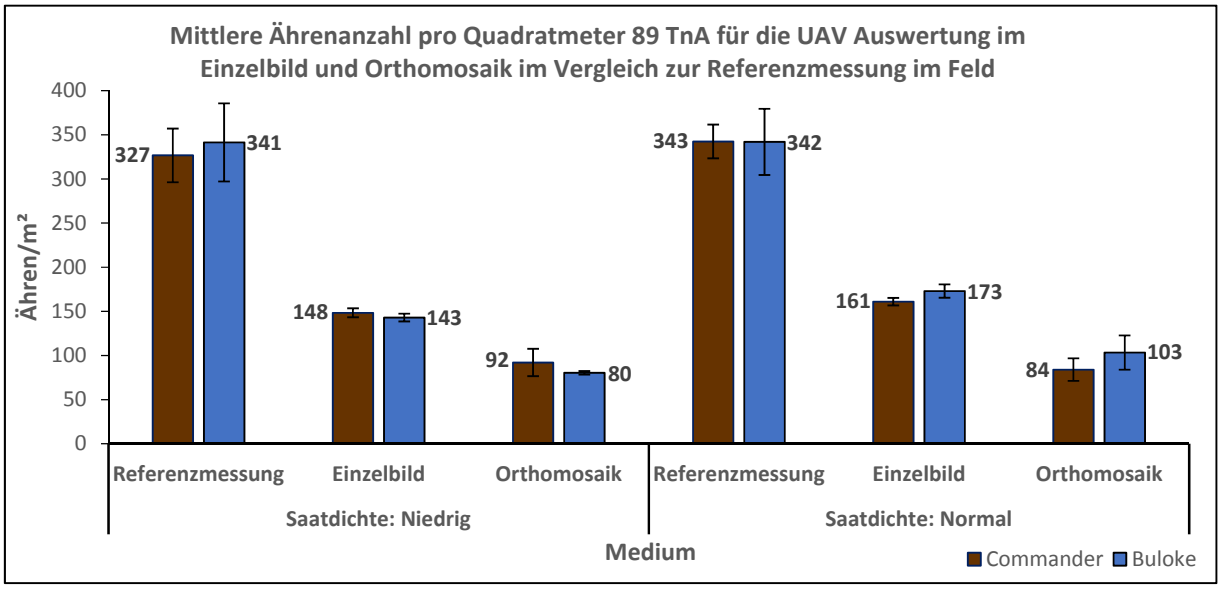


Abb. 8: UAV Ergebnis der mittleren Ährenanzahl pro Quadratmeter und Vergleich der UAV-Auswertung im Einzelbild und Orthomosaik mit der Referenzmessung im Feld 89 TnA (n = 3)

Im Orthomosaik kann lediglich etwas mehr als ein Viertel der Ähren im Vergleich zur Referenzmessung im Feld identifiziert werden (Abb. 8).
Bei der Analyse der Ergebnisse muss berücksichtigt werden, dass die Referenzmessung im Feld ebenfalls fehlerbehaftet sein kann. Um die Erfassung der Ähren zu vereinfachen und mögliche Fehler zu vermeiden, sollte in zukünftige Studien der verwendete Quadratmeter in mehrere Raster von beispielsweise 10*10 Zentimetern unterteilt werden. Durch die Definition eines kleineren Bereiches wird die Wahrscheinlichkeit verringert, dass Ähren im Feld doppelt gezählt werden oder unberücksichtigt bleiben. Diese Fehlerquelle kann in der UAV-Auswertung durch die farbliche Markierung der Ähren ausgeschlossen werden.
Der Vergleich zwischen der Referenzmessung im Feld und der Zählung im UAV-Einzelbild zeigt deutlich, dass durch auftretende Überlappungen und durch unterschiedliche Höhenebenen der Ähren (Abb. 9), nur ein Teil der Ähren mit Hilfe einer RGB Kamera identifiziert werden können. Anders als bei der Untersuchung von KOVACHEVA et al. (2014) ist es in der vorliegenden Studie durch die Nadir Perspektive nicht möglich, jede Ähre zu erfassen. Im Vergleich zu COINTAULT et al. (2008) trägt die limitierende Auflösung und der deutlich höhere Abstand zwischen Bestand und RGB Aufnahme von 10 Metern (zwölffacher Abstand) dazu bei, dass deutlich weniger Ähren erfasst werden konnten. Durch die höhere Distanz muss bei der Auswertung zusätzlich die korrekte Ährenposition berücksichtigt werden, wie im vorherigen Kapitel bereits erläutert wurde.

Abb. 9: Hohe Sprosszahl und unterschiedliche Höhenebenen der Ähren erschweren die Quantifizierung der Ähren mittels UAV-Aufnahmen

Bei der Analyse der Diskrepanz zwischen Referenzmessung und Einzelbild war auffällig, dass die Werte hinsichtlich der Genotypen und Saatdichten konstant zwischen 50% und 56% abweichen und die Standardabweichung (± 3%) gering ist (Abb. 8).

Wird dieser Wert von zukünftigen Untersuchungen bestätigt, könnte die aus den UAV-Bildern ermittelte Ährenanzahl mit einem Faktor (von ca. zwei) multipliziert werden, um eine genauere Erfassung der Ährenanzahl zu ermöglichen.
Die Resultate für das Orthomosaik sind stark davon abhängig, wie gut die Oberfläche anhand der UAV-Einzelbilder rekonstruiert werden kann. Durch die geringe Flughöhe von 10 Metern ist die Fläche reduziert, die von einem Einzelbild abgedeckt werden kann. Die Bildanzahl für die notwendige Überlappung und die Komplexität das Orthomosaik zu berechnen, nimmt dadurch deutlich zu (GÓMEZ-CANDÓN et al. 2014). Zusätzlich sind die Ähren durch Windeinflüsse auch bei niedrigen Windgeschwindigkeiten oftmals in Bewegung. In unterschiedlichen Aufnahmen befinden sich die Ähren somit nicht zwangsläufig in den gleichen Positionen. Für die Rekonstruktion des Orthomosaiks ist dies jedoch eine Voraussetzung, da die berechnete 3D Oberfläche durch die Bewegung der Ähren unter Umständen bereits fehlerhaft sein kann. Dies führt zu unscharfen Bereichen und Artefakten im Orthomosaik, die die Erfassung einer einzelnen Ähre erschweren und in einer Abweichung im Vergleich zum Einzelbild resultiert. Zur Veranschaulichung der Problematik ist in Abb. 10 (A) ein Bereich eines Orthomosaiks mit geringer Rekonstruktionsgüte visualisiert. Dem gegenüber steht ein Ausschnitt mit einer qualitativ hochwertigen Rekonstruktion (Abb. 10 B).

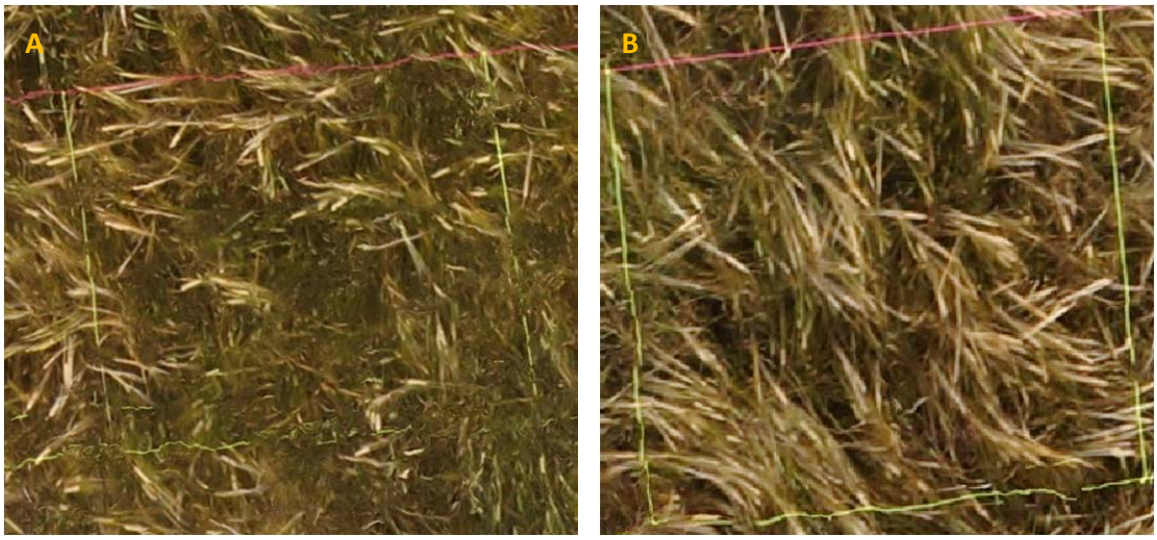


Abb. 10: Auswirkung von externen Faktoren wie Wind auf die korrekte Rekonstruktion im Orthomosaik mit schlechter Rekonstruktionsgüte (A) in Gegenüberstellung mit einer qualitativ hochwertigen Rekonstruktion (B)

4 Fazit & Ausblick

Die vorliegende Studie hat resümierend aufgezeigt, dass die Quantifizierung der Ähren über UAV RGB-Aufnahmen eingeschränkt möglich ist. Für zukünftige Studien können die ermittelten Resultate berücksichtigt werden, um das Verfahren zu optimieren. Letztendlich kann mittels UAV jedoch lediglich das Kronendach eines Bestandes bonitiert werden. Dementsprechend könnte die Ährenanzahl, die über die UAV-Bilddaten erfasst werden, mit einem entsprechenden Parameter (Ährenzahl im Kronendach) neu definiert werden.

Alternativ sollte das Verfahren bei einer höheren Saatdichte angewandt werden. Durch die geringe Saatdichte und die erhöhte Sprossanzahl pro Pflanze wird die Problematik der unterschiedlichen Höhenebenen der Ähren verstärkt (Abb. 8). Je höher die Saatdichte, desto geringer die Sprossanzahl und desto weniger Ähren befinden sich in den tieferen Bereichen, die über die UAV-Aufnahmen nicht erfasst werden können. Die erhöhte Ährenanzahl, die im UAV-Einzelbild für die normale Saatdichte detektiert werden konnten, unterstreicht dies. Zusätzlich könnte eine Identifizierung der Ähren bei Genotypen mit mehrzeiligen sowie größeren Ähren mit weniger dominanten Grannen erleichtert werden.
Unter Berücksichtigung der Resultate des optimalen Aufnahmezeitpunktes könnten die UAV-Ergebnisse verbessert werden, wenn die Quantifizierung bereits direkt zu Beginn des Entwicklungsstadiums (BBCH 61) durchgeführt wird. Der deutliche Farbkontrast zwischen Ähren und Pflanzen zum genannten Zeitpunkt könnte ebenfalls dazu verwendet werden, eine automatisierte Erkennung zu entwickeln. Durch die Erfassung der Ähren über die Farbinformationen reduziert sich unter Umständen die erforderliche Auflösung. Die Flughöhe von zuvor 10 Metern könnte erhöht und die praxisnähe dadurch zusätzlich gesteigert werden. Untersucht werden muss dabei jedoch, inwiefern die Aufnahme- und Ährenposition das Verfahren im genannten Entwicklungsstadium beeinflussen. Alternativ könnte bei der Datenaufnahme die Kameraausrichtung von 90° (Nadir) auf beispielsweise 45° geändert werden, um die dargelegte Problematik zu vermeiden. Zusätzlich besteht die Möglichkeit, die Ähren mit Hilfe einer Multispektralkamera zu erfassen. Die multispektrale Information im NIR-Bereich erleichtert unter Umständen die Separierung von Pflanze und Ähre und könnte die Ergebnisse optimieren.
Resümierend kann die Ährenanzahl im Vergleich zum Orthomosaik im Einzelbild genauer erfasst werden. Das Orthomosaik ist ein Produkt aus überlappenden Bildern aufgenommen zu verschiedenen Zeitpunkten, wohingegen das Einzelbild eine Momentaufnahme ist. Die Entwicklung einer Methode, durch die die Flughöhe zukünftig gesteigert werden kann, vergrößert die durch ein Einzelbild abgedeckte Fläche und würde die notwendige Bildüberlappung verringern. Mögliche Fehlerquellen, die zu unscharfen Bereichen und dadurch zu einer erhöhten Abweichung führen, können dementsprechend reduziert werden.

5 Literaturverzeichnis

Atzberger, C., 2013: Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. Remote Sensing, **5**(2), 949-981.

Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S. & Bareth, G., 2014: Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-based RGB Imaging. Remote Sensing, **6**(11), 10395-10412.

Burkart, A., Aasen, H., Alonso, L., Menz, G., Bareth, G. & Rascher, U., 2015: Angular Dependency of Hyperspectral Measurements over Wheat Characterized by a Novel UAV Based Goniometer. Remote Sensing, **7**(1), 725-746.

Candiago, S., Remondino, F., De Giglio, M., Dubbini, M. & Gattelli, M., 2015: Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images. Remote Sensing, **7**(4), 4026-4047.

Cointault, F., Guerin, D., Guillemin, J□P. & Chopinet, B., 2008: In-field Triticum aestivum ear counting using colour-texture image analysis. New Zealand Journal of Crop and Horticultural Science, **36**(2), 117-130.

Eling, C., Wieland, M., Hess, C., Klingbeil, L. & Kuhlmann, H., 2015: Development and Evaluation of a UAV based Mapping System for Remote Sensing and Surveying Applications. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, **40**(1), 233-239.

Gómez-Candón, D., De Castro, A. I. & López-Granados, F., 2014: Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. Precision Agriculture, **15**, 44-56.

Grenzdörffer, G. J., Engel, A. & Teichert, B., 2008: The photogrammetric potential of low-cost UAVs in forestry and agriculture. International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, **37**, 1207-1213.

Kersten, T. P., 2016: Auf Knopfdruck – Auswertung und Resultate UAV-gestützter Bildflugdaten. UAV 2016 - Vermessung mit unbemannten Flugsystemen, Schriftenreihe des DVW, **82**, 135-158.

Khot, L. R., Sankaran, S., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J. & Miklas, P. N., 2015: Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. European Journal of Agronomy, **70**, 112-123.

Kovacheva, I., Kerssen, S., Mentrup, D., Rosemann, T., Bilges, K. & Ruckelshausen, A., 2014: Bildgebende Lichtschrankensensoren zur Detektion von Pflanzenhöhen, Ähren und Grannen. Bornimer Agrartechnische Berichte, **88**, 179-190.

Lelong, C. C. D., Burger, P., Jubelin, G., Roux, B., Labbé, S. & Baret, F., 2008: Assessment of Unmanned Aerial Vehicles Imagery for Quantitative Monitoring of Wheat Crop in Small Plots. Sensors. **8**, 3557-3585.

Liu, H. Y., Yang, G. J. & Zhu, H. C., 2014: The Extraction of Wheat Lodging Area in UAV's Image Used Spectral and Texture Features. Applied Mechanics and Materials, 651-653, 2390-239.

Lowe, D. G. 2004: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, **60**(2), 91-110.

Sankaran, S., Khot, L. R. & Carter, A. H., 2015: Field-based crop phenotyping: Multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. Computers and Electronics in Agriculture, **118**, 372-379.

Torres-Sánchez, J., Peña, J. M., De Castro, A. I. & López-Granados, F., 2014: Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. Computers and Electronics in Agriculture, **103**, 104-113.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J. & Reynolds, J. M., 2012: Structure- from- Motion photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, **179**, 300-314.

Turner, D., Lucieer, A., & Watson, C., 2012: An automated technique for generating georectified mosaics from ultra-high resolution Unmanned Aerial Vehicle (UAV) imagery, based on Structure from Motion (SFM) point clouds. Remote Sensing, **4**(5), 1392-1410.

Introducing a New Concept for Grassland Monitoring: The Multi-temporal Grassland Index (MtGI)

ULRIKE LUSSEM[1] & GEORG BARETH[1]

*Abstract: The monitoring of forage mass in managed grasslands is of key importance because it equals yield. The spatially explicit determination of removed forage mass in multi-cut grasslands is a non-solved problem and the knowledge would enable improved fertilizer management. Therefore, the objective of this contribution is the development of a remote sensing-based method to derive forage mass after grassland is cut. For this purpose, we propose a new concept for vegetation monitoring which is based on the multi-temporal acquisition of super high resolution UAV-based RGB-imagery. From the acquired data it is possible to derive the RGB-Vegetation Index (RGBVI) (Bendig et al. 2015) and sward height (Bareth et al. 2015). The latter is based on the concept of Crop Surface Models (CSMs) which is based on multiple UAV campaigns during phenology (Hoffmeister et al. 2010). The study was carried out at the long-term Rengen Grassland Experiment (RGE) in Germany in 2014. The RGE is conducted by the INRES (*Institute of Crop Science and Resource Conservation*) of the University of Bonn, Germany (Schellberg et al. 1999). From the combination of the UAV-derived data sets, sward height and RGBVI, the computation of the Grassland Index (*Grass*I) is possible (Bareth et al. 2015). In contrast, the newly proposed concept for grassland monitoring is based on two UAV-campaigns per growth: direct before (t_1) and direct after (t_2) cut. From the two time windows, the Multi-temporal Grassland Index (MtGI) is calculated. The MtGI is a NDVI-like vegetation index but includes temporal change in its equation. Applied on remote sensing data, it can be considered as a spatio-temporal vegetation index. The MtGI is highly correlated (R^2 = 0.74) to destructively measured forage mass and therefore, bears a promising potential for non-destructive and automated forage mass mapping in managed grasslands with low-cost UAVs.*

1 Introduction

Monitoring forage mass on managed/multi-cut grasslands, such as extensive or intensive pastures or hay and silage meadows, is an important step towards optimising yield and nutrient balancing/budgeting. Deriving spatially high resolution information about removed forage mass with each cut is a non-solved problem and the knowledge would enable improved fertilizer management.

Site specific fertilizer application as part of Precision Agriculture (PA) can aid in the decision making process (SCHELLBERG et al. 2008). However, manual measurements of forage mass, such as clipping or disc meters/rising plate meters and spectroradiometer measurements cannot meet the requirements to produce a high resolution assessment of the within-field heterogeneity of standing biomass and nutrient status (BARETH et al. 2015; HOLLBERG & SCHELLBERG 2017).

Therefore, the objective of this contribution is to further investigate the potential of RGB imaging with low-cost UAVs for managed grasslands as introduced by BARETH et al. (2015). For

[1] University of Cologne, Institute of Geography, GIS & RS Group, 50923 Cologne, Germany, E-Mail: [ulrike.lussem,g.bareth]@uni-koeln.de

this purpose, we propose a new concept for vegetation monitoring in general which is based on the multi-temporal acquisition of super high resolution UAV-based RGB-imagery. From the acquired data it is possible to derive the RGB-Vegetation Index (RGBVI) (BENDIG et al. 2015) and sward height (BARETH et al. 2015). Sward height can be spatially derived by applying the concept of Crop Surface Models (CSMs) which is based on multiple UAV campaigns during distinct phenological stages (HOFFMEISTER et al. 2010) on grasslands. From the combination of the UAV-derived data sets, sward height and RGBVI, the computation of the Grassland Index (GrassI) is possible (BARETH et al. 2015). The focus of this contribution is the analysis of RGB image data after grass is cut and knowledge on harvested forage mass is demanded for managing purposes like improving fertilization schemes.

2 Study Area and Methods

2.1 The Rengen Grassland Experiment (RGE)

The study was conducted in 2014 on the Rengen Grassland Experiment (RGE) managed by the INRES (Institute of Crop Science and Resource Conservation) of the University of Bonn. The RGE is a long-term experiment established in 1941 in the Eifel Mountains in Germany (Rhineland-Palatinate, Germany, 50°13`N, 6°51`E). The design, floristic composition, soil properties and research foci are extensively described in previous publications (CHYTRÝ et al. 2009; HEJCMAN et al. 2007; SCHELLBERG et. al. 1999).

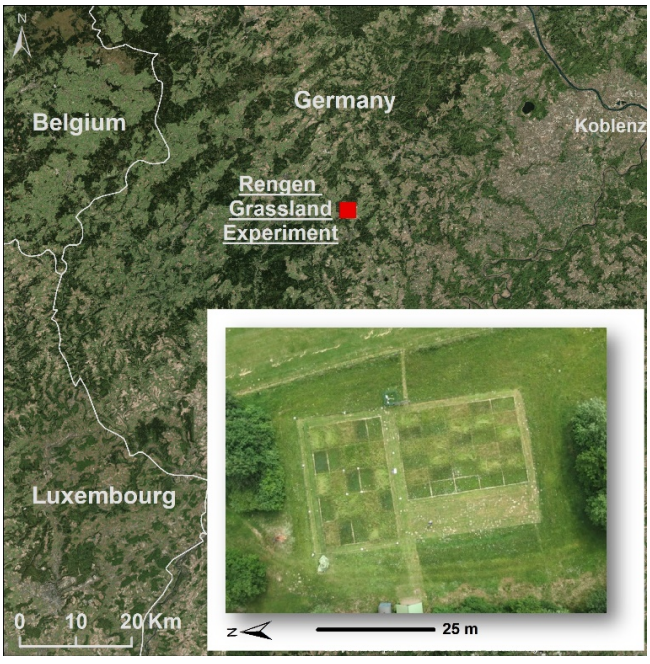


Fig. 1: Overview of the study area: The Rengen Grassland Experiment (data source: ESRI-Basemaps, Eurostat; Photo: G. Bareth)

The RGE is arranged in a randomized block design of five fertilizer treatments (Ca, CaN, CaNP, $CaNPK_2O$, $CaNPK_2SO_4$) with ten replicates per treatment on a plot size of 3×5 meter. The sward is cut twice per growing season, in early July and in mid-October. Due to the constant fertilizer application a distinct floristic composition per treatment and plot has evolved (CHYTRÝ et al. 2009).
For this study, one block was chosen for investigation (five replicates per treatment resulting in 25 plots). The pre-cut sampling date (t1) was on June 17th 2014 and the corresponding post-cut sampling date (t2) on July 6th 2014. From data of both sampling dates the RGBVI and the sward height (SH) was derived. Forage mass sampling was conducted in between on July 2nd 2014. Subsamples of forage mass from 6.5 m^2 of the 15 m^2 plots were taken and dry matter forage mass in kg per hectare was calculated. Additionally, rising plate meter (RPM) measurements were carried out with ten evenly distributed measurements per plot on June 23rd 2014. Forage mass sampling and RPM measurements were conducted by the INRES (HOLLBERG & SCHELLBERG 2017; BARETH et al. 2015).

2.2 The low-cost UAV system

In parallel to the RPM sampling campaigns image data was collected with a low-cost unmanned aerial vehicle (UAV), a DJI Phantom 2. The UAV was equipped with a consumer-grade RGB camera (Canon Powershot S110 with 12 MP) attached in nadir viewing position on a fixed camera mount being aware of image distortion due to UAV movements (Fig.2). The camera was equipped with a Canon Hack Development Kit (CHDK) to allow for continuous image acquisition. Shutter speed was set to 1/2000 with fixed aperture and ISO settings. For this analysis, we used the image data of June 17th because lodging already occurred on June 23rd.

Fig. 2: DJI Phantom 2 equipped with Canon Powershot 110 (BARETH et al. 2015)

The System was operated manually with a random flight pattern in about 20 m above ground level (AGL) to capture multiple overlapping images for Structure-from-Motion (SfM) processing with Agisoft Photoscan. Additionally, one overview image from about 40 m AGL was captured to minimize illumination effects when computing RGB-based vegetation indices for the whole area of interest. Evenly distributed ground control points for georeferencing were measured using a RTK-GPS.

2.3 Crop Surface Models (CSMs)

The concept of multitemporal crop surface models (CSM) was introduced by HOFFMEISTER et al. (2010) to derive plant height with terrestrial laser scanning. CSMs represent the increase in elevation due to plant growth of a certain crop in distinct phenological stages (Fig.3). To derive absolute plant height metrics, a DEM acquired at the beginning of the growing season shortly after sowing (t_0) is subtracted from the CSM of each succeeding sampling date (t_1, t_2, t_n) in a GIS environment. CSMs derived from UAV-based imagery are proofed to be a reliable estimator for biomass in barley as shown by BENDIG et al. (2014). Recently, the concept of CSMs was also successfully transferred to experimental grassland-sites to estimate sward height and forage mass (BARETH et al. 2015; POSSOCH et al. 2016; LUSSEM et al. 2017b).

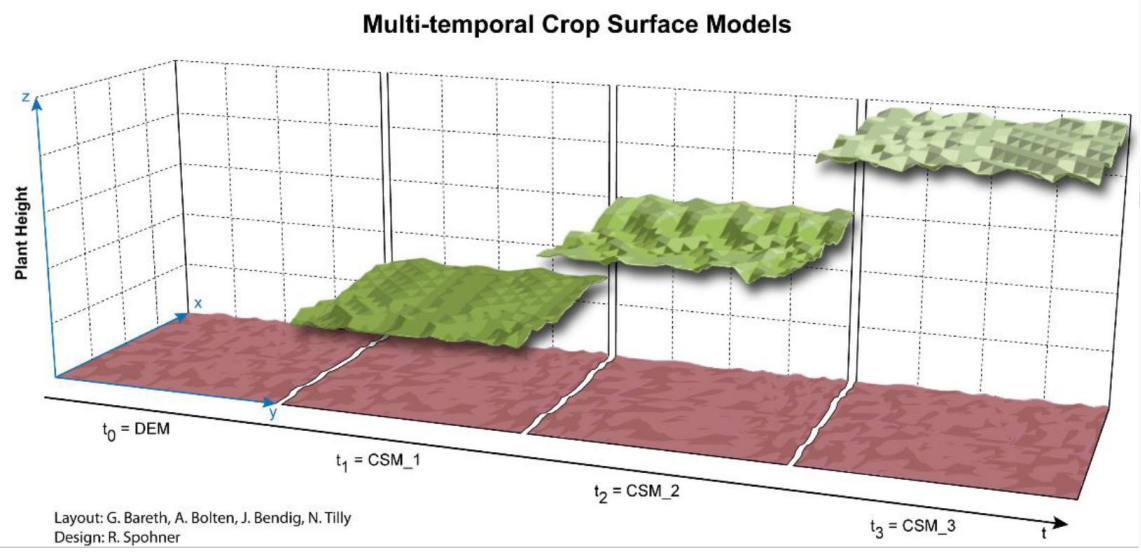


Fig. 3: Multi-temporal Crop Surfac Models (CSMs) (BENDIG et al. 2013)

2.4 Vegetation indices using the VIS domain

Vegetation indices (VIs) calculated from the visible (VIS) domain of the electromagnetic spectrum can be derived in high spatial resolution from UAV-based RGB images to assess vegetation-parameters, since they are sensitive to subtle changes in greenness by using the chlorophyll-absorption features in the red and blue part of the spectrum (HUNT et al. 2013). VIS-VIs have shown good results in estimating chlorophyll content (HUNT et al. 2011, 2013), biomass and LAI (GITELSON et al. 2003) or vegetation fraction (GITELSON et al. 2002). A recently introduced VIS-VI, the Red-Green-Blue Vegetation Index (RGBVI), performed moderately well to estimate biomass in barley (BENDIG et al. 2015).

$$RGBVI = ((R_G)^2 - (R_B * R_R)) / ((R_G)^2 + (R_B * R_R)) \quad (1)$$

with reflectance of the red (R), green (G) and blue (B) camera band.

Compared to the well-known NDVI the RGBVI showed less saturation effects in later growing stages for heterogenous grassland communities (LUSSEM et al. 2017a) and reached good correlations in grasslands (BARETH et al. 2015). Similar observations in terms of saturation effects for VIS-VIs are also shown by HUNT et al. (2013).

2.5 The multi-temporal Grassland Index (MtGI)

BARETH et al. (2015) introduced the Grassland Index (*Grass*I) as a combination of SH from CSMs and RGBVI (see equation 2) for each sampling date to differentiate the five RGE fertilization managements.

$$\textit{Grass}\text{I} = \text{SH} + \text{RGBVI} \quad (2)$$

In contrast to the *Grass*I, the index proposed for grassland monitoring in this study, the Multi-temporal Grassland Index (MtGI), has a NDVI-like structure, but includes temporal changes in its equation.

$$\text{MtGI} = ((SH_{t1} + RGBVI_{t1}) - (SH_{t2} + RGBVI_{t2})) / ((SH_{t1} + RGBVI_{t1}) + (SH_{t2} + RGBVI_{t2})) \quad (3)$$

with sward height from crop surface models (SH), pre-cut (t_1) and post-cut (t_2) sampling date, and RGBVI see equation (1)

As can be seen in Fig. 4, the relationship between forage yield and appearance of greenness is inverse from t1 to t2. Plots with high fertilizer levels and therefore high sward height and forage yield (CaNP, CaNPKCl, $CaNPK_2SO_4$) appear less green in t2. In contrast, lower fertilizer levels with low sward heights (Ca, CaN) appear greener after the cut than the aforementioned plots. Based on these patterns, the MtGI was developed to incorporate the changes of structural and physiological information over time from RGBVI and CSM-based SH data.

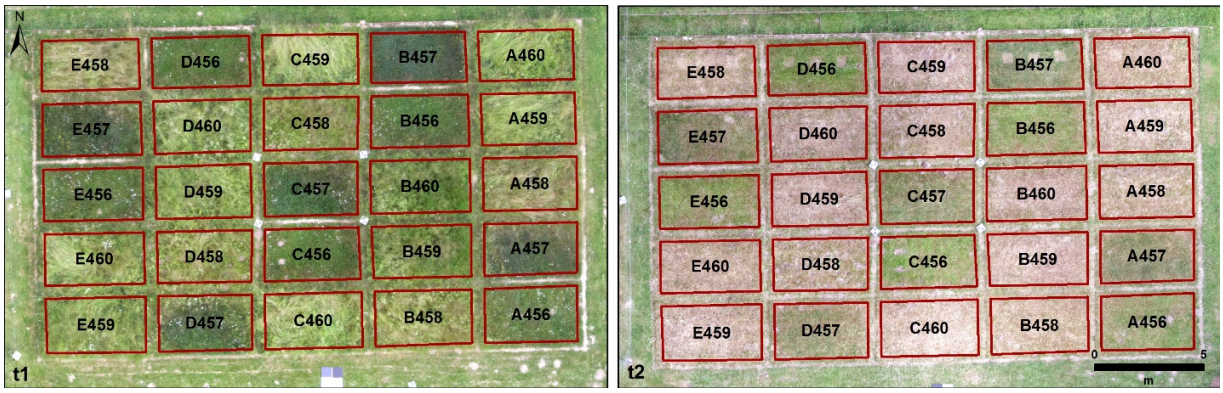


Fig. 4: t1: pre-cut (17.06.2014) & t2: post-cut (06.07.2014); capital letters indicate rows; treatment abbreviations: 456 (Ca); 457 (CaN); 458 (CaNP); 459 (CaNPKCl); 460 ($CaNPK_2SO_4$)

The RGBVI was calculated per pixel (see equation (1)) in a spatial resolution of 2 cm from the overview images of the RGE field experiment. For both sampling dates, digital surface models

were computed in the Structure from Motion (SfM) software Agisoft PhotoScan to derive sward height from CSMs (see Fig.3) (BENDIG et al. 2013; BENDIG et al. 2015).
Analysis of the UAV-based data as estimators of forage mass was conducted on the plot level. Therefore, the values of SH-CSM and RGBVI were averaged per plot using the analysis tool *Zonal Statistics as Table* in ArcGIS. Statistical analysis was computed in MS Excel with linear regression models for the above-mentioned estimators (SH-CSM, RGBVI, RPM, MtGI) of forage mass including the coefficient of determination (R^2).

3 Results

The spatial resolution of the CSMs for t1 and t2 and for the overview images of the RGE field experiment to calculate the RGBVI is 2 cm. The scatterplot in Fig. 5 (left) shows a strong correlation between the CSM-derived sward height (SH) and rising plate meter measurements (RPM) for compressed SH with a R^2 of 0.77. The RPM measurements are also averaged per plot from the ten individual measurements. However, the mean SH from CSMs per plot is higher compared to the values of the RPM. Fig. 5 (right) emphasis the strong correlation between forage mass and RPM measurements (R^2 = 0.85). The latter is well known in grassland sciences and proves the applicability of the method for the RGE.

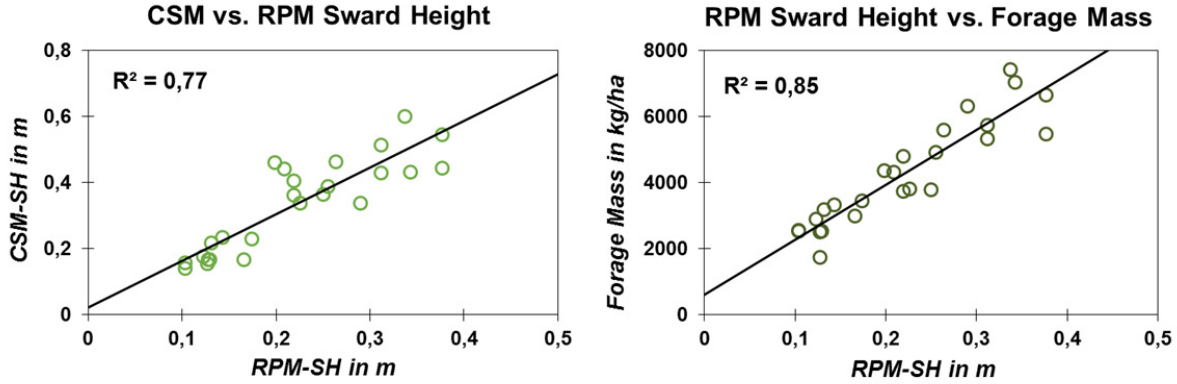


Fig. 5: left: Sward height (SH) from crop surface models (CSMs) in meter plotted against compressed SH from rising plate meter measurements (RPM) in meter for t1 (n=25). Right: Compressed SH from RPM measurements in meter plotted against dry matter forage mass in kg/ha (n=25)

In Fig. 6, the averaged plot data show a moderate correlation of the RGBVI for the pre-cut sampling date (t1, left) with forage mass and the trend is clearly visible. For the post-cut sampling date (t2, right), the t2 plot data with high forage mass yield are clustered around an RGBVI value of 0.19. The later indicates a kind of saturation of non-green values. However, a general trend is given resulting in a moderate to high R^2 of 0.7 which is higher than the pre-cut (t1) analysis.
The negative slope for t2 can be explained by the inversed greenness of the plots. Plots with high fertilizer input and high forage mass yield show higher values for the RGBVI (and higher SH) before the cut and lower values after the cut, because of the chlorophyll distribution and shading effects in taller grasses.

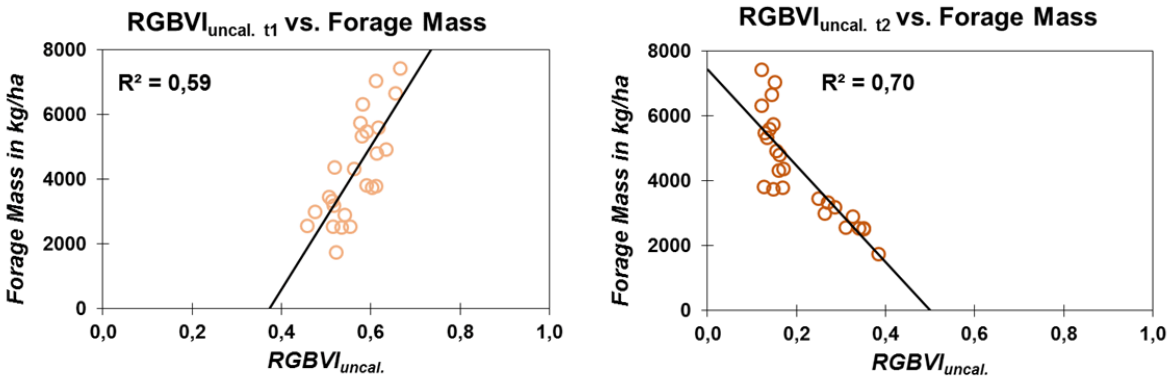


Fig. 6: left: scatterplot of RGBVI values for pre-cut (t1) against dry matter forage mass in kg/ha (n=25). right: scatterplot of RGBVI values for post-cut (t2) against dry matter forage mass in kg/ha (n=25)

In Fig. 7 (left), the averaged plot data of the Multi-temporal Grassland Index (MtGI) is plotted against forage mass (kg/ha) with a good correlation ($R^2 = 0.74$). However, the MtGI shows some severe saturation for higher forage yield, which is indicated by the clustering of data points between 0.6 and 0.8. A similar pattern can be observed when the MtGI is compared to the RPM measurements in Fig. 7 (right), although the R^2 value is slightly higher (0.78), the saturation is slightly weaker but still clearly visible.

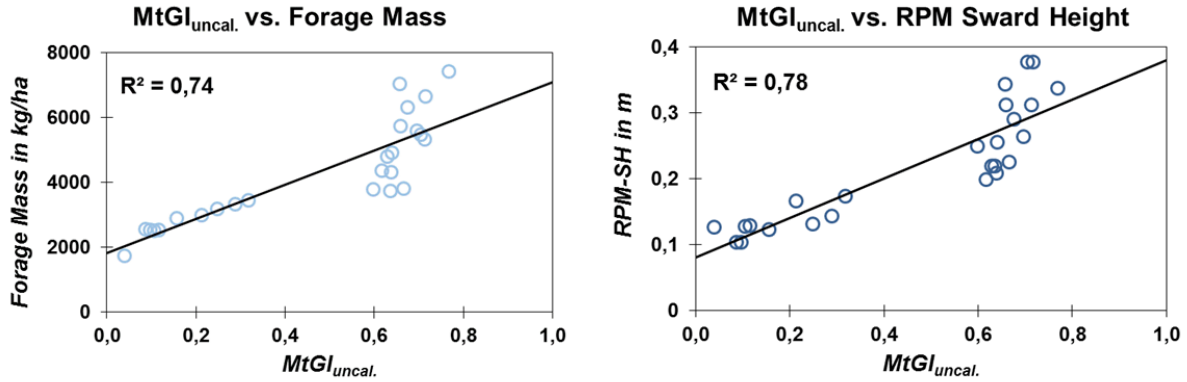


Fig. 7: left: scatterplot of MtGI values for pre-cut (t1) against dry matter forage mass in kg/ha (n=25). right: scatterplot of RGBVI values for post-cut (t2) against dry matter forage mass in kg/ha (n=25).

4 Discussion, Conclusion & Outlook

The newly introduced vegetation monitoring approach using a multi-temporal vegetation index, the Multi-temporal Grassland Index (MtGI), performed moderate to well. The innovative idea of the MtGI is the UAV-based grassland monitoring before (t1) and after (t2) cut with a low-cost system. The approach of remote sensing data acquisition after cut is rather unique and could lead to optimized grassland management strategies. However, the presented results are based only on one cutting date and consequently on one UAV data acquisition campaign before and one after

cut. The hypothesis of an inverse correlation pattern of the spectral data analysis (RGBVI) before and after cut is proofed in our contribution and bears a potential which should be further investigated. Especially when considering that the RGBVI data analysis is based on uncalibrated RGB images. In this context, Lussem et al. (2017a) proofed that empirical line corrected RGB images perform significantly better than uncalibrated (Bareth et al. 2015). Especially, the described saturation effects of the RGBVI were reduced which of course are also incorporated in the MtGI.

Compared to the results of the spectral analysis, the R^2 of the CSM-derived SH serving as a predictor for forage mass is significantly better, having a R^2 of 0.78, compared to the RGBVI ($R^2 = 0.59$). This fits well to other studies (Possoch et al. 2016; Bendig et al. 2015). However, while CSM-derived SH performs well to RPM data ($R^2 = 0.77$), RPM still is the strongest predictor for forage mass ($R^2 = 0.85$) in our study. Bareth et al. (2015) concluded that CSM-derived sward height correlated very well with RPM measurements considering data analysed for a whole growing season ($R^2 = 0.89$). This also supports the potential of further studies for multi-cut seasons. Finally, more sophisticated multivariate regression analysis should be investigated for the combination of CSM-derived SH with the RGBVI for grassland monitoring (Marshall and Thenkabail 2015; Bendig et al. 2015).

5 Acknowledgements

We sincerely thank Jürgen Schellberg (INRES, University of Bonn) for introducing us to the Rengen Grassland Experiment (RGE), for the very good cooperation, for supporting us in the last years, and for sharing his impressive expertise with us. Additionally, we thank Jens Hollberg for the RPM data of 2014 and the INRES for maintaining the RGE.

6 References

BARETH, G., BOLTEN, A. HOLLBERG, J., AASEN, H., BURKART, A. & SCHELLBERG, J., 2015: Feasibility study of using non-calibrated UAV-based RGB imagery for grassland monitoring: Case study at the Rengen Long-term Grassland Experiment (RGE), Germany. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 24, T. Kersten (Hrsg.), 55-62.

BENDIG, J., BOLTEN, A. & BARETH, G., 2013: UAV-based imaging for multi-temporal, very high resolution Crop Surface Models to monitor crop growth variability. Photogrammetrie, Fernerkundung, Geoinformation (PFG), **2013**(6), 551-562. doi: 10.1127/1432-8364/2013/0200.

BENDIG, J., YU, K., AASEN, H., BOLTEN, A., BENNERTZ, S., BROSCHEIT, J., GNYP, M.L. & BARETH, G., 2015: Combining UAV-based Crop Surface Models, Visible and Near Infrared Vegetation Indices for Biomass Monitoring in Barley. International Journal of applied Earth Observation and Geoinformation, **39**, 79-87. doi:10.1016/j.jag.2015.02.012.

CHYTRÝ, M., HEJCMANN, M., HENNEKENS, S. & SCHELLBERG, J., 2009: Changes in vegetation types and Ellenberg indicator values after 65 years of fertilizer application: evidence from

the Rengen Grassland Experiment, Germany. Applied Vegetation Science, **12**, 167-176. doi:10.1111/j.1654-109X.2009.01011.x.

GITELSON, A.A., KAUFMANN, Y.J., STARK, R. & RUNDQUIST, D., 2003: Novel algorithms for remote sensing of vegetation fraction. Remote Sensing of Environment, **80**, 76-87.

GITELSON, A.A., VINA, A., ARKEBAUER, T.J., RUNDQUIST, D.C., KEYDAN, G. & LEAVITT, B., 2003: Remote estimation of leaf area index and green leaf biomass in maize canopies. Geophysical Research Letters, **30**, 1248. doi:10.1029/2002GL016450.

HEJCMAN, M., KLAUDISOVA, M., SCHELLBERG, J. & HONSOVA D., 2007: The Rengen Grassland Experiment: Plant species composition after 64 years of fertilizer application. Agriculture, Ecosystems and Environment, 122, 259-266. doi:10.1016/j.agee.2006.12.036.

HOFFMEISTER, D., BOLTEN, A., CURDT, C., WALDHOFF, G. & BARETH, G., 2010: High resolution Crop Surface Models (CSM) and Crop Volume Models (CVM) on field level by terrestrial laserscanning. Proc. SPIE, **7840**, 78400E: 6 p. doi: 10.1117/12.872315.

HOLLBERG, J.L. & SCHELLBERG, J., 2017: Distinguishing Intensity Levels of Grassland Fertilization Using Vegetation Indices. Remote Sensing, **9**(1), 81. doi: 10.3390/rs9010081.

HUNT, E.R. JR., DAUGHTRY, C.S.T., EITEL, J.U.H. & LONG, D.S., 2011: Remote sensing leaf chlorophyll content using a visible band index. Agronomy Journal, **103**(4), 1090-1099.

HUNT, E.R. JR., DORAISSWAMY, P.C., MCMURTREY, J.E., DAUGHTRY, C.S.T., PERRY, E.M. & AKHMEDOV, B., 2013: A visible band index for remote sensing leaf chlorophyll content at the canopy scale. International Journal of applied Earth Observation and Geoinformation, **21**, 103-112.

LUSSEM, U., HOLLBERG, J., MENNE, J., SCHELLBERG, J. & BARETH, G., 2017a: Using calibrated RGB imagery from low-cost UAVs for grassland monitoring: Case study at the Rengen Grassland Experiment (RGE), Germany. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **42**(2/W6), 229-233, doi.org/10.5194/isprs-archives-XLII-2-W6-229-2017.

LUSSEM, U., BOLTEN, A., BARETH, G. & SCHELLBERG, J., 2017b: Feasibility study of directly georeferenced images from low-cost unmanned aerial vehicles for monitoring sward height in a longterm experiment on grassland. Grassland Science in Europe, **22**, 354-356.

MARSHALL, M. & THENKABAIL, P., 2015: Developing in situ non-destructive estimates of crop biomass to address issues of scale in remote sensing. Remote Sens., **7**, 808-835. doi: 10.3390/rs70100808.

POSSOCH, M., BIEKER, S., HOFFMEISTER, D., BOLTEN, A., SCHELLBERG, J. & BARETH, G. 2016: Multi-temporal crop surface models combined with the RGB vegetation index from UAV-based images for forage monitoring in grassland. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **41**(B1), 991-998. doi:10.5194/isprsarchives-XLI-B1-991-2016.

SCHELLBERG, J., HILL, M.J., GERHARDS, R., ROTHMUND, M. & BRAUN, M., 2008: Precision agriculture on grassland: applications, perspectives and constraints. Europ. J. Agronomy, 29 (2-3), 59-71. doi: 10.1016/j.eja.2008.05.005.

SCHELLBERG, J., MÖSELER, B.M., KÜHBAUCH, W. & RADEMACHER, I., 1999: Long-term effects of fertilizer on soil nutrient concentration, yield, forage quality and floristic composition of a hay meadow in the Eifel Mountains, Germany. Grass and Forage Science, **54**, 195-207. DOI: 10.1046/j.1365-2494.1999.00166.x.

Inklusive Web-Kartographie: Ausgewählte Methoden und Ansätze zur Erfüllung (gesetzlicher) Forderungen nach Barrierefreiheit

SABINE HENNIG[1]

Zusammenfassung: Die Forderung, dass Webinhalte für jeden, d.h. auch für Menschen mit Behinderung, zugänglich und nutzbar sein müssen, ist auch für Online-Karten von Relevanz. Während zahlreiche Materialien die Bereitstellung barrierefreier Webinhalte unterstützen und damit zur digitalen Inklusion der Gesellschaft beitragen, fehlen Richtlinien und Empfehlungen, die die Entwicklung barrierefreier Online-Karten ermöglichen, noch weitgehend. Doch welche Methoden und Ansätze bieten sich an, um zur Umsetzung barrierefreier Online-Karten und im Weiteren zur Entwicklung einer inklusiven Web-Kartographie beizutragen? In Projekten, welche die Bereitstellung barrierefreier Online-Karten fokussierten, wurden entsprechende Erkenntnisse gesammelt. Sie unterstreichen die Bedeutung des Einsatzes partizipativer Ansätze, die Bezugnahme auf Standards und Guidelines hinsichtlich Web Accessibility i.w.S. und die Nutzung von Design Patterns.

1 Einleitung und Fragestellung

Eine Reihe von Gesetzen, Verordnungen, Richtlinien und Programmen (vgl. Tab. 1) fordern bzw. unterstreichen – u.a. in Deutschland – die Forderungen nach der Umsetzung barrierefreier Infrastrukturen und Produkte. Im Hinblick auf eine inklusive Gesellschaft soll jeder Mensch, d.h. auch Menschen mit Behinderung, an allen Aktivitäten des täglichen Lebens teilnehmen können (BMAS 2011). Dies beinhaltet die Nutzung digitaler und web-basierter Anwendungen. In Deutschland verpflichtet speziell die Barrierefreie Informationstechnik-Verordnung BITV die Träger öffentlicher Gewalt (d.h. Behörden der Bundes-/ Landesverwaltungen, die Bundesrecht ausüben), bundesunmittelbare Körperschaften sowie Anstalten und Stiftungen öffentlichen Rechts ihre Internetangebote barrierefrei zu gestalten. Für andere Einrichtungen ist dies derzeit noch nicht verpflichtend, allerdings unterstreichen zahlreiche Aspekte die Notwendigkeit inklusive bzw. barrierefreie Produkte bereitzustellen: Zum einen sind dies gesellschaftliche, ethnische Gründe. Zum anderer spielen ökonomische Überlegungen eine zunehmend wichtige Rolle. Durch mehr Barrierefreiheit und infolgedessen durch ein selbstbestimmteres und unabhängigeres Leben der Betroffenen können steigende Kosten bei Pflege und Betreuung verringert werden (BMVIT 2013).

Bei der Umsetzung von Online-Karten werden die Bedürfnisse von Menschen mit Behinderung bisher oft nicht (ausreichend) berücksichtigt. Zugang zu und Nutzung „traditioneller“ Online-Karten sind für diese Nutzer in vielen Fällen schwierig, teilweise sogar unmöglich. Die Nutzer sehen sich mit zahlreichen und je nach Behinderung unterschiedlichen Barrieren konfrontiert. Beispiele hierfür sind eine ungeeignete Farbauswahl, schlechte Farbkontraste, kleine Schriften,

[1] Paris Lodron Universität Salzburg, Interfakultärer Fachbereich für Geoinformatik – Z_GIS, Schillerstr. 30, A-5020 Salzburg, Austria, E-Mail: sabine.hennig@sbg.at.at

unklar strukturierte Elemente, zu komplexe Inhalte und zu schwierig zu bedienende Funktionalitäten (NEUSCHMID et al. 2012).

Tab. 1: Ausgewählte Gesetze, Verordnungen, Richtlinien und Programme, die in Deutschland die Forderungen nach Barrierefreiheit und Inklusion fordern bzw. unterstreichen

Gesetze, Verordnungen, Richtlinien	In Kraft treten
Grundgesetz (GG; Artikel 3 Abs. 3)	1994
Behindertengleichstellungsgesetz (BGG)	2011
Allgemeines Gleichbehandlungsgesetz (AGG)	2006
Barrierefreie Informationstechnik-Verordnung (BITV 1.0/ BITV 2.0)	2002/ 2011
EU-Richtlinie zur Festlegung eines allgemeinen Rahmens für die Verwirklichung der Gleichbehandlung in Beschäftigung und Beruf	2000
EU Web Accessibility Richtlinie	2016
UN Übereinkommen über die Rechte von Menschen mit Behinderungen	2008
Digital Agenda für Europa 2020 (Bereich Digitale Gesellschaft: e-Health & Altern)	2010
European i2010 initiative on e-Inclusion - to be part of the information society	2007

Während zahlreiche Materialien die Realisierung barrierefreier Webinhalte unterstützen, fehlen solche hinsichtlich der Umsetzung barrierefreier Online-Karten bisher weitgehend. Doch welche Methoden und Verfahren bieten sich an, um die Entwicklung einer inklusiven Webkartographie zu unterstützen. Diese Frage wird anhand von Erkenntnissen diskutiert, die in Projekten mit Fokus auf die Bereitstellung von Online-Karten, geeignet für Menschen mit Behinderung und ältere Menschen, gewonnen wurden: (1) AccessibleMap (2012-2014; gefördert im Österreichischen FFG Programm Benefit), (2) senTour (2012-2014; gefördert im Österreichischen FFG Programm Benefit) und (3) Geomedien 55-PLUS (seit 2011; Initiative des Interfakultären Fachbereich für Geoinformatik-Z_GIS gemeinsam mit der Uni 55-PLUS, Universität Salzburg).

2 Hintergründe zu Inklusion und Barrierefreiheit

1.1 Die Zielgruppe: Menschen mit Behinderung

In Deutschland gilt nach dem Bundessozialhilfegesetz (BSHG) § 124 Abs. 4, Satz 1-4 eine Behinderung als "... eine nicht nur vorübergehende erhebliche Beeinträchtigung der Bewegungsfreiheit, die auf dem Fehlen oder auf Funktionsstörungen von Gliedmaßen oder auf anderen Ursachen beruht […] weiterhin liegen Behinderungen bei einer nicht nur vorübergehenden erheblichen Beeinträchtigung der Seh-, Hör-, und Sprachfähigkeit und bei einer erheblichen Beeinträchtigung der geistigen oder seelischen Kräfte vor." Dementsprechend gelten als Handicap, mit denen sich Menschen konfrontiert sehen können, Körperbehinderungen (Geh-/ Greifbehinderung etc.), Hörbehinderungen und Gehörlosigkeit, Sehbehinderungen und Blindheit, Lernbehinderungen, kognitiven Einschränkungen und Sprachbehinderungen, Allergien und Photosensibilität, innere Erkrankungen, Epilepsie, psychische Erkrankungen, Suchterkrankungen, sowie Mehrfachbehinderungen, d.h. die Kombinationen aus diversen Behinderungen (BEHINDERUNG.ORG n.d.; W3C 2009).

In Deutschland wird zwischen amtlich anerkannten und nicht amtlich anerkannten Behinderungen differenziert. Rund 10 % der Bevölkerung gelten als Menschen mit amtlich anerkannten Behinderungen und 15 % als Menschen mit nicht amtlich anerkannten Behinderungen. Insgesamt

sind damit 25 % der Gesellschaft von amtlich anerkannten und/ oder nicht anerkannten Behinderungen betroffen (BMAS 2013; DEMOS 2013). Dabei sind in Deutschland 52 % der Bevölkerungen mit Körperbehinderungen konfrontiert. Rund 4 % leiden unter Beeinträchtigungen der visuellen Wahrnehmung. 2,5 % sind von Beeinträchtigungen der akustischen Wahrnehmung betroffen. Ebenfalls 2,5 % haben psychische Beeinträchtigungen und etwa 1 % der Bevölkerung leiden unter Lernschwierigkeiten bzw. kognitive Beeinträchtigungen (STATISTISCHES BUNDESAMT 2014).
Die Anzahl von Menschen mit Behinderung ist nicht nur in Deutschland, sondern in vielen (industrialisierten) Ländern steigend. Dies steht im Zusammenhang mit dem demographischen Wandel der Gesellschaft, d.h. der prozentualen Zunahme älterer Menschen am Anteil der Gesamtbevölkerung aufgrund sinkender Geburtenzahlen und zunehmender Lebenserwartung. Da zahlreiche Beeinträchtigungen eine altersbedingte Erscheinung sind, sind ältere Menschen eine zentrale Zielgruppe für inklusive bzw. barrierefreie Produkte (BSI 2003; STATISTISCHES BUNDESAMT 2014).

1.2 Definition und Abgrenzungen

1.2.1. Inklusion und Barrierefreiheit

Prinzipiell bedeutet Inklusion, dass alle Menschen, unabhängig ihrer Herkunft, ihres Geschlechtes, ihres Alters oder einer Behinderung, an allen Prozessen des täglichen Lebens teilhaben und sie diese mitgestalten können (UN 2016). Für Menschen mit Behinderung bedeutet Inklusion, dass alle Barrieren, mit denen sie in ihrem beruflichen wie privaten Alltag konfrontiert sind, abgebaut werden, dass Infrastrukturen und Produkte barrierefrei gestaltet sind. Von Barrierefreiheit wird dabei gesprochen, wenn durch den Menschen gestaltete Lebensbereiche für behinderte Personen in der allgemein üblichen Weise, ohne besondere Erschwernis und ohne fremde Hilfe zugänglich sind (BSI 2003; W3C 2009).

1.2.2. Web Accessibility im engeren und weiteren Sinn

Lange Zeit fokussierte Barrierefreiheit primär bauliche Elemente (z.B. Abflachung von Bordsteinen, Bereitstellung von Rampen, Aufzügen und Leitsystemen). Bedingt durch die rasante Entwicklung im Bereich der Informations- und Kommunikationstechnologien (IKT) und deren Bedeutungszuwachs für alle Aspekte des Lebens bezieht sich Barrierefreiheit (als sog. Web Accessibility) heute ebenfalls auf die Gestaltung digitaler und web-basierter Produkte und Dienste. Diese sind derart zu gestalten, dass sie die Bedürfnisse aller Menschen gleichermaßen berücksichtigen und Menschen mit Behinderung von einer Nutzung nicht ausgeschlossen sind (KERKMANN 2013). Hierfür sind alle Hindernisse, die den Zugang zu Webinhalten erschweren, abzubauen (EU 2006; W3C 2005).
Web Accessibility ist insofern von weitreichender Bedeutung, als dass IKT großes Potential zugesprochen wird, die Unabhängigkeit von Menschen mit Behinderung sowie deren Teilhabe an allen Aspekten des Lebens zu erhöhen und infolgedessen deren Lebensqualität zu verbessern (HARRIS 2010; THORNTON 1993; WHO 2007). Dies gilt aber nicht per se: Um das bestehende Potential nutzen zu können, müssen Anwendungen die Anforderungen der Nutzer erfüllen und v.a. barrierefrei implementiert sein (MACDONALD & CLAYTON 2013). Dafür sind weitere Aspek-

te, die durch Web Accessibility nicht explizit adressiert werden, mit dieser jedoch in engem Zusammenhang stehen, zu berücksichtigen. Dies bezieht sich bspw. auf Usability: Auch wenn Usability und Accessibility verschiedene Aspekte fokussieren, überschneiden sich ihre Zielsetzungen, Kriterien und Empfehlungen vielfach. Die W3C (2010) betont, die Berücksichtigung von Accessibility kann zu besserer Usability für unterschiedliche Nutzergruppen führen. Weiters gilt, dass Webseiten/ -applikationen nicht nur zugänglich und über eine gute Usability verfügen müssen, sondern sie müssen auch hinsichtlich Inhalten, Funktionalitäten und der Nutzung von Ein-/ Ausgabegeräten sowie Assistiven Technologien[2] den Anforderungen der Zielgruppe entsprechen: Zum Beispiel verlangen Menschen mit Behinderung – neben Informationen, die generell in Karten verfügbar sind (z.B. Straßen, Gebäude, Grünflächen und Gewässer) – auch Angaben zu speziell für sie relevanten Infrastrukturen (z.B. Leitsysteme, barrierefreie Toiletten und Verkehrsmittel). Infolge ist die Berücksichtigung von Web Accessibility (im engeren Sinn) nicht genug: Web Accessibility muss in einem weiteren Kontext bzw. Sinn gesehen werden. Um Online-Karten zu entwickeln, die für Menschen mit Behinderung geeignet sind, sind neben Kriterien der Accessibility auch Aspekte bzgl. Usability und Utility zu beachten (ITU/G3ICT 2014a; ITU/G3ICT 2014b; LEPORINI & PATERNÒ 2008).

3 Methoden und Ansätze zur Gestaltung barrierefreier Online-Karten

In den verschiedenen Projekten (AccessibleMap, senTour und Geomedien 55-PLUS; vgl. Kap. 1) wurde eine Reihe ausgewählter Methoden und Ansätzen verwendet (siehe Tab. 2). Sie trugen wesentlich zur Entwicklung barrierefreier Online-Karten für die jeweilige Zielgruppe bei.

Tab. 2: Methoden bzgl. Nutzerverständnis/ Anforderungserhebung verwendet in verschiedenen Projekten mit Fokus auf die Entwicklung von inklusiven/ barrierefreien Online-Karten

Projekt	Methoden & Ansätze
AccessibleMap	*Befragungstechniken*: Fragebogen Literatur-/ Internetrecherche: Standards und Richtlinien *Analysis of Similar Systems AoSS* *Partizipative Ansätze*: Stakeholder Beteiligung
senTour	*Befragungstechniken*: Fragebogen Literatur-/ Internetrecherche: Standards und Richtlinien *Analysis of Similar Systems AoSS* *Partizipative Ansätze*: bewusste systematische Teilnahme und Beobachtung
Geomedien 55-PLUS	*Befragungstechniken*: Conceptual Inquiry *Beobachtungstechniken*: direkte Beobachtung *Partizipative Ansätze*: Participatory Design

[2] Bei Assistiven Technologien handelt es sich um verschiedene Hilfsmittel bzw. Technologien, die helfen, die Funktionseinschränkungen von Menschen mit Behinderung zu überwinden. Indem Assistive Technologien die Fähigkeiten der Nutzer unterstützen, kann die Behinderung der Betroffenen (teilweise) kompensiert werden. Während klassische Assistive Technologien Hilfsmittel wie Stöcke, Rollatoren und Rollstühle, Brillen, Vergrößerungsgläser, Blindenstöcke, Hörgeräte oder Implantate umfassen, beziehen sich digitale Assistive Technologien auf Hard- und Software-Angebote wie alternative Tastaturen und Zeigegeräte, Vergrößerungssoftware (z.B. Bildschirmlupen), Vorlese-Anwendungen (Screenreader), Sprachausgabe oder Braillezeile. Diese Angebote erleichtern bzw. ermöglichen für Menschen mit Beeinträchtigungen den Zugriff auf digitale Medien und Webinhalte (BSI 2003; TOLAR 2008; W3C 2009).

3.1 Klassische Verfahren der Anforderungserhebung und Partizipative Ansätze

Klassische Methoden der Anforderungserhebung, die auch für die Entwicklung barrierefreier Online-Karten relevant sind, sind u.a. Befragungs- (Interviews, Fragebögen, Conceptual Inquiry etc.) und Beobachtungstechniken (direkte/ indirekte, aktive/ passive Beobachtung). In der Literatur werden diese umfassend beschrieben (u.a. bei PRESSMAN 2009; RICHTER & FLÜCKIGER 2010; SOMMERVILLE 2011; USABILTY NET 2006).

Zwischen Nutzern und Entwicklern bestehen in vielen Fällen Kommunikationsprobleme (Missverständnissen, Verwendung unterschiedlichen Vokabulars bzw. von Fachbegriffen etc.). Nutzer sind sich zudem nicht immer ihrer Anforderungen (voll) bewusst und können diese auch nicht immer ausreichend (gut) beschreiben (FIRESMITH 2007; HENNIG & VOGLER 2016). Hier hat sich der Einsatz partizipativer Ansätze als nützlich erwiesen. Neben der Beteiligung von Stakeholdern am Entwicklungsprozess (z.B. von Vertreterorganisationen; vgl. u.a. USABILTY NET 2006) und der teilnehmenden Beobachtung (vgl. u.a. FRIEDRICHS 1990) lassen sich speziell im Rahmen von Participatory Designs die Anforderungen und Bedürfnisse der Nutzer – auch bzgl. Web Accessibility i.w.S. – besser und umfassender identifizieren und verstehen als dies durch klassische Methoden in der Regel der Fall ist (HENNIG & VOGLER 2016).

Participatory Design zielt darauf ab, Repräsentanten der Zielgruppe direkt und aktiv am Design- bzw. Entwicklungsprozess einer Anwendung teilhaben zu lassen. Dies kann auf verschiedene Weise und in verschiedener Intensität erfolgen (BAEK et al. 2007). Unterschieden wird zwischen *weak* und *strong* Participatory Design: Beim w*eak* Participatory Design sind die Nutzer in den gesamten Entwicklungsprozess involviert, Entscheidungen werden aber ausschließlich von den Entwicklern getroffen. Im *strong* Participatory Design werden Entscheidungen dem hingegen gemeinsam mit den Nutzern getroffen. Weitere Informationen zu Participatory Design finden sich u.a. bei ENERSON (2013), MAZZONE & READ (2005), MULLER & DRUIN (2012) und STEEN et al. (2007).

3.2 Standards und Richtlinien

Standards und Richtlinien, die eigens die Entwicklung barrierefreier Online-Karten unterstützen, fehlen bislang weitgehend. Bestehenden Standards und Richtlinien bzgl. Web Accessibility i.w.S., die Informationen zur entsprechenden Gestaltung von Webinhalten liefern, sind allerdings auch für die Umsetzung barrierefreier Online-Karten hilfreich (vgl. u.a. FOWLER 2001; HENNIG et al. 2015). Beispiele sind:

- Web Content Accessibility Guidelines WCAG 2.0[3],
- Standards for disabled and elderly peoples' access to information and communications technologies (ICT) products and services including design for all[4],
- Accessibility requirements suitable for public procurement of ICT products and services in Europe EN 301 549[5],
- Guidance on World Wide Web user interfaces ISO 9241-151[6] und

[3] https://www.w3.org/TR/WCAG20/
[4] http://www.etsi.org/
[5] http://mandate376.standards.eu/standard
[6] https://www.iso.org/standard/37031.html

- Guidance on Usability ISO 9241-11[7]

Unter diesen kommt der von der Web Accessibility Initiative WAI (Bereich innerhalb des World Wide Web W3C) 2009 veröffentlichten WCAG 2.0 (Nachfolger der WCAG 1.0, veröffentlicht 1999) zentrale Bedeutung zu. Die WCAG 2.0 besteht aus vier Prinzipien, zwölf Richtlinien und 61 Erfolgskriterien (vgl. Abb. 1). In vielen Ländern ist sie die Basis für gesetzliche Vorgaben zur Gestaltung barrierefrei zugänglicher Webinhalte.

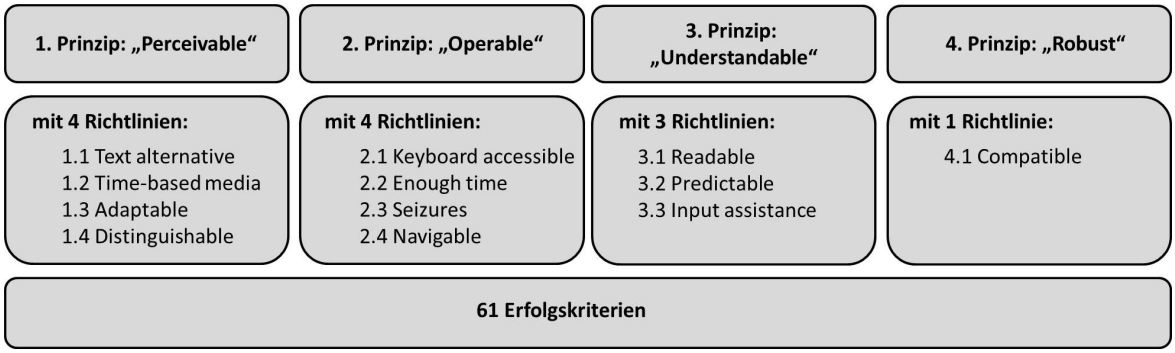


Abb. 1: WCAG 2.0: Prinzipien, Richtlinien und Erfolgskriterien

3.3 AoSS und Design Patterns

Durch die Analyse vergleichbarer Systeme (Analysis of Similar Systems AoSS) kann ein Einblick gewonnen werden, wie andere Systeme gestaltet sind, welche Inhalte, Komponenten und Funktionen idealerweise zu implementierten sind und wie dies erfolgen könnte bzw. sollte. Außerdem zeigt sich, welche Lösungen weniger geeignet sind und welche Probleme zu vermeiden sind. Die Analyse der Anwendungen basiert in der Regel auf einer eigens für die jeweilige Fragestellung erstellten Liste an Kriterien (NEMETH 2004; USABILITY NET 2006).

Mit Hinblick auf die Entwicklung barrierefreier Online-Karten sind eine Reihe verschiedener Anwendungen von Interesse: einerseits Webseiten/ -anwendungen, die barrierefrei gestaltet sind bzw. barrierefreie Elemente enthalten; andererseits Online-Karten mit Lösungen und Komponenten, die im Kontext einer Web Accessibility i.w.S. beachtenswerte Lösungen beinhalten. Kriterien relevant für die Analyse von Anwendungen beziehen sich z.B. auf das Design der Nutzeroberfläche und der Karte bzw. der Kartenkomponente, inhaltliche Aspekte der Karte, Art und Umfang der implementierten Funktionen, Zugang zu Assistiven Technologie sowie Art und Umfang der Berücksichtigung und Umsetzung der diversen Standards und Richtlinien, v.a. der WCAG 2.0.

Ein weiteres Hilfsmittel bei der Gestaltung barrierefreier Online-Karten sind Design Patterns. Für bekannte und bei der Applikationsentwicklung häufig auftretende Probleme beschreiben sie in Textform, ergänzt durch exemplarische Abbildungen, Lösungen, die sich in der Praxis als erfolgreich erwiesen haben. Sie fassen Wissen und Erfahrungen im Hinblick auf bewährter Anwendungen und Komponenten zusammen (PRESSMAN 2009; SOMMERVILLE 2011). Im Software Engineering werden Design Patterns in verschiedenem Kontext genutzt: z.B. als Architectural

[7] https://www.iso.org/standard/16883.html

Patterns, Interaction Patterns oder User Interface Patterns. Letztgenannte beschreiben Lösungen für gängige Probleme im Design von Nutzeroberflächen, die aus Nutzersicht als geeignet gelten (BRECHT 2010).
Design Patterns werden anhand von Templates (Vorlagen) beschrieben. Sie definieren, welche Informationen bereitzustellen sind wie z.B. Name des Design Pattern, Problembeschreibung, Lösungsbeschreibung (inkl. Screenshot), Verwendungszweck, Umsetzungsart, Einsatzgrund und weitere Lösungsbeispiele (ev. mit Screenshot). Verwaltet werden Design Patterns in der Regel in web-basierten Bibliotheken und Repositories (PRESSMAN 2009). Web-basierte Bibliotheken/ Repositories, die Design Patterns graphischer Nutzerschnittstellen verwalten, sind bspw.: UI Patterns[8], Welie.com[9] und Pattern Tap[10].
Hinsichtlich der Gestaltung barrierefreier Webinhalte sowie (traditioneller) Online-Karten stehen User Interface Design Patterns zur Verfügung. So hat das W3C Design Patterns zur Umsetzung barrierefreier Webinhalte veröffentlicht (W3C 2013). Beispiele für Design Patterns bzgl. Online-Karten sind ESRI's Design-Patterns-for-Web-Maps[11], Oracle's Spatial Visualization Design Patterns[12] und Welie's Interaction Design Pattern Library[13]. Konkret für barrierefreie Online-Karten finden sich derzeit keine Design Patterns öffentlich zugänglich. Für die Entwicklung barrierefreier Online-Karten sind sie durchaus von Relevanz.

4 Potential und Herausforderungen bzgl. der genutzten Methoden und Ansätze

Im Hinblick auf das Fehlen von Materialien, die die Entwicklung barrierefreier Online-Karten unterstützen, bieten die beschriebenen Methoden und Ansätze zahlreiche Vorteile. Grundsätzlich ermöglicht ihre Verwendung, Anwendungen zu entwickeln, die besser auf die Nutzer, ihre Bedürfnisse und Ziele ausgerichtet sind.
Nutzer sind Experten für ihre Anforderungen. Durch die Zusammenarbeit mit ihnen im Rahmen von Participatory Design können die Entwickler die zukünftigen Nutzer, ihre Bedürfnisse, Fähigkeiten und Lebensumstände kennenlernen. Dies erlaubt es, die Anforderungen der Zielgruppe umfassend(er) zu identifizieren und zu verstehen. Durch den kontinuierlichen Austausch mit den Nutzern können unerwünschte Entwicklung vermieden werden. Es kann gewährleistet werden, dass die zu entwickelnden Anwendungen den Anforderungen der Nutzer tatsächlich entsprechen. Dabei bringen Nutzer oft neue und innovative Ideen in den Entwicklungsprozess ein (vgl. u.a. HENNIG & VOGLER 2016; STEEN et al. 2007).
Durch die Bereitstellung gleicher bzw. vergleichbarer Lösungen – was insbesondere durch Design Patterns ermöglicht wird – werden Anwendungen in ihrem Design konsistenter und damit für die Nutzer verständlicher und vorhersagbarer. Indem sie praktische Erfahrungen und angewandtes Wissen weitergegeben, ermöglichen sie Wissenstransfer (PRESSMAN 2009; SOMMER-

[8] ui-patterns.com/patterns
[9] http://www.welie.com/patterns/index.php
[10] http://patterntap.com/library
[11] http://blogs.esri.com/esri/arcgis/2008/08/05/design-patterns-for-web-maps/
[12] http://www.oracle.com/webfolder/ux/applications/uxd/endeca/content/library/en/home.html
[13] http://www.welie.com/patterns/showPattern.php?patternID=map-navigator

VILLE 2001). Dies ist insofern bedeutsam, als dass Entwickler, denen die Anforderungen von Menschen mit Behinderungen unbekannt sind (vgl. u.a. HENNIG & VOGLER 2016; TSOU 2003), ein entsprechender Zugang und Einstieg in die Themen Barrierefreiheit und barrierefreie Online-Karten gegeben wird (BRÜNNING 2014). Als Grundlage für Diskussionen verbessern Design Patterns die Kommunikation unter den an einer Produktentwicklung beteiligten Akteuren (BRÜNNING 2014). Diese Punkte gelten auch hinsichtlich der Nutzung bestehender Standards und Guidelines bzgl. Web Accessibility i.w.S.

Neben diesen Vorteilen steht der Einsatz der beschriebenen Methoden und Ansätze im Zusammenhang mit einer Reihe von Herausforderungen. Einige werden im Folgenden kurz skizziert: Als wichtige Grundlage für den Erfolg von Participatory Designs verlangt die Zusammenarbeit mit den Nutzern, die keine Experten bzgl. Geoinformation und Kartographie sind, Sorgfalt bei der Auswahl der Werkzeuge, die bei der Applikationsentwicklung zum Einsatz kommen. Ferner sind geeignete Arbeitsmaterialien bereitzustellen, um die Nutzer bei der Mitarbeit im Entwicklungsprozess anzuleiten und zu unterstützen. Die Arbeitsmaterialien müssen barrierefrei zugänglich sein und sie müssen dem Kenntnisstand der Nutzer gerecht werden und die Ausbildung relevanter Kompetenzen ermöglichen. Dies verlangt bereits ein gewisses Verständnis für die Zielgruppe. Hier kann die Kooperation mit Vertreterorganisationen oder die Mitarbeit speziell geschulter Mitarbeiter relevant sein (HENNIG & VOGLER 2016).

Tab. 3: Auswahl an Prinzipien und Richtlinien der WCAG 2.0. und ihre Bedeutung bzw. Anwendung bei der Gestaltung barrierefreier Online-Karten (Quelle: FOWLER 2001; HENNIG et al. 2015)

Prinzip	Richtlinien	Bedeutung bzgl. Online-Karten
Perceivable	***Text alternative***	Nicht-Text-Inhalte (z. B. Grafiken) verlangen nach Textalternativen
		Verbale Beschreibung des Karteninhalts
	Distinguishable	Präsentierte Informationen müssen so leicht wie möglich wahrnehmbar sein (adäquater Farbgebrauch, gut lesbare Zeichensätzen, Text-Schriftgröße etc.)
		Geeignete Farbwahl/ Größe zur Feature-Darstellung in Karten, Kontrast zwischen den Features etc.
Operable	***Keyboard accessible***	Funktionalitäten sollten immer auch über die Tastatur (u.a. Tastaturkurzbefehlen) ausführbar sein
		Tastaturkurzbefehle: Kartennavigation, Feature-Auswahl etc.
	Navigable	Unterstützung beim Finden von Inhalten, Überblick über die momentane Position
		Werkzeuge (zoom/ pan, Auswahl von Hintergrundkarten/ Layern) müssen angemessen und optimal wahrnehmbar (Größe, Farbwahl, Beschriftung etc.) sein
Understandable	***Readable***	Textinhalte müssen lesbar und verständlich sein; sie sollten von Nutzern und Assistiven Technologien gelesen werden können; für das Verständnis notwendige Informationen müssen verfügbar sein
		Alternative Karte/ Kartenbild („vereinfachte“/ „leichte“ Version): visuell vereinfachtes Kartenbild mit wesentlichem Karteninhalt und einfacher Sprache
	Predictable	Webinhalte sollten vorhersehbar sein, d. h. Inhalte sind in einer vorhersehbaren Anordnung von Webseite zu Webseite zu präsentieren; das Verhalten von funktionalen und interaktiven Komponenten muss vorhersehbar sein
		Online-Karten sind im Aufbau vergleichbar zu gestalten: Nutzeroberflächen- und Kartenkomponenten (Navigationstool, Auswahltool bzgl. Hintergrundkarten/ Layern, Maßstab, Legende etc.) sollten stets an derselben Stelle positioniert sein

Standards und Richtlinien bzgl. Web Accessibility i.w.S. sind für die Gestaltung der jeweiligen Nutzeroberfläche an sich hilfreich. Ihr konkreter Einsatz im Rahmen der Umsetzung der Kartenkomponenten (inkl. zugehöriger Werkzeuge) verlangt jedoch eine intensive Auseinandersetzung mit diesen und eine entsprechende Interpretation und Anwendung. Beispiele hierzu sind in Tab. 3 beschrieben.
Die Erarbeitung von Design Patterns beruht in der Regel auf der langjährigen Erfahrung von Entwicklern (SOMMERVILLE 2011). Da barrierefreie Online-Karten ein relativ junges Phänomen sind, sind Lösungen und Erkenntnisse, auf die bei der Erarbeitung von Design Patterns zurückgegriffen werden kann, derzeit wenig umfangreich. Durch AoSS können Best-Practice-Beispiele beschrieben werden, die als wichtige Informationsquelle dienen können, um User Interface Design Patterns bzgl. barrierefreier Online-Karten bereitzustellen. Die in den Design Patterns beschriebenen Lösungen sollten dabei nicht nur von Experten, sondern unbedingt auch von Vertretern der Zielgruppe evaluiert werden. Gerade bzgl. Barrierefreiheit ist die Evaluierung durch die zukünftigen Nutzer ein zentraler Aspekt (ZOBL et al. 2013).

5 Fazit und Ausblick

Damit Menschen mit Behinderung an allen Prozessen des täglichen Lebens teilhaben können, spielt Barrierefreiheit (d.h. die barrierefreie Gestaltung von Produkten) im Sinn einer inklusiven Gesellschaft eine zentrale Rolle. Barrierefreiheit betrifft auch die Gestaltung von web-basierten Produkten und Diensten, wobei Web Accessibility hier in einem weiteren Sinn – einschließlich der Berücksichtigung von Usability- und Utility Kriterien – zu verstehen ist. Aufgrund von Gesetzen und Verordnungen ist die Berücksichtigung von Barrierefreiheit für Träger öffentlicher Gewalt in Deutschland obligatorisch. Für andere staatliche und nicht staatliche Einrichtungen wird sie zunehmend bedeutsamer, u.a. bedingt durch gesellschaftliche, ethische und ökonomische Aspekte.
Während zur Umsetzung barrierefreier Webinhalte umfangreiche Materialien vorliegen, gilt dies nicht für barrierefreie Online-Karten. Hier können Entwickler, denen die Zielgruppe und ihre Bedürfnisse oftmals wenig bekannt sind, durch die Verwendung ausgewählter Methoden und Ansätze unterstützt werden. Im Rahmen verschiedener Projekte hat sich, neben klassischen Methoden (Befragungen, Beobachtung etc.), der Einsatz partizipativer Ansätze (u.a. Participatory Design) als hilfreich erwiesen. Zudem sind Standards und Richtlinien (Accessibility, Usability etc.) für die Gestaltung barrierefreier Online-Karten eine wichtige Informationsquelle. Für die Verwendung in diesem Kontext müssen sie aber entsprechend interpretiert und angewendet werden. Als Anleitung für die Gestaltung barrierefreier Online-Karten bieten sich prinzipiell auch Design Patterns an. Während für barrierefreie Webinhalte und (traditionelle) Online-Karten Design Patterns existieren, sind sie bzgl. barrierefreier Online-Karten noch zu entwickeln. Erkenntnissen gewonnen bei der AoSS können hier als Informationsquelle dienen.
Durch die vorgestellten Methoden und Ansätze können Erfahrungen gewonnen und Erkenntnisse bereitgestellt werden, um Entwickler bei der Umsetzung barrierefreier Online-Karten zu unterstützen. Dennoch sollte die Erarbeitung konkreter Empfehlungen und Richtlinien für eine inklusive Web-Kartographie angestrebt werden.

6 Literaturverzeichnis

BAEK, E.-O., CAGILTAYIK, K., BOLING, E. & FRICK, T., 2007: User-Centered Design and Development. Handbook of Research on Educational Communications and Technology, Spector, J.M., Merrill, M.D. van Merrienboer, J. J. and Driscoll, M. F. (Hrsg.), Routledge Chapman & Hall, New York, 659-670.

BEHINDERUNG.ORG, n.d.: Die Vielseitigkeit des Begriffs „Behinderung". https://behinderung.org/definition-behinderung.htm (25.010.2018).

BMAS BUNDESMINISTERIUM FÜR ARBEIT UND SOZIALES, 2011: Unser Weg in eine inklusive Gesellschaft. Der Nationale Aktionsplan der Bundesregierung zur Umsetzung der UN-Behindertenrechtskonvention.

BMVIT BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION UND TECHNOLOGIE, 2013: Mobilität im Alter. Ein Handbuch für PlanerInnen, EntscheidungsträgerInnen und InteressensvertreterInnen.

BRECHT, R., 2010: User Interface Patterns for Digital Libraries. Bulletin of IEEE Technical Committee on Digital Libraries **6**(1).

BRÜNNING, W., 2014: Bauanleitung für eine Pattern Library – Teil 1: Warum braucht man eigentlich eine Pattern Library? http://www.produktbezogen.de/bauanleitung-pattern-library-1/ (25.01.2018)

BSI PROJEKTGRUPPE E-GOVERNMENT IM BUNDESAMT FÜR SICHERHEIT IN DER INFORMATIONSTECHNIK, 2003: Barrierefreies E-Government. Leitfaden für Entscheidungsträger, Grafiker und Programmierer.

DEMOS DEMOGRAPHIE, ENTWICKLUNG, MIGRATION ONLINE SERVICE, 2013: Zu Unrecht benachteiligt. http://www.berlin-institut.org/newsletter/160_16_Okt_20013.html.html (25.01.2018).

ENERSON, M., 2013: User-Centered Design & User Participatory Design — What's the Difference. http://www.enervisionmedia.com/user-experience-monitor/2013/08/User-Centered-Design-And-User-Participatory-Design (25.01.2018).

EU EUROPEAN UNION, 2006: Barrierefreier Zugang (eAccessibility). http://europa.eu/legislation_summaries/information_society/strategies/l24226h_de.htm (25.01.2018).

FIRESMITH, D., 2007: Common Requirements Problems, Their Negative Consequences, and the Industry Best Practices to Help Solve Them. Journal of Object Technology, **6**(1), 17-33.

FOWLER, A., 2011: Transcript for Web Mapping Accessibility Presentation. https://ajfowler.wordpress.com/2011/02/14/transcript-of-web-mapping-accessibility-presentation/ (15.1.2016).

FRIEDRICHS, J., 1990: Methoden empirischer Sozialforschung. Opladen.

HARRIS, J., 2010: The Use, Role and Application of Advanced Technology in the Lives of Disabled People in the UK. Disability & Society, **25**(4), 427-439.

HENNIG, S., HÖCKER, K. & MARANO, D., 2015: Online-Karten und Barrierefreiheit. Online-Karten im Fokus, Hennig, S. (Hrsg.), Wichmann, 209-226.

HENNIG, S. & VOGLER, R. 2016. User-Centred Map Applications Through Participatory Design: Experiences Gained During the 'YouthMap 5020' Project. The Cartographic Journal, **53**(3), 213-229.

HENNIG, S. & WASSERBURGER, W.W., 2016: Design Patterns für barrierefreie Online-Karten. AGIT - Journal für Angewandte Geoinformatik, **2016**(2), 308-317.

HENNIG, S., ZOBL, F. & WASSERBURGER, W.W., 2017: Accessible Web Maps for Visually Impaired Users: Recommendations and Example Solutions. Cartographic Perspectives, **88**.

ITU/G3ICT (INTERNATIONAL TELECOMMUNICATION UNION/ GLOBAL INITIATIVE FOR INCLUSIVE ICTS), 2014a: What is an Accessible Product? http://www.e-accessibilitytoolkit.org/toolkit/product_development_and_design/accessible_product (15.1.2018)

ITU/G3ICT (INTERNATIONAL TELECOMMUNICATION UNION/ GLOBAL INITIATIVE FOR INCLUSIVE ICTS), 2014b: Basic Accessibility Principles. http://www.e-accessibilitytoolkit.org/toolkit/eaccessibility_basics/basic_accessibility_principles (15.1.2018)

KERKMANN, F., 2013: Web Accessibility (Barrierefreiheit). http://www.gi.de/nc/service/informatiklexikon/detailansicht/article/web-accessibility-barrierefreiheit.html (25.01.2018).

LEPORINI, B. AND F. PATERNO. 2008: Applying Web Usability Criteria for Vision-Impaired Users: Does It Really Improve Task Performance? International Journal of Human-Computer Interaction, **24**(1), 17-47.

MACDONALD, S. J. & CLAYTON, J., 2013: Back to the Future, Disability and the Digital Divide. Disability & Society, **28**(5), 705-718.

MAZZONE, E., & READ, J., 2005: Not Just Bits of Paper — The Outcomes of Participatory Design Sessions with Children. http://www.chici.org/references/not_just_bits_of_paper.pdf (25.01.2018).

MULLER, M. J. & DRUIN, A., 2012: Participatory Design. The Third Space in HCI. The Human-Computer Interaction Handbook, Jacko, J. (Hrsg.), Lawrence Erlbaum Associates, Hillsdale, 1051-1068.

NEMETH, C. P. 2004: Human Factors Methods for Design. Making Systems Human-Centered. Boca Raton, London, New York, Washington DC, CRC Press.

NEUSCHMID, J., HENNIG, S., SCHRENK, M., WASSERBURGER, W. & ZOBL, F., 2012: Barrierefreiheit von online Stadtplänen – das Beispiel AccessibleMap. Angewandte Geoinformatik, Strobl, J., Blaschke, T. & Griesebner, G. (Hrsg.), Wichmann, 339-347.

PRESSMAN, R., 2009: Software Engineering. Mcgraw Hill Book Co.

RICHTER, M. & FLÜCKIGER, M., 2010: Usability Engineering. Springer

SOMMERVILLE, I., 2011: Software Engineering. Addison-Wesley.

STATISTISCHES BUNDESAMT, 2014: Statistik der schwerbehinderten Menschen 2013. Wiesbaden.

STEEN, M., KUIJT-EVERS, L. & KLOK, J., 2007: Early user involvement in research and design projects – A Review of Methods and Practices. 23rd EGOS Colloquium, Vienna, July 5–7.

THORNTON, P., 1993: Communications Technology – empowerment or disempowerment Disability. Handicap & Society, **8**(4), 339-349.

TOLAR, M., 2008: Assistive Technologien. Studie im Auftrag des Bundeskanzleramtes. Institut für Gestaltungs- und Wirkungsforschung, Technische Universität Wien.

TSOU, M.H., 2003: An Intelligent Software Agent Architecture for Distributed Cartographic Knowledge Bases and Internet Mapping Services. Maps and the Internet, Peterson, M.P. (Hrsg.), Elsevier Press, Oxford, 229–243.

UN UNITED NATIONS, 2016: Identifying social inclusion and exclusion. www.un.org/esa/socdev/rwss/2016/chapter1.pdf (25.01.2018).

USABILITY NET, 2006: Evaluate existing systems. http://www.usabilitynet.org/tools/existing.htm (25.01.2018).

WHO WORLD HEALTH ORGANIZATION, 2007: Global Age-Friendly Cities: A Guide. Geneva, WHO Press.

W3C WORLD WIDE WEB CONSORTIUM, 2005: Introduction to Web Accessibility. https://www.w3.org/WAI/intro/accessibility.php (25.01.2018).

W3C WORLD WIDE WEB CONSORTIUM, 2009: Web Content Accessibility Guidelines (WCAG) 2.0. https://www.w3.org/Translations/WCAG20-de/ (25.01.2018).

W3C WORD WIDE WEB CONSORTIUM, 2010: Web Accessibility and Usability Working Together, at https://www.w3.org/WAI/intro/usable (25.01.2018).

W3C WORLD WIDE WEB CONSORTIUM, 2013: Usability of Accessible UI Design Patterns, https://www.w3.org/WAI/RD/wiki/Usability_of_Accessible_UI_Design_Patterns (25.01.2018).

ZOBL, F., HENNIG, S., NEUSCHMID, J. & WASSERBURGER, W.W., 2013: Barrierefreie Karten: Entwicklung einer web-basierten Desktop- und mobilen Anwendung für seh-beeinträchtigte und blinde Personen. Kartographische Nachrichten, **2013**(6), 319-325.

NLGA-Map: Web-gestützte thematische Karten mit JavaScript und Leaflet

THOMAS HURTIG[1] & HOLGER SCHARLACH[1]

Zusammenfassung: Für die Darstellung von webbasierten thematischen interaktiven Karten wurde die JavaScript Bibliothek NLGAMap entwickelt. NLGAMap basiert auf der Open Source Bibliothek Leaflet und bietet neben den in Leaflet enthaltenen grundlegenden webmapping Funktionen eine spezielle API um Choroplethen- und Signaturenkarten zu erstellen. NLGAMap wurde vor dem Hintergrund entwickelt, eine möglichst einfache, aber dennoch flexible Open Source Lösung für die verschiedenen Kartendarstellungen in der Gesundheitsberichterstattung des NLGA zu schaffen. NLGAMap bietet unter anderem Funktionen für die Einbindung einer Basiskarte, die Klassifizierung von Daten anhand statistischer Verfahren, die Auswahl von Farbschemata sowie die Möglichkeit, zeitbezogene Daten darzustellen. Der Quellcode von NLGAMap wurde als Open Source Projekt im Internet veröffentlicht.

1 Einleitung

Geht es um Fragestellungen zu Gesundheit/Krankheit, spielt bei vielen Themen der Raumbezug eine zentrale Rolle. Wo sind die meisten Fälle einer Erkrankung aufgetreten? Warum ist die Lebenserwartung in einer Region höher als in einer anderen? Sollen diese raumbezogenen Daten dargestellt werden, sind Karten das Mittel der Wahl. Gegenüber tabellarischen und textlichen Darstellungen haben Sie den Vorteil, dass der Betrachter räumliche oder raum-zeitliche Ereignisse und Strukturen auf einen Blick erfassen kann. Das in Karten erkennbare räumliche Muster ist oft der Ausgangspunkt für Fragen nach dem „Warum?“.

Bereits seit dem 19. Jahrhundert werden Karten verbreitet zur Darstellung von Krankheiten eingesetzt, beginnend mit den Cholera-Epidemien zu dieser Zeit. Dabei stand der Einsatz von Karten zunächst in engem Zusammenhang mit der krankheitsökologischen Sichtweise, die Zusammenhänge zwischen regionaltypischen Umweltbedingungen (Klima, Wasser, Boden), Lebensgewohnheiten (zum Beispiel Ernährung) und Gesundheit sieht (KISTEMANN & SCHWEIKART 2010). Als neues Thema kam in den 1970er Jahren die gesundheitliche Versorgungsforschung hinzu (WENNBERG & GITTELSOHN 1971).

Dabei nehmen Karten überwiegend zwei Aufgaben wahr. Sie dienen der Analyse räumlicher Zusammenhänge und der Erkennung regionaler Risikofaktoren sowie der Kommunikation von Sachverhalten und den Ergebnissen dieser Analysen (SCHARLACH & DRESSMAN 2013). Einen Überblick über die kartographische Visualisierung in der Gesundheitsgeographie geben PIEPER & FRANKE (2017).

[1] Niedersächsisches Landesgesundheitsamt, Roesebeckstraße 4-6,
D-30449 Hannover, E-Mail: [Thomas.Hurtig, Holger.Scharlach]@nlga.niedersachsen.de

1.1 Karten in der Gesundheitsberichterstattung

Zum Aufgabenbereich der modernen Gesundheitsberichterstattung (GBE) gehört die Ermittlung des vordringlichen politischen Handlungsbedarfs im Hinblick auf die gesundheitliche Lage und Versorgung von Bevölkerungsgruppen. „GBE zieht dazu gesundheitsbezogene Daten und Informationen heran, bewertet sie hinsichtlich ihrer Aussagekraft, analysiert sie mit wissenschaftlichen Methoden und stellt sie verdichtet und adressatenorientiert dar. GBE ist auf Wiederholbarkeit und Vergleichbarkeit angelegt." (KELLERHOFF 1996) Damit folgt sie auf die Medizinalstatistik, die Gesundheitsdaten bis in die 1980er Jahre überwiegend in Form von Tabellen veröffentlicht hat.

Die Gesundheitsberichterstattung richtet sich insbesondere an zwei Zielgruppen: die Politik und die Öffentlichkeit. Sie unterstützt die Politik, damit „Entscheidungen im Gesundheitswesen datengestützt, rational und optimierungsorientiert gefällt werden können" (KUHN et al. 2005) und informiert Bürgerinnen und Bürger über die gesundheitliche Situation. Bei beiden Zielgruppen handelt es sich um fachliche Laien, so dass „insbesondere auch ein klares Layout sowie ein an den Empfänger angepasster Sprachstil, Grafiken und Tabellen zur Illustration der Ergebnisse gefordert" wird (KUHN et al 2005).

Eine zunehmende Anzahl an verfügbaren Daten, kürzere Aktualisierungszyklen und eine steigende Bedeutung von wirtschaftlichen Gesichtspunkten erforderten neue technische Konzepte für die Veröffentlichung von Gesundheitsdaten. Die Veröffentlichung von Basisdaten der Gesundheitsberichterstattung, die jährlich oder in kürzeren Abständen aktualisiert werden, erfolgt heute überwiegend durch datenbankgestützte Internetseiten. Darüber hinaus bietet das Internet im Vergleich mit einer gedruckten Publikation die Möglichkeit, Daten direkt zum Download in Form von Tabellen zur Verfügung zu stellen.

Raumbezogene Daten spielen seit dem ausgehenden 20. Jh. eine zunehmend wichtigere Rolle in der Gesundheitsberichterstattung, was auf zwei Ursachen zurückgeführt werden kann. Erstens der Wandel des Gesundheitssystems von der Angebots- hin zur Nachfrageorientierung und zweitens der zunehmend regionale Blick auf die Gesundheitsversorgung und Public Health (ZWACKHALS et al 2004). Animierte Karten und Grafiken erlauben eine anschauliche Darstellung zeitlicher Veränderungen und die interaktive Verknüpfung mit weiteren Informationen erlaubt eine zielgerichtete Datenauswertung in wenigen Schritten (PETERSON 1999).

Ein frühes Beispiel für eine internetbasierte Veröffentlichung von raumbezogenen Daten der Gesundheitsberichterstattung ist der nationale Gesundheitsatlas der Niederlande (Zorgatlas), der seit Ende der 1990er Jahre entwickelt wurde und 2013 mehr als 1000 Karten umfasste (ZWACKHALS et al. 2013). Der Atlas ist heute in das Gesundheitsportal *Volksgezondheidenzorg.info* integriert, das umfangreiche Information zur Gesundheitsversorgung und zum Gesundheitszustand der Bevölkerung bereitstellt.

In Deutschland werden die Daten der Gesundheitsberichterstattung neben dem zentralen Portal GBE-Bund von den einzelnen Bundesländern in unterschiedlichem Umfang veröffentlicht. NRW stellt Daten der Gesundheitsberichterstattung in einem internetbasierten Gesundheitsatlas (LANDESZENTRUM FÜR GESUNDHEIT 2017) zur Verfügung, ebenso wie Bayern (BAYERISCHES LANDESAMT FÜR GESUNDHEIT UND LEBENSMITTELSICHERHEIT 2015). In beiden Bundesländern wird für die Darstellung die proprietäre Software InstantAtlas verwendet.

1.2 Karten als Teil der Gesundheitsberichterstattung im NLGA

Im Land Niedersachsen ist das Niedersächsische Landesgesundheitsamt (NLGA) die zentrale Fachbehörde für übertragbare und nichtübertragbare Erkrankungen sowie Gesundheitsrisiken der Bevölkerung. Das NLGA hat u.a. die gesetzliche Aufgabe, Fachberichte zur gesundheitlichen Situation der Bevölkerung im Rahmen der Gesundheitsberichterstattung zu erstellen (Niedersächsisches Gesetz über den öffentlichen Gesundheitsdienst (NGöGD), §8 (4)). Die Gesundheitsberichterstattung dient nach dem NGöGD der Planung und Durchführung von Maßnahmen, die die Gesundheit fördern und Krankheiten verhüten. In den Berichten sollen die Daten und Informationen zielgruppenbezogen und geschlechterspezifisch dargestellt und bewertet werden (NGöGD, §8 (1)).
Das NLGA hat 2003 damit begonnen, die Daten der gedruckten Jahresberichte zu Infektionskrankheiten und Impfungen in Form von interaktiven Anwendungen auf der Internetseite des NLGA zur Verfügung zu stellen. Eine Übersicht geben SCHARLACH und SCHARLACH (2008).
Umgesetzt wurden diese Informationssysteme mit der proprietären Software Adobe Flash. Für die Wahl der Technik waren folgende Argumente ausschlaggebend (SCHARLACH & SCHARLACH 2008).

- Vektordatenformat und damit kleine Dateigrößen
- Datenimport über eine Datenbankanbindung oder Textdatei
- vom Programm bereitgestellte umfangreiche Funktionalität zum Erstellen komplexer, interaktiver, multimedialer Internet-Anwendungen, ergänzt durch die programmeigene Skriptsprache Action Script
- zum damaligen Zeitpunkt weite Verbreitung des Flash-Players, standardmäßige Integration in vielen Internetbrowsern

Allerdings verursachte der Flash-Player in den letzten Jahren zunehmend Sicherheitsprobleme. Als Konsequenz kündigte Apple 2016 an, mit der Version 10 seines Internetbrowsers Safari den Flash-Player nicht mehr automatisch zu unterstützen. Zu diesem Zeitpunkt wurde absehbar, dass die Zukunft des Flash-Players fraglich ist. Im Sommer 2017 folgte dann die Ankündigung von Adobe, die Weiterentwicklung und Verbreitung des Flash-Players Ende 2020 einzustellen. Somit entstand die Notwendigkeit, die kartenbasierten Informationssysteme des NLGA auf ein neues technisches System zu portieren. Damit ergab sich die Möglichkeit, die Ziele für eine solche Kartenanwendung neu zu definieren:

- Erstellung webgestützter thematischer Karten mit möglichst wenig Programmieraufwand
- Bereitstellung einer an den Bedürfnissen des NLGA orientierten Basisfunktionalität (Animationen, verschiedene Kartentypen etc.)
- Verwendung von Open-Source-Software
- Veröffentlichung der Karten sowohl auf den Internetseiten als auch Ableitung im Vektorformat für gedruckte Berichte
- Flexible Datenanbindung (Datei oder Datenbank)

Ähnliche Ziele verfolgt die an der Universität Wien entwickelte Web Mapping API mapmap.js (LEDERMANN & GARTNER 2015), die auf der von Mike Bostock entwickelten D3-Bibliothek (https://d3js.org/) aufsetzt. Mapmap.js bietet Möglichkeiten zur Erstellung verschiedener thema-

tischer Kartentypen. Die Entscheidung fiel jedoch zugunsten einer eigenen Lösung unter Verwendung von Leaflet (http://leafletjs.com) aus, da Leaflet bereits eine umfangreiche Funktionalität für das Interface bereitstellt (Zoom, Vollbildmodus, Ebenenkonzept).

2 Grundkonzept

NLGAMap ist eine JavaScript Bibliothek, die die Erstellung von webbasierten thematischen Karten vereinfachen soll. NLGAMap basiert dabei auf der Open Source Bibliothek Leaflet und nutzt deren grundlegenden Mapping-Funktionen. NLGAMap wurde so konzipiert, dass möglichst wenig Programmcode geschrieben werden muss, die Darstellung aber dennoch individuellen Bedürfnissen angepasst werden kann. Um diese Anforderung zu erfüllen, wird die Erstellung einer Karte hauptsächlich über ein Konfigurationsobjekt realisiert, das bei der Initialisierung einer neuen Karteninstanz übergeben wird. Für alle optionalen Eigenschaften im Konfigurationsobjekt wurden Standardwerte festgelegt, so dass eine neue Karte mit möglichst wenig Aufwand erstellt werden kann. Neben dem Konfigurationsobjekt stehen zusätzlich API-Funktionen zur Verfügung, mit denen die Kartendarstellung weiter angepasst werden kann. Außerdem kann über NLGAMap auf die jeweilige Leaflet-Instanz zugegriffen werden wodurch alle Leaflet-Funktionen verwendet werden können.
Grundlegend unterscheidet NLGAMap zwischen den Ebenentypen Choropleth und Marker. Beide Ebenentypen können einzeln oder auch kombiniert verwendet werden. Die Basis bildet dabei eine Umrisskarte, die im *GeoJSON* (BUTLER et al. 2016) bzw. *TopoJSON*-Format (BOSTOCK & METCALF 2017) vorliegen muss und im Konfigurationsobjekt angegeben wird. In dieser *JSON*-Datei sind die einzelnen Flächen der Karte in einer sogenannten *FeatureCollection* hinterlegt. Jede Fläche ist dabei als *Feature* vom Typ *Polygon* oder *MultiPolygon* angelegt und definiert anhand von geographischen Koordinaten die Umrisse der Basiskarte. Für eine spätere Referenzierung muss außerdem jedes *Feature* eine *id* aufweisen, die in der Eigenschaft *properties.id* definiert wird. Über diese *id* wird der entsprechende Datensatz der Choroplethenkarte mit der Basiskarte verknüpft (siehe Abbildung 1).

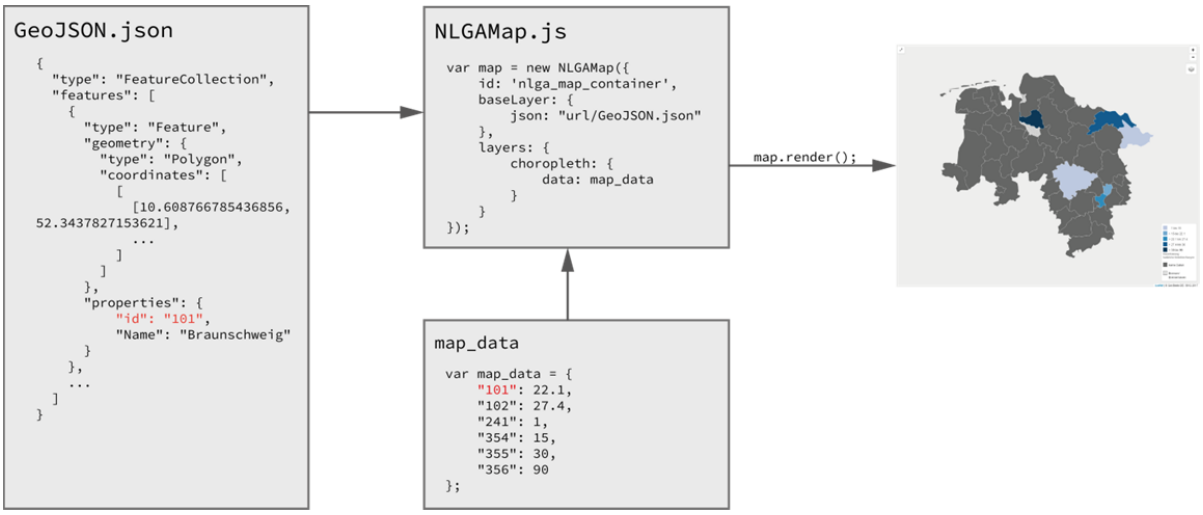


Abb. 1: Grundkonzept NLGAMap

Die Trennung von geographischen Koordinaten und darzustellenden Datenwerten bei Choroplethenkarten bietet den Vorteil, dass die Darstellungsart leicht verändert werden kann. So kann anstelle einer geographischen Basiskarte jegliche polygone Form als Kartengrundlage genutzt und über die *id* mit dem jeweiligen Datensatz verknüpft werden. Ein Beispiel hierfür zeigt Abbildung 2.
Weiterhin kann die zu übertragende Datenmenge deutlich reduziert werden, da bei verschiedenen Datensätzen nur die Datenwerte und nicht die geographischen Koordinaten übertragen werden müssen. Das wirkt sich besonders positiv aus, wenn häufig die gleiche Basiskarte verwendet wird.

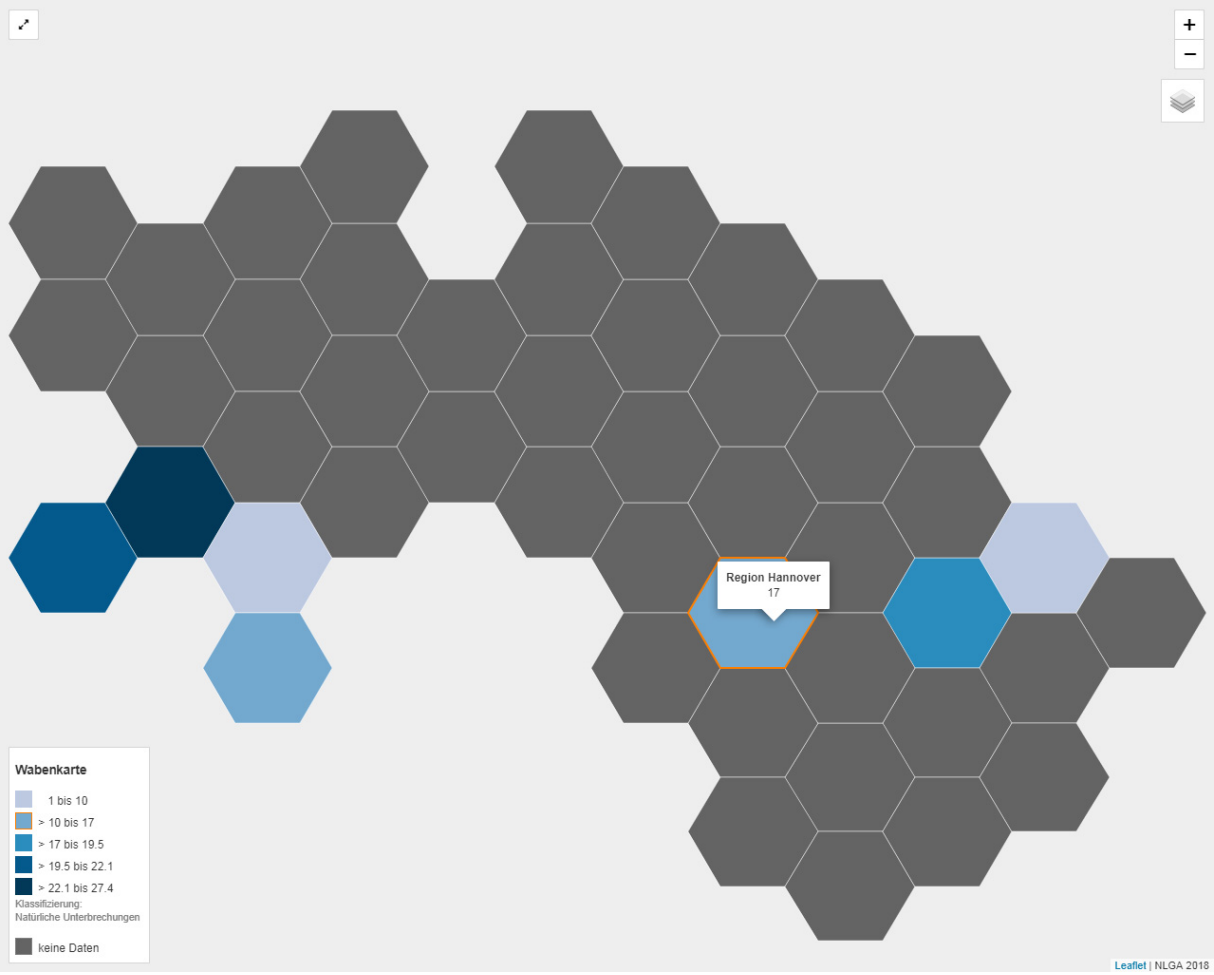


Abb. 2: Beispielhafte Darstellung als Wabenkarte

Bei Markerebenen werden im Gegensatz zu Choroplethebenen die Daten direkt im *GeoJSON*-Format angegeben. Das *GeoJSON*-Objekt enthält dabei sowohl die geographischen Koordinaten als auch den Datenwert. Die Daten werden dann in Form einer graphischen Signatur räumlich angeordnet dargestellt.

3 Basisfunktionen

NLGAMap nutzt die grundlegenden Funktionalitäten von Leaflet, um interaktive webbasierte Choroplethen- und Diagrammkarten zu erstellen. Dazu gehören die Positionierung und das Bewegen der Karte, die Zoom-Funktionen, eine Vollbild-Funktion sowie die Steuerung der Ebenen. Aufbauend auf diesen Funktionen nutzt NLGAMap die Leaflet-API um standardmäßig eine Legende und eine Popup-Funktion für jede Ebene bereitzustellen.
Die Legende wird dabei entsprechend des Ebenentyps angepasst. Bei Markerebenen wird standardmäßig der Kreismarker in der Legende angezeigt. Bei Choroplethenebenen werden die entsprechenden Klassengrenzen und die dazugehörigen Farben angezeigt. Das Erscheinungsbild der Legende kann beliebig angepasst werden. Über das Konfigurationsobjekt können der Titel, die Einheit aber auch die Texte für Standardfelder wie zum Beispiel "keine Daten" gesetzt oder ausgeblendet werden. Weiterhin lassen sich statistische Informationen, wie die verwendete Klassifizierungsmethode, das Minimum, das Maximum, der Mittelwert, der Median sowie die Standardabweichung in der Legende anzeigen. Außerdem kann ein eigenes HTML-Template für die Legende angegeben werden.
Für jeden Teilbereich und für jeden einzelnen Marker wird ein Popup erstellt, das den jeweiligen Namen und Wert anzeigt, wenn sich der Mauszeiger über der entsprechenden Fläche befindet. Auch für das Popup lassen sich verschiedene Einstellungen über das Konfigurationsobjekt vornehmen und ein eigenes HTML-Template angeben.

4 Choroplethenkarten

Choroplethenkarten bilden die Grundlage für alle Kartendarstellungen im Anwendungsbereich des NLGA. Deshalb wurde auf die einfache Umsetzung dieser Darstellungsart besonderen Wert gelegt. Wie unter dem Punkt Grundkonzept erläutert, wird für die Erstellung einer Choroplethenkarte zunächst eine Basiskarte anhand von Geodaten erstellt. Über eine *id* werden anschließend den einzelnen Teilflächen der Basiskarte Datenwerte zugeordnet.
Die Färbung der Teilflächen wird durch die Klassifizierung der Daten und durch das gewählte Farbschema festgelegt. Bei der Klassifizierung werden die Daten in Gruppen eingeteilt und jeder Gruppe wird eine Farbe zugeordnet. Die Wahl der Klassengrenzen hat damit einen erheblichen Einfluss auf den optischen Eindruck der Karte. So können ungünstig gewählte Klassengrenzen zu einer starken Kartenfärbung führen und so zum Beispiel den Eindruck einer akuten Ausbruchssituation entstehen lassen. Um dem Vorzubeugen, wurden in NLGAMap statistische Klassifikationsverfahren mit Hilfe der JavaScript-Bibliothek *geostats* (https://github.com/simogeo/geostats) integriert, die für die Erstellung von Choroplethenkarten geeignet sind. Es kann zwischen den Methoden "Gleiches Intervall", "Quantil", "Natürliche Unterbrechungen" und "Standardabweichung" ausgewählt und die Anzahl der benötigten Klassen angegeben werden. Durch die Wahl eines für die jeweilige Datengrundlage und Datenverteilung geeigneten Klassifizierungsverfahrens können die eingangs erwähnten Effekte verringert werden. Um möglichst flexibel zu sein, können außerdem eigene Klassengrenzen angegeben werden. Die Klassengrenzen können dabei auch mit einer offenen oberen und unteren Grenze definiert werden.

Für die Erstellung von Choroplethenkarten stehen verschiedene Farbschemata zur Auswahl. NLGAMap verwendet dabei die Farbschemata von Colorbrewer (http://colorbrewer2.org). Diese Farbwerte wurden speziell für die Verwendung in kartographischen Abbildungen entwickelt. Jedes Farbschema wurde dabei für bestimmte Anwendungsfälle und für verschiedene Klassenanzahlen entworfen. Über das Konfigurationsobjekt können außerdem eigene Farbwerte definiert werden. Weiterhin besteht die Möglichkeit, eine kontinuierliche Farbskala zu verwenden. Dabei wird jedem Wert eine eigene Farbe zugeordnet ohne die Werte in Klassen zusammenzufassen.

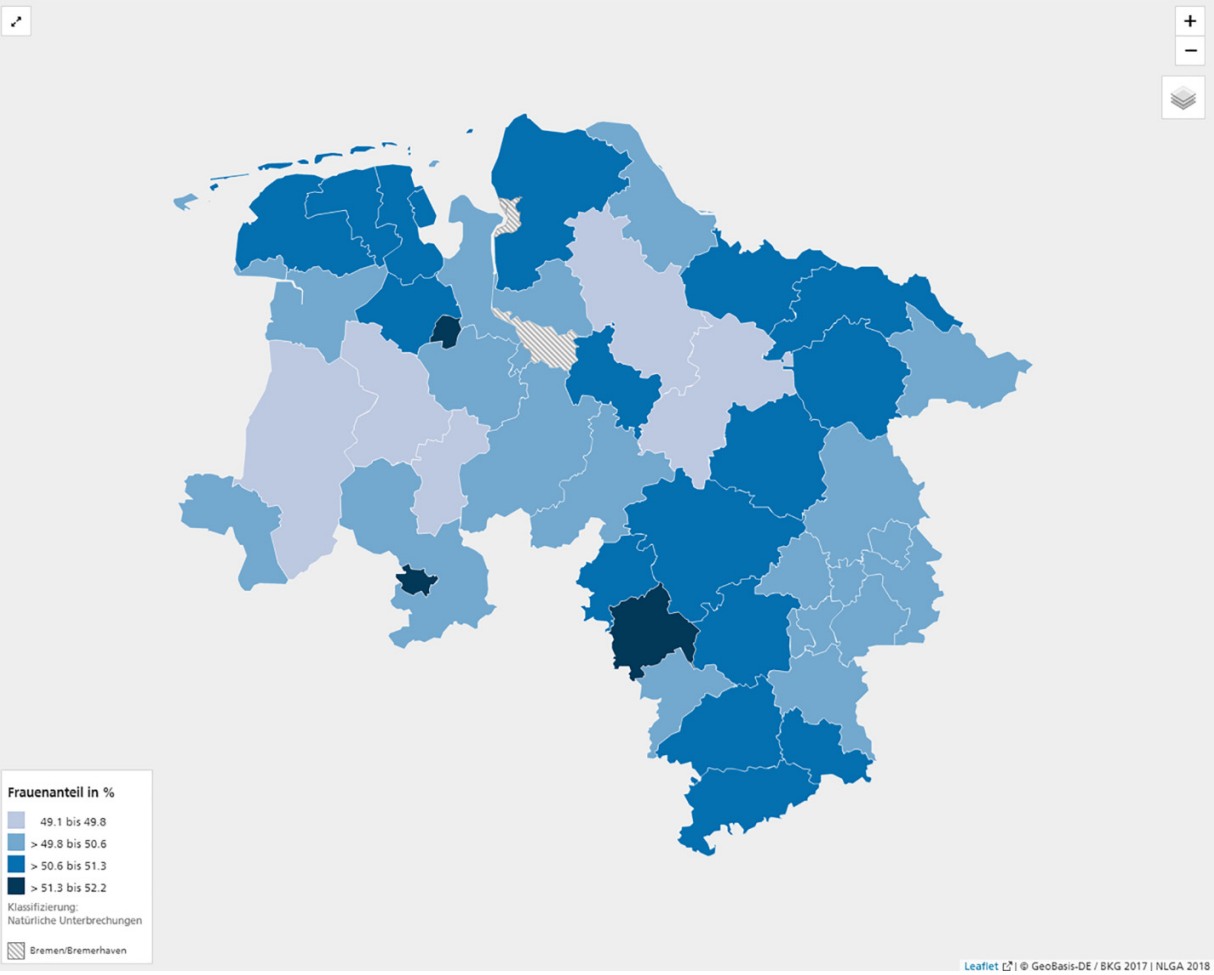


Abb. 3: "Frauenanteil an der Bevölkerung in %"; Choroplethenkarte; 4 Klassen; Klassifizierung "Natürliche Unterbrechungen"; Farbschema NLGA-Blau

Eine weitere Besonderheit sind "Textklassen". Damit können Daten dargestellt werden, die bereits in Kategorien wie zum Beispiel "gering", "mittel", "hoch" eingeteilt wurden. Die Textklassen müssen dafür zusammen mit dem zu der jeweiligen Klasse gehörenden Farbwert im Konfigurationsobjekt definiert werden. Die Datenwerte enthalten dann den Index der jeweiligen Textklasse und optional den realen Datenwert für die Anzeige im Popup. Abbildung 5 zeigt ein Beispiel für die Verwendung von Textklassen.

5 Signatur- und Diagrammkarten

Die Erstellung von Signatur- und Diagrammkarten wird in NLGAMap über sogenannte Markerebenen umgesetzt. Für die Erstellung einer Markerebene müssen die Daten neben dem Wert die geographischen Koordinaten enthalten und im *GeoJSON*-Format vorliegen. Jeder Datenpunkt wird dann an der entsprechenden geographischen Position als Markierung auf der Karte dargestellt. Die Größe der Markierung wird dabei anhand des Datenwerts errechnet. Bisher steht als Markierung eine Kreisdarstellung zur Verfügung. Der Radius des Kreises wird dabei abhängig vom Datenwert dargestellt. Es kann dabei zwischen einer linearen oder logarithmischen Skalierung gewählt sowie der maximale Radius angegeben werden. Die Farbe des Kreises kann per CSS beeinflusst werden. Zu einem späteren Zeitpunkt sollen weitere Signaturarten hinzugefügt werden. Weiterhin können durch die Angabe einer Callback-Funktion eigene Signaturen erstellt werden. Ein Beispiel für eine Signaturenkarte mit Kreissignaturen ist in Abbildung 4 zu sehen.

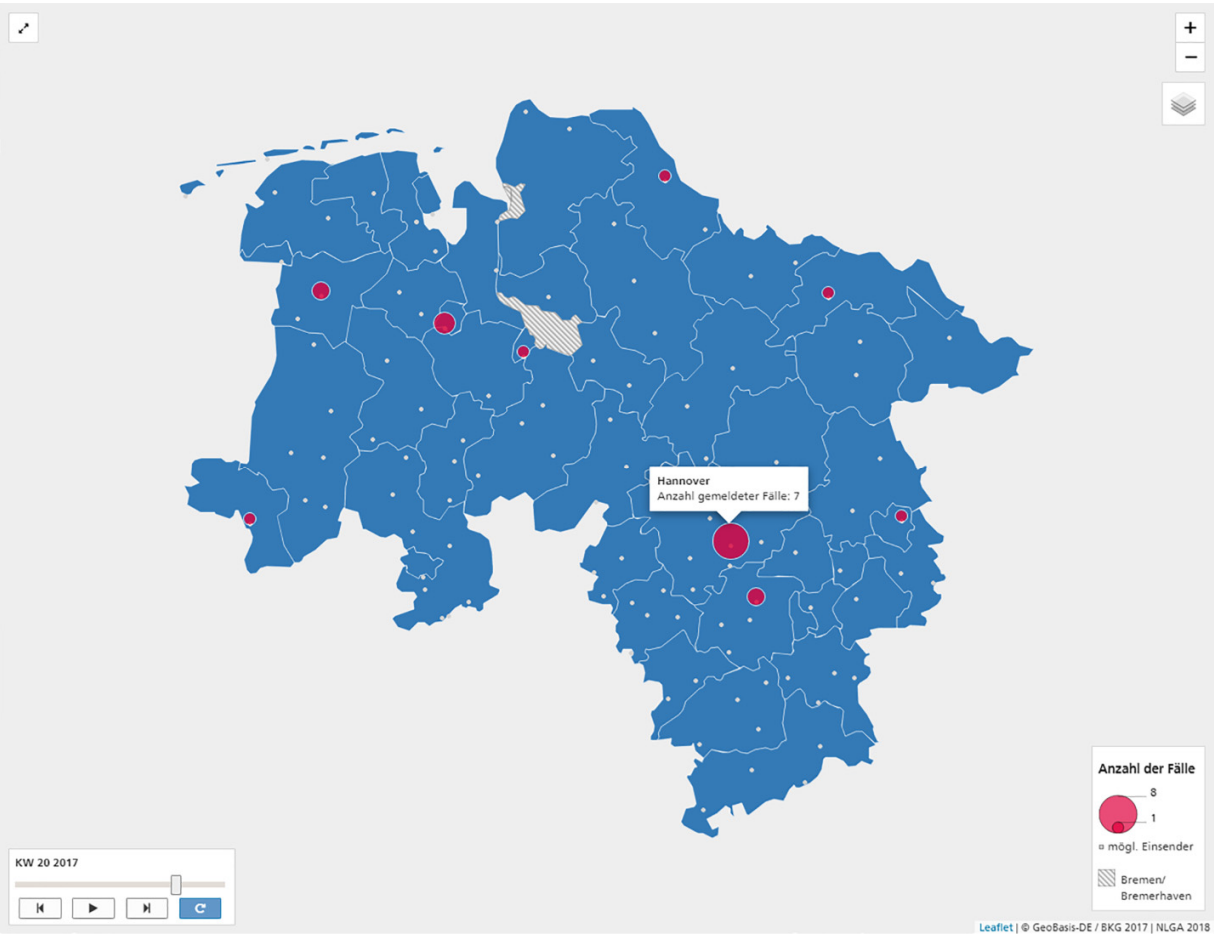


Abb. 4: "Meningitis-Verdachtsfälle: Gemeldete Fälle pro Kalenderwoche"; Signaturenkarte mit Zeitverlauf

6 Zeitbezogene Kartendarstellungen

NLGAMap erlaubt die Darstellung von zeitbezogenen Datensätzen. Der Zeitbezug wird als beliebige Zeichenkette angegeben und kann in einer oder in allen Ebenen vorhanden sein. Es muss nur darauf geachtet werden, dass alle Ebenen die gleichen Zeichenketten als Zeitbezug verwenden.
Bei Choroplethenebenen wird der Zeitbezug in der Datenstruktur abgebildet. Jeder Zeitabschnitt wird durch ein neues Objekt innerhalb des Datenobjektes gekennzeichnet. Der jeweilige Schlüsselwert repräsentiert dabei die Zeitreferenz. Bei Markerebenen wird der Zeitbezug durch eine zusätzliche Eigenschaft im jeweiligen *GeoJSON-Feature* angegeben. Der Name der Eigenschaft muss dabei im Konfigurationsobjekt der jeweiligen Markerebene als *timeKey* definiert werden.

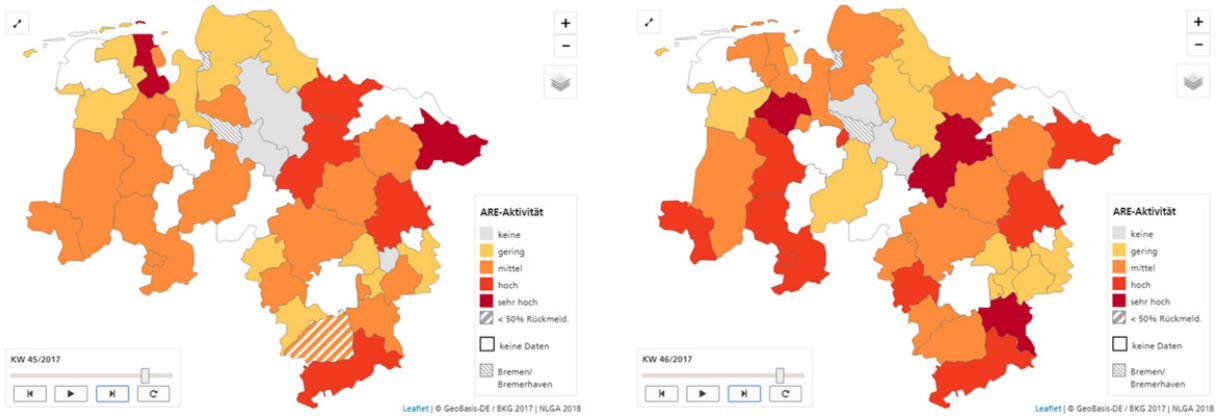


Abb. 5: "ARE-Aktivität in Niedersachsen"; Choroplethenkarte im Zeitverlauf; Textklassen

NLGAMap fügt der Darstellung automatische eine Zeitleistensteuerung hinzu, wenn es sich um eine zeitbezogene Darstellung handelt. Über die Zeitleistensteuerung wird der aktuell Zeitpunkt ausgewählt und die entsprechenden Daten in der Kartendarstellung gezeigt. Der zeitliche Verlauf der Datensätze kann außerdem als Animation abgespielt werden. Zusätzlich kann über das Konfigurationsobjekt eine Callback-Funktion angegeben werden, die bei jeder Änderung des dargestellten Zeitbezugs aufgerufen wird. NLGAMap stellt über *NLGAMap.timeline* eine API zur Steuerung der Zeitleiste bereit. Damit kann zum Beispiel der aktuell dargestellte Zeitraum direkt per JavaScript gesetzt werden. Durch die API und die Möglichkeit einer Callback-Funktion können weitere interaktive Elemente auf der Internetseite implementiert werden, die auf die Änderung der Zeitleiste bzw. des Diagramms reagieren.

7 Fazit und Ausblick

Mit NLGAMap wurde eine JavaScript Bibliothek auf Basis von Leaflet und weiteren Open Source JavaScript Bibliotheken erstellt. NLGAMap reduziert den Aufwand für die Erstellung von interaktiven webbasierten thematischen Karten indem viele Basisfunktionen standardmäßig bereitgestellt werden. Neben der Erstellung der Kartengrundlage wurden außerdem Funktionen für die Klassifizierung der Daten implementiert. Durch die Verwendung eines Konfigurationsob-

jekts ist die Bibliothek flexibel genug um die Darstellung individuellen Bedürfnissen anzupassen und verschiedene Kartentypen umzusetzen.
NLGAMap profitiert sehr stark von Leaflet und anderen bereits vorhandenen Open Source Projekten und wäre ohne diese nicht möglich. Um die interaktive Darstellung von thematischen Karten auf Basis von Webtechnologien weiter voran zu treiben, wurde der Quellcode von NLGAMap als Open Source Bibliothek unter der BSD-Clause-3-Lizenz im Internet veröffentlicht (https://github.com/nlga/NLGAMap). Auch wenn das NLGA die Software zur Verfügung stellt, kann kein Support gewährleistet werden, da hierzu keine personellen Kapazitäten vorhanden sind. Dennoch haben durch die Veröffentlichung andere Programmierer*innen die Möglichkeit, an der Bibliothek weiter zu arbeiten und deren Funktionalität zu erweitern.
NLGAMap wurde unter der Prämisse erstellt, die Erstellung von webbasierten thematischen Karten im Umfeld des NLGAs zu erleichtern. Dadurch wurden bisher primär Funktionen umgesetzt die häufig in NLGA-Webanwendungen genutzt werden. Zukünftig soll NLGAMap aber um neue Funktionen erweitert werden. In einem nächsten Schritt soll eine API für den Grafik-Export implementiert werden. Dabei soll es möglich sein, die generierten Karten im PNG und SVG Format zu speichern, damit sie zum Beispiel für den Druck verwendet werden können. Weiterhin soll die API dahingehend erweitert werden, dass Web Map Services und Web Feature Services, wie sie zum Beispiel vom Bundesamt für Kartographie und Geodäsie (http://geodatenzentrum.de) bereitgestellt werden, leichter verwendet werden können. Die Legende der Choroplethenebene soll ähnlich wie bei LEDERMANN und GARTNER (2015) eine Funktion bekommen, die es erlaubt die Häufigkeitsverteilung der Daten in Bezug auf die verschiedenen Klassen darzustellen.

8 Literaturverzeichnis

BAYERISCHES LANDESAMT FÜR GESUNDHEIT UND LEBENSMITTELSICHERHEIT, 2015: GESUNDHEITSATLAS BAYERN. https://www.lgl.bayern.de/gesundheit/gesundheitsberichterstattung/gesundheitsatlas/ia_report/atlas.html, letzter Zugriff 22.01.2018.

BOSTOCK, M. & METCALF, C., 2017: The TopoJSON Format Specification. https://github.com/topojson/topojson-specification, 21.01.2018

BUTLER, H., DALY, M., DOYLE, A., GILLIES, S., HAGEN, S. & SCHAUB, T., 2016: The GeoJSON Format, RFC 7946. https://tools.ietf.org/rfc/rfc7946.txt, letzter Zugriff 22.01.2018

KELLERHOFF, M, 1996: Ein wenig Theorie vorweg … In: Hamburger Projektgruppe Gesundheitsberichterstattung: Praxishandbuch Gesundheitsberichterstattung. Ein Leitfaden für GesundheitsberichterstatterInnen und solche, die es werden wollen. Akademie des öffentlichen Gesundheitswesens in Düsseldorf (Hrsg.), Band 18, 15-23.

KISTEMANN, TH. & SCHWEIKART, J., 2010: Von der Krankheitsökologie zur Geographie der Gesundheit. Geographische Rundschau, **62**(7-8), 4-10.

KUHN, J., WILDNER, M. & ZAPF, A., 2005: Gesundheitsberichterstattung – Ziele, Stand und Perspektiven. Arbeitsmedizin Sozialmedizin Umweltmedizin, **40**(1), 28-34

LANDESZENTRUM FÜR GESUNDHEIT, 2017: GESUNDHEITSATLAS NORDRHEIN-WESTFALEN. https://www.lzg.nrw.de/ges_bericht/ges_indi/gesundheitsatlas_nrw/index.html, letzter Zugriff 22.01.2018

LEDERMANN, F., & GARTNER, G., 2015: mapmap.js: A Data-Driven Web Mapping API for Thematic Cartography. Proceedings of the 27th International Cartographic Conference (ICC2015), Rio de Janeiro, Brasil.

PETERSON, M., 1999: Elements of Multimedia Cartography. Cartwright, W., Peterson, M.P., Gartner, G. (Hrsg.): Multimedia Carography. 31-40.

PIEPER, J . & FRANKE, C., 2017: Kartografische Visualisierung in der Gesundheitsgeographie. In: Augustin, J. und D. Koller (Hrsg.): Geographie der Gesundheit. Die räumliche Dimension von Epidemiologie und Versorgung, 124-145.

SCHARLACH, H. & SCHARLACH, M., 2008: Kartenbasierte Informationssysteme im Gesundheitswesen am Beispiel Niedersachsens. In: Kartographische Nachrichten, **58**(4), 171-178.

SCHARLACH, M. & DREESMAN, J., 2013: Kartographische Darstellungen in der modernen Infektionsepidemiologie. In: Kartographische Nachrichten, **63**(1), 12-19.

WENNBERG, J. & GITTELSOHN, A., 1973: Small Area Variations in Health Care Delivery. Science, **182**, 1102-108.

ZWACKHALS, S.L.N., GIESBERS, H., MAC GILLAVRY, E., VAN BOVEN, P.F. & VAN DER VEEN, A.A., 2004: The Dutch National Atlas of Public Health. In: Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz, **47**(9), 882-889.

ZWACKHALS, S.L.N., GIESBERS, H., BAKEL, A.M. & VAN OERS, J.A.M., 2013: The Dutch National Atlas of Public Health. In: Kartographische Nachrichten, **63**(1), 27-31.

Ein Konzept zur integrierten Modellierung von Detailstufen für 3D-Gebäudemodelle

IZABELA KARUT[1], ANDREAS WICHMANN[1], YEVGENIYA FILIPPOVSKA[1] & MARTIN KADA[1]

Zusammenfassung: 3D-Stadtmodelle werden in unterschiedlichsten Anwendungsgebieten verwendet. Die darin enthaltenen Gebäudemodelle weisen oftmals einen hohen Detaillierungsgrad auf, der jedoch nicht immer benötigt wird. Damit der Detaillierungsgrad den Bedürfnissen der Anwendung entsprechend angepasst werden kann, wurden verschiedene Datenmodelle entwickelt, welche mehrere Detailstufen eines Objektes vorhalten. Die gesonderte Speicherung dieser Detailstufen beinhaltet jedoch zwangsläufig Redundanzen, welche infolge von Aktualisierungsvorgängen zu Inkonsistenzen führen können. In diesem Artikel wird ein Konzept vorgestellt, welches eine integrierte Modellierung mehrerer Detailstufen eines Gebäudes ermöglicht, um Redundanzen zu vermeiden. Die Gültigkeit dieses Konzepts wird anhand eines Testdatensatzes demonstriert.

1 Einleitung

Die Nachfrage nach 3D-Stadtmodellen ist in den letzten Jahren deutlich gestiegen. Sie werden mittlerweile in vielfältigen Anwendungsfeldern wie der Stadt- und Raumplanung, Umweltanalyse und Navigation eingesetzt. Obgleich die darin enthaltenen Gebäude eine immer höhere Detaillierung aufweisen können, entstehen durch die Vielfalt der Anwendungen häufig unterschiedliche Anforderungen diesbezüglich. Aus diesem Grund wurden Datenmodelle wie beispielsweise CityGML (KOLBE et al. 2005) eingeführt, welches es ermöglicht mehrere Detailstufen (LODs) von ein und demselben Gebäude vorzuhalten. Somit kann je nach Aufgabenstellung das für den gewünschten Zweck geeignetste Modell ausgewählt werden. Diesem Konzept liegt jedoch die gesonderte Speicherung jeder vorhandenen Detailstufe zugrunde, was bei verschiedenen geometrischen Ausprägungen eines Gebäudes zwangsläufig zu Redundanzen führt. Die negativen Folgen einer solchen redundanten Speicherung zeigen sich einerseits in einer deutlichen Erhöhung der abzuspeichernden Datenmenge. Andererseits wirkt sich eine redundante Speicherung auch negativ auf den Aufwand für das Management der Daten aus, da u.a. die Vermeidung von Inkonsistenzen infolge von Aktualisierungsvorgängen sichergestellt werden muss. Beide Problematiken werden insbesondere in solchen Fällen noch weiter verschärft, in denen sehr viele diskrete Detailstufen eines Gebäudes mit jeweils nur geringfügigen Änderungen vorgehalten werden.
Um Redundanzen in der abzuspeichernden Datenmenge zu vermeiden, werden in den unterschiedlichsten Disziplinen Modellierungskonzepte zur Speicherung und Darstellung von 3D-Objekten mit mehreren LODs verwendet. In der Computergrafik beispielsweise werden häufig kontinuierliche Detailstufen eingesetzt, um die darzustellenden Details einer Szene zu reduzieren. Dabei werden solche Details eines Objektes ausgeblendet, welche aufgrund der Entfernung zum Beobachter kaum oder nicht mehr wahrnehmbar sind. Um die abzuspeichernde Datenmenge

[1] Technische Universität Berlin, Institut für Geodäsie und Geoinformationstechnik, Straße des 17. Juni 135, D-10623 Berlin, E-Mail: [Vorname.Nachname]@tu-berlin.de

für kontinuierliche Detailstufen gering zu halten, wird jedes darzustellende Objekt nur ein einziges Mal, und zwar in seiner höchsten Detailstufe, abgespeichert. Die Reduzierung von Details erfolgt dann durch die Anwendung unterschiedlicher Flächensimplifizierungsalgorithmen. In Bezug auf die Darstellung von Gebäuden ergeben sich jedoch wesentliche Nachteile in der realistischen Darstellung. So lassen sich zwar unterschiedliche LODs eines Gebäudemodelles erzeugen und abspeichern, jedoch bleiben gebäudecharakteristische Eigenschaften nicht immer zwangsläufig erhalten (KADA et al. 2016).
In diesem Artikel wird ein Konzept zur integrierten Modellierung von Detailstufen für 3D-Gebäudemodelle vorgestellt, das nicht auf eine diskrete Anzahl von Detailstufen limitiert ist. Es basiert auf einer Zellenzerlegung, die das abzuspeichernde Gebäude in disjunkte Zellen aufteilt, sodass jeder Zelle eine oder mehrere Detailstufen zugeordnet werden können. Die Zuordnung mehrerer Detailstufen ermöglicht die wiederholte Verwendung der in der Zelle repräsentierten Geometrie für unterschiedliche Detailstufen. Somit lassen sich Redundanzen in der Datenspeicherung vermeiden. Die Vorteile der vorgestellten Modellierung zeigen sich neben der redundanzfreien Datenhaltung auch in deren Flexibilität. Um dies zu demonstrieren, wird des Weiteren ein Konzept zur Erzeugung von Gebäudemodellen mit gemischten Detailstufen vorgestellt. Es basiert auf der Dekomposition des vom abzuspeichernden Gebäude eingenommenen Raumes und ermöglicht es jedem dieser Teilräume eine eigene Detailstufe zuzuordnen.

2 Relevante Arbeiten

Eines der bekanntesten Datenmodelle zur Speicherung von Gebäuden mit mehreren Detailstufen ist CityGML, das u.a. in KOLBE & GRÖGER (2003) vorgestellt wird. CityGML wird seit 2008 als OGC-Standard zur Repräsentation und zum Austausch virtueller 3D-Stadtmodelle eingesetzt. Als de-facto Standard ist CityGML international anerkannt und wird in vielen Ländern für unterschiedlichste Anwendungsfelder herangezogen (LÖWNER et al. 2016). Es ermöglicht das Vorhalten diskreter Gebäudemodelle in fünf verschiedenen Detailstufen, wobei drei dieser fünf Detailstufen (LOD_1 - LOD_3) bereits in früheren Ansätzen definiert wurden (COORS & FLICK 1998).
Obgleich CityGML als Standard viele Vorteile in Bezug auf die vielseitige Verwendung von Gebäudemodellen bietet, werden bereits in unterschiedlichen wissenschaftlichen Arbeiten die mit den fünf diskreten Detailstufen einhergehenden Einschränkungen diskutiert. Zu den häufigsten Kritikpunkten zählen dabei u.a. die geringe Anzahl an definierten Detailstufen, das Fehlen einer präzisen Definition der einzelnen Detailstufen und das Fehlen von Beziehungen zwischen den verschiedenen Detailstufen (STOTER et al. 2011). Um diesem Umstand Rechnung zu tragen, wurden von BILJECKI et al. (2016) eine Erweiterung des Detailierungskonzepts für 3D-Gebäudemodelle von fünf auf sechzehn Detailstufen vorgeschlagen. Das Problem der redundanten Datenhaltung durch den Anstieg auf sechzehn Detailstufen wird dadurch jedoch nur noch weiter verschärft. Eine ähnliche Problematik bezüglich einer redundanzfreien Datenhaltung weist auch die in BENNER et al. (2013) vorgestellte Modellierung von 3D-Stadtmodellen auf, bei der zwischen einer geometrischen Detailebene und einer semantischen Detailebene unterschieden wird.
Neben der Verwendung von diskreten Detailstufen werden beispielsweise in der Computergrafik auch kontinuierliche Detailstufen verwendet. Dabei wird oftmals nur die höchste Detailstufe ei-

nes Modells als Dreiecksnetz abspeichert und dieses dann durch Anwendung von Methoden der Flächensimplifizierung iterativ vereinfacht. Ein häufig eingesetztes Flächensimplifizierungsverfahren ist zum Beispiel das Edge-Collapse-Verfahren. Hierbei werden zwei durch eine Kante verbundene Knoten zu einem einzigen Knoten zusammengeführt, sodass die ursprünglich vorhandene Kante aus dem Modell entfernt wird. Durch die iterative Anwendung mehrerer Edge-Collapse-Operationen können so Details aus dem zu repräsentierenden Objekt kontinuierlich entfernt werden. Varianten dieses Flächensimplifizierungsverfahrens werden in HOPPE (1996), GUÉZIEC (1995) und RONFARD & ROSSIGNAC (1996) näher beschrieben und unterscheiden sich im Wesentlichen durch die Selektion der zu reduzierenden Kanten. Ein weiteres Flächensimplifizierungsverfahren stellt das Vertex-Clustering-Verfahren dar, das eine Menge nahe beieinanderliegender Knoten zu einem einzigen Knoten vereint; siehe z.B. ROSSIGNAC & BORREL (1993) und (SCHAEFFER 2007). Einen umfassenden Überblick über weitere Methoden der Flächensimplifizierung bieten LUEBKE (1997) und BOTSCH et al. (2010). Ein wesentlicher Vorteil dieser Konzepte ist, dass Objekte nur in ihrer höchsten Detailstufe zusammen mit einer Simplifizierungsvorschrift vorgehalten werden müssen, wodurch sich eine redundanzfreie Speicherung ergibt. Eine Datenstruktur, die sich für die Speicherung kontinuierlicher Detailstufen und insbesondere auch allgemein für die Ergebnisse eines Generalisierungsprozesses eignet, ist beispielsweise die tGap-Datenstruktur (VAN OOSTEROM 2005). In dieser werden kontinuierliche Detailstufen in einer Baumstruktur repräsentiert, wobei in den Blättern die höchste Detailstufe und in den Elternknoten nur die Generalisierungsmethode für die Kindknoten und nicht die Geometrie selbst repräsentiert ist. Auch wenn mit solchen Konzepten generell eine redundanzfreie Speicherung ermöglicht wird, so ist jedoch ein wesentlicher Nachteil, dass für die Generierung einer geringen Detailstufe oftmals zunächst alle Repräsentationen höheren Details erzeugt werden müssen.

3 Konzeptuelles Modell

In dem hier vorgestellten Konzept zur integrierten Modellierung von 3D-Gebäudemodellen wird auf die gesonderte Speicherung jeder Detailstufe verzichtet. Dazu werden alle einem Gebäude zugehörigen LODs mittels einer Zellenzerlegung in eine Menge von disjunkten Komponenten zerlegt. Diese Komponenten repräsentieren jeweils eine Teilmenge, die in einer oder mehreren Detailstufen vorkommt. Dabei werden die Komponenten in eine Basiskomponente und mehrere Detailstufenkomponenten unterteilt. Die Basiskomponente repräsentiert den Teil, den alle LODs gemeinsam haben, während die Detailstufenkomponenten nur Teil bestimmter Detailstufen sind. Eine genaue Definition zur Generierung dieser Komponenten wird in Abschnitt 3.1 gegeben.
Für die Rückgewinnung der ursprünglichen Detailstufen eines Gebäudes wird in Abschnitt 3.2 ein Verfahren vorgestellt, das die Basiskomponente des wiederherzustellenden Gebäudes mit den notwendigen Detailstufenkomponenten vereinigt.
Um die Flexibilität und Vorteile des hier vorgestellten integrierten Modells für verschiedenste Anwendungen zu demonstrieren, wird eine Erweiterung für die Rückgewinnung von gemischten Detailstufen vorgestellt. Dazu wird der vom Gebäude eingenommene Raum in disjunkte Zellen unterteilt und jeder dieser Zellen eine der Anwendung erforderlichen Detailstufe zugewiesen. Die gemischte Detailstufe ergibt sich dann durch die in Abschnitt 3.3 vorgestellte Vereinigung.

Somit lassen sich nicht nur die ursprünglichen Detailstufen wiederherstellen, sondern auch verschiedenste Detailstufenkombinationen generieren.

3.1 Basiskomponente & Detailstufenkomponenten

Basierend auf den standardisierten Grundlagen der Objektmodellierung werden Objekte im weiteren Verlauf als Teilmengen eines topologischen Raums betrachtet, wodurch mengentheoretische Boolesche Operatoren angewendet werden können. Sei dazu $S = \{LOD_1, \dots, LOD_n\}$ eine Menge von n Repräsentationen eines Gebäudes in unterschiedlichen Detailstufen und C die Menge der durch das integrierte Modell erzeugten Komponenten. Jedes Element $C_I \in C$ mit $I \subseteq \{1, \dots, n\} = N$ ist dann definiert durch:

$$C_I = \bigcap_{i \in I} LOD_i \setminus \bigcup_{j \in N \setminus I} LOD_j \tag{1}$$

Im Folgenden wird aus Gründen der Lesbarkeit auf Mengenklammern in Indizes verzichtet. Unter Berücksichtigung von Gleichung (1) wird folglich für die Erzeugung einer Komponente die Schnittmenge der LODs in S gebildet, deren Index ein Element im Index der zu erzeugenden Komponente ist. Anschließend wird von dieser Schnittmenge die Vereinigung der in S verbliebenen Elemente subtrahiert. Die Komponente bei der beispielsweise $I = N$ gilt repräsentiert die Komponente $C_{1,\dots,n}$. Da diese Komponente in allen Detailstufen enthalten ist, wird sie im Folgenden als Basiskomponente bezeichnet und ist dementsprechend allen LODs in S zugeordnet. Die übrigen Komponenten, die somit Elemente der Menge $C \setminus C_N$ sind, bilden die Menge der Detailstufenkomponenten. Im Gegensatz zur Basiskomponente ist eine Detailstufenkomponente nicht jedem LOD zugeordnet, sondern nur einer echten Teilmenge von S. Die Zuordnung einer Detailstufenkomponente zu bestimmten LODs ergibt sich durch deren Index I. So ist beispielsweise die Komponente $C_{1,3,9}$ den LODs der Menge $\{LOD_1, LOD_3, LOD_9\}$ zugeordnet.

Um eine redundanzfreie Speicherung der LODs zu ermöglichen, müssen die Komponenten in C disjunkt sein. Dies wird durch die Differenzbildung in Gleichung (1) sichergestellt, bei der einzig im Falle der Basiskomponente der zweite Operand der Differenz der leeren Menge entspricht. Somit beinhaltet jede Detailstufenkomponente nur solche Details, die ausschließlich in allen mit dieser Komponente assoziierten LODs repräsentiert sind. Generell ist bei der Anwendung von Gleichung (1) zur Erzeugung der Komponenten festzustellen, dass die Anzahl der zu berechnenden Kombinationen exponentiell mit der Anzahl der Detailstufen wächst. Dies ist darauf zurückzuführen, dass die Anzahl an zu erzeugenden Komponenten durch die Elementanzahl der Potenzmenge $\mathcal{P}$ von S gegeben ist, sodass $|C| = |\mathcal{P}(S)| = 2^n$ gilt. In der Praxis jedoch entspricht ein großer Anteil der Elemente in C der leeren Menge, sodass diese Komponenten nicht weiter berücksichtigt oder abgespeichert werden müssen.

In den folgenden Abbildungen wird exemplarisch an einem konkreten Beispiel für die abzuspeichernden LODs der Menge $S = \{LOD_1, LOD_2, LOD_3, LOD_4\}$ aufgezeigt, wie die Basiskomponente $C_{1,2,3,4}$ und die Detailstufenkomponenten C_1 und $C_{2,3,4}$ anhand von Gleichung (1) generiert werden. In Abbildung 1 ist die Erstellung der Basiskomponente dargestellt. Der erste Operand der Differenz bildet die Schnittmenge aller Detailstufen eines Gebäudemodells ab. Da jede vorhandene Detailstufe bereits in den ersten Operanden eingeht, bildet der zweite Operand die leere Menge. Somit ergibt sich die Basiskomponente aus der Schnittmenge aller Detailstufen eines

Gebäudemodells. Diese Aussage trifft nicht nur auf das aufgezeigte Beispiel zu, sondern ist allgemeingültig für die Generierung von Basiskomponenten.

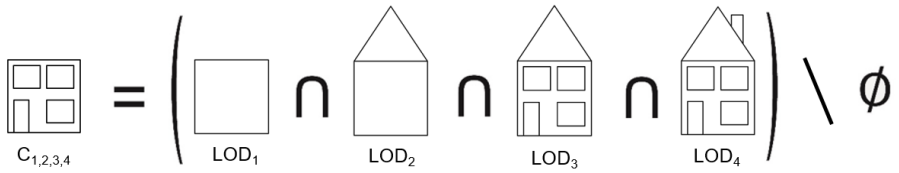


Abb. 1: Generierung der Basiskomponente $C_{1,2,3,4}$ aus den vier Detailstufen eines Gebäudes basierend auf Gleichung (1)

Die Generierung der Detailstufenkomponente C_1 ist in Abbildung 2 dargestellt. Im Gegensatz zur Generierung der Basiskomponente besteht hier die zu verschneidende Menge nicht aus allen Komponenten, sondern, dem Komponentenindex I folgend, aus der Menge $\{LOD_1\}$. Dadurch müssen im zweiten Operanden der Differenz die restlichen Detailstufen LOD_2, LOD_3 und LOD_4 vereinigt und anschließend von LOD_1 abgezogen werden. Dadurch repräsentiert die Detailstufenkomponente C_1 die leere Menge und muss somit nachfolgend nicht weiter berücksichtigt werden.

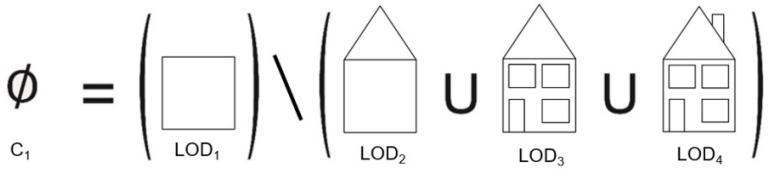


Abb. 2: Generierung der Detailstufenkomponente C_1 basierend auf Gleichung (1)

Die Generierung der Detailstufenkomponente $C_{2,3,4}$ wird in Abbildung 3 illustriert. Der erste Operand der Differenz ergibt sich in diesem Fall aus der Verschneidung der Menge $\{LOD_2, LOD_3, LOD_4\}$, wodurch der zweite Operand aus der einelementigen Menge $\{LOD_1\}$ besteht. Aus den Elementen des ersten Operanden wird die Schnittmenge gebildet und anschließend der zweite Operand abgezogen. Das Ergebnis liefert die Detailstufenkomponente $C_{2,3,4}$, welche ausschließlich für die Generierung der Detailstufen LOD_2, LOD_3 und LOD_4 benötigt wird.

Abb. 3: Generierung der Detailstufenkomponente $C_{2,3,4}$ basierend auf Gleichung (1)

Werden die Basiskomponente und alle Detailstufenkomponenten des vorangegangenen Beispiels mit dem Konzept des vorgestellten integrierten Modells generiert, so müssen insgesamt sechzehn Komponenten berechnet werden. Die Anzahl der zu berechnenden Komponenten ergibt sich aus der Potenzmenge der vier vorhandenen LODs. In diesem Beispiel entsprechen zwölf der sech-

zehn Komponenten der leeren Menge. Somit müssen insgesamt nur die in Abbildung 4 dargestellten vier Komponenten bestehend aus der Basiskomponente $C_{1,2,3,4}$ und den drei Detailstufenkomponenten $C_{1,2}$, $C_{2,3,4}$ und C_4 abgespeichert werden.

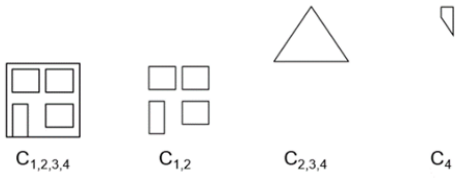


Abb. 4: Die Basiskomponente und Detailstufenkomponenten für das Beispielgebäude in vier verschiedenen Detailstufen

3.2 Rückgewinnung von Detailstufen

Die Rückgewinnung einer Detailstufe erfolgt durch das Zusammenfügen sämtlicher Komponenten, die mit der jeweiligen Detailstufe assoziiert sind. Sei eine Menge $S = \{LOD_1, \dots, LOD_n\}$ gegeben, die entsprechend Abschnitt 3.1 in Basiskomponente und Detailstufenkomponenten vorliegt, so ist die Rückgewinnung einer Detailstufe LOD_k mit $k \in \{1, \dots, n\} = N$ wie folgt definiert:

$$LOD_k = \bigcup_{I \mid k \in I \subseteq N} C_I \tag{2}$$

Da für die Basiskomponente $I = \{1, \dots, n\}$ und somit für alle $k \in I$ gilt, wird sie für die Rückgewinnung jeder Detailstufe benötigt. Detailstufenkomponenten werden hingegen entsprechend ihrem Index I einzig für die Rückgewinnung einer oder mehrerer aber nicht für alle Detailstufen benötigt. Die Basiskomponente wird Gleichung (2) folgend zusammen mit den entsprechenden Detailstufenkomponenten vereinigt, wodurch sich das Gebäudemodell einer bestimmten Detailstufe ergibt. Eine positive Eigenschaft des hier vorgestellten Konzepts zur integrierten Modellierung von Detailstufen und deren Rückgewinnung ist, dass jede Detailstufe direkt erzeugt werden kann, ohne dass Abhängigkeiten zu anderen Detailstufen berücksichtigt werden müssen. Ein Beispiel für die Rückgewinnung der Detailstufe LOD_2 für das in Abschnitt 3.1 eingeführte Gebäude mit vier Detailstufen ist in Abbildung 5 dargestellt. Hierzu wird die Basiskomponente $C_{1,2,3,4}$ mit den zwei Detailstufenkomponenten $C_{1,2}$ und $C_{2,3,4}$ vereinigt.

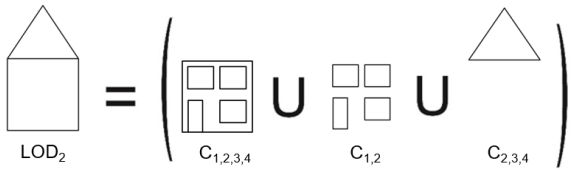


Abb. 5: Rückgewinnung der Detailstufe LOD_2 durch die Vereinigung der Basiskomponente mit den entsprechenden Detailstufenkomponenten

In Abbildung 6 ist ein weiteres Beispiel für die Rückgewinnung einer Detailstufe (LOD_3) dargestellt. Da nur zwei der in Abbildung 4 dargestellten Komponenten im Index I das Element 3

aufweisen, ist für die Rückgewinnung dieser Detailstufe die Vereinigung der Basiskomponente mit der Detailstufenkomponente $C_{2,3,4}$ ausreichend.

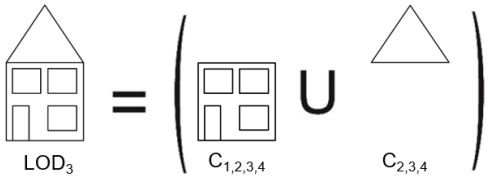


Abb. 6: Rückgewinnung der Detailstufe LOD_3 durch die Vereinigung der Basiskomponente mit der Detailstufenkomponente $C_{2,3,4}$

3.3 Rückgewinnung von gemischten Detailstufen

Um ein Gebäudemodell auch in gemischten Detailstufen darstellen zu können, wird der vom Gebäude eingenommene Raum zunächst so in überlappungsfreie Teilräume aufgeteilt, dass sie nach MÄNTYLÄ (1988) der Definition eines Dekompositionsmodells folgend den zu modellierenden Raum vollständig repräsentieren. In der Praxis setzt sich das Dekompositionsmodell häufig aus Teilräumen einfacher Geometrien (z.B. Quader) zusammen, die beispielsweise die Stockwerke oder Treppenhausbereiche des Gebäudes repräsentieren. Jedem dieser Teilräume wird dann, in Abhängigkeit von der späteren Verwendung des Gebäudemodells, eine eigene Detailstufe zugewiesen, sodass eine Vielzahl unterschiedlichster Dekompositionsmodelle konstruiert werden kann.
Sei für die Rückgewinnung von gemischten Detailstufen ein Dekompositionsmodell $D = \{d_1, \dots, d_m\}$ mit m Teilräumen und eine Funktion f gegeben, welche den Index der mit dem Element $d \in D$ assoziierten Detailstufe aus der Menge $S = \{LOD_1, \dots, LOD_n\}$ liefert, so ist die Rückgewinnung einer gemischten Detailstufe wie folgt definiert:

$$LOD_{gemischt} = \bigcup_{d \in D} \bigcup_{I | f(d) \in I} d \cap C_I \tag{3}$$

Durch die Anwendung dieser Gleichung auf eine gegebene Menge S an Detailstufen und einem Dekompositionsmodel D können Gebäude entsprechend dem jeweiligen Anwendungszweck mit individuell gemischten Detailstufen generiert werden.

4 Ergebnisse & Diskussion

Die Gültigkeit des vorgeschlagenen Konzepts wurde exemplarisch anhand eines manuell konstruierten 3D-Straßenblocks mit insgesamt sechs Gebäudeobjekten verifiziert (vgl. Abbildung 7). Aus den insgesamt vier modellierten Detailstufen wurden die Basiskomponente und die Detailstufenkomponenten anhand von Gleichung (1) gemäß der dargestellten Methodik erzeugt. Jede dieser Komponenten besteht aus einer Sammlung von Objekten, die in ihrer Gesamtheit disjunkt sind. Unter Verwendung von Gleichung (2) können mit diesen Komponenten erwartungsgemäß die ursprünglichen Detailstufen rückgewonnen werden, was in dem dargestellten Beispiel durch die vier Detailstufen gezeigt wird.

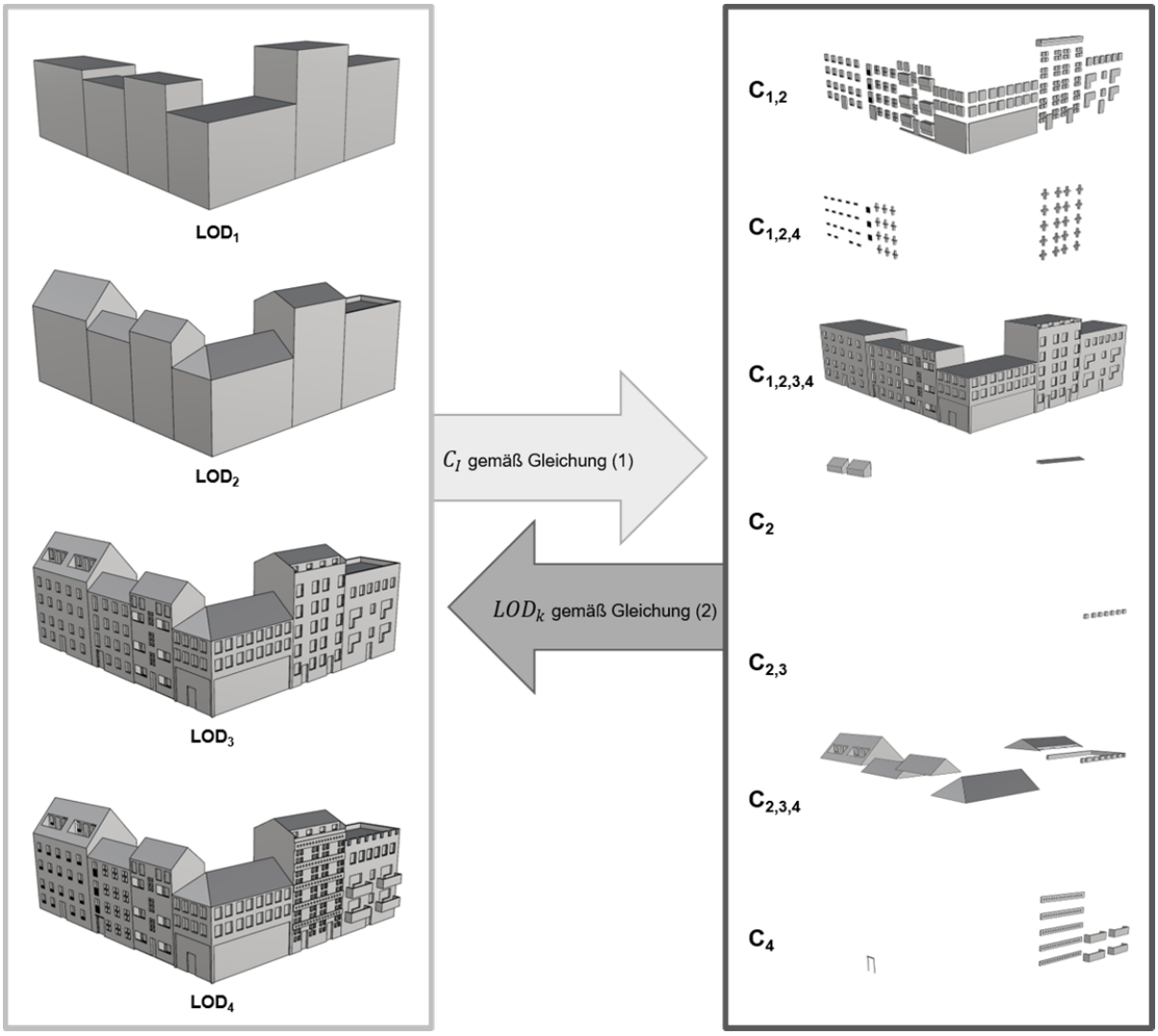


Abb. 7: Ein Gebäudeblock in vier Detailstufen und die sich durch das integrierte Modell ergebenen Komponenten

Ein Beispiel zur Generierung von Gebäuden in gemischten Detailstufen ist in Abbildung 8 dargestellt. Hierzu wurde ein Dekompositionsmodell erstellt, das aus neun Teilräumen besteht. Diese repräsentieren die Aufteilung der einzelnen Stockwerke des Gebäudes und unterteilen, mit Ausnahme des Daches, jedes Stockwerk in einen linken und rechten Wohnbereich. Jedem dieser Teilräume wurde dann ein Detaillierungsgrad zugewiesen. Anschließend wurden die mit dem Dekompositionsmodell assoziierten Komponenten gemäß Gleichung (3) vereinigt und zu einem Gesamtmodell zusammengeführt. Das Ergebnis des zusammengeführten Gesamtmodells ist das in Abbildung 8 dargestellte Gebäude in gemischten Detailstufen. Durch diese Flexibilität wird eine individuelle Anpassung eines Gebäudes ermöglicht, die auf eine spezifische Aufgabenstellung angepasst werden kann.

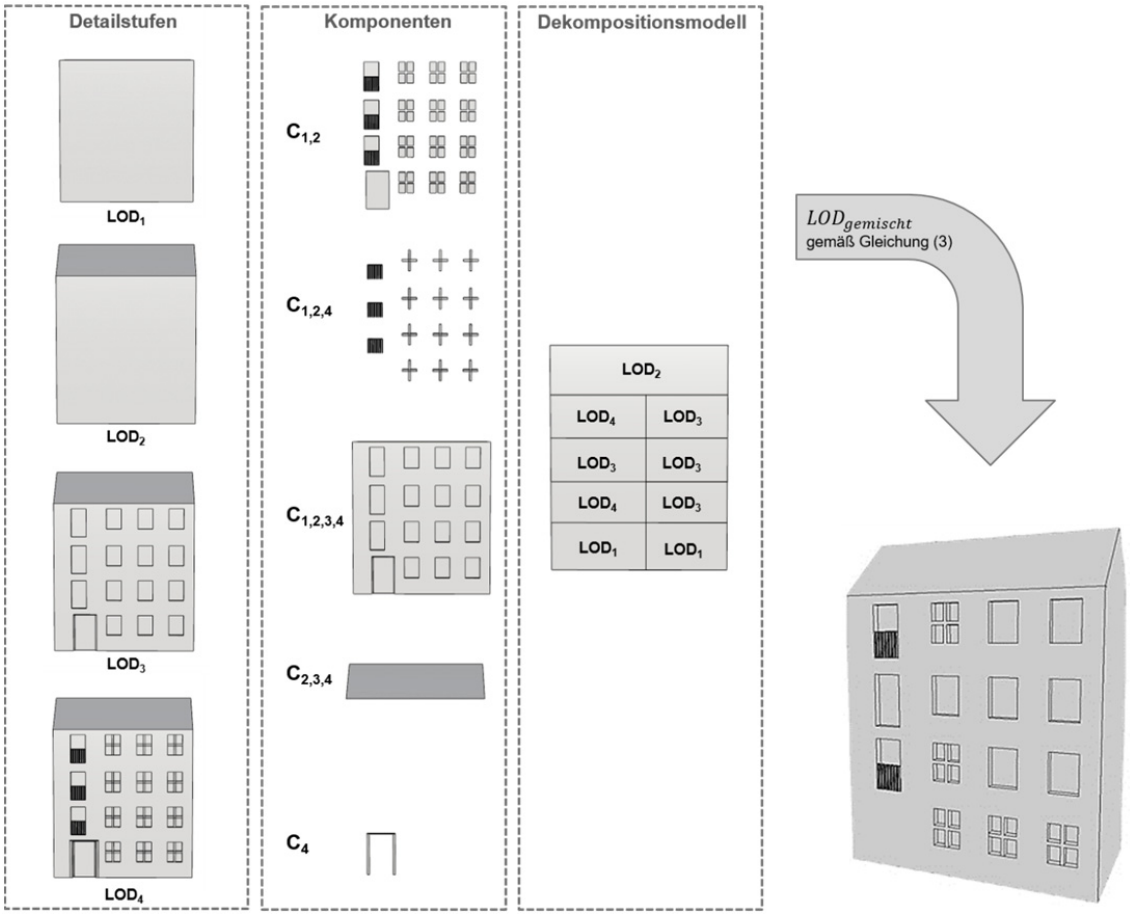


Abb. 8: Erstellung eines Gebäudes in Mischformdarstellung unter Verwendung der entsprechenden Komponenten für jede Zelle des Dekompositionsmodells

Die durchgeführten Experimente belegen, dass durch die Umsetzung des integrieren Modells nicht jede Detailstufe einzeln abgespeichert werden braucht und dass dadurch eine redundante Datenhaltung vermieden werden kann. Jedoch zeigt sich, dass durch das integrierte Modell neue Flächen eingeführt wurden, die nicht in den ursprünglichen Detailstufen vorhanden waren. Diese Flächen sind dadurch begründet, dass die Objekte einer Komponente als geschlossene Festkörper abgespeichert werden. Bei der Rückgewinndung der ursprünglichen Detailstufen fallen diese durch die Vereinigung weg. Verwendung können diese Flächen jedoch bei der Generierung gemischter Detailstufen finden.

Ein weiterer Aspekt, den es auch zu untersuchen gilt, ist die realistische Darstellung eines Gebäudemodells in gemischten Detailstufen. So sollte das Dekompositionsmodell eine Aufteilung aufweisen, die der geometrischen Zusammensetzung des Gebäudes entspricht. Wird dies nicht eingehalten, so kann dies in Bezug auf Fassaden, Außenkomponenten und Dachstrukturen folgende negative Auswirkungen haben. Fenster und Türen einer Fassade können beispielsweise in mehrere Zellen unterteilt werden, wodurch die Darstellung in unterschiedlichen Detailstufen erfolgen kann, was wiederum das Gebäude realitätsfremd erscheinen lassen kann. Ähnlich verhält es sich auch wenn Außenkomponenten (z.B. Feuertreppen), Verbindungen (z.B. Bautenverbindungsbrücken) zwischen verschiedenen Gebäuden und Dachstrukturen in mehreren unter-

schiedlichen Detailstufen repräsentiert werden sollen. Übergänge zwischen den einzelnen Detailstufen können aber nicht nur aus geometrischer Sicht unrealistisch wirken, sondern können vor allem auch bei texturieren Objekten erkennbar auftreten.

5 Fazit & Ausblick

Das hier vorgestellte integrierte Modell ermöglicht die redundanzfreie Speicherung von Gebäudemodellen in unterschiedlichen Detailstufen. Die einzelnen Detailstufen dieser Gebäude können dabei fehlerfrei und ohne Verlust von Details direkt rückgewonnen werden, ohne dass Gebäudemodelle mit einer höheren Detailstufe zuvor erzeugt werden müssen, wie es beispielsweise bei den kontinuierlichen Detailstufen oftmals der Fall ist. Die vorgestellte Erweiterung des integrierten Modells ermöglicht es zudem gemischte Detailstufen anwendungsspezifisch zu generieren und gestattet dadurch eine sehr flexible Darstellung von Gebäuden.
Obgleich das integrierte Modell eine effiziente Speicherung bezüglich redundanzfreier Datenhaltung ermöglicht, besteht noch weiterer Forschungsbedarf im Hinblick auf den Speicherbedarf und möglichen realitätsfernen Darstellungen von Gebäuden in gemischten Detailstufen. So müssen für die im experimentellen Teil diskutierten Fehldarstellungen von Außenkomponenten noch Lösungsansätze entwickelt werden. Zum Beispiel könnte im Fall einer Feuertreppe der Detailgrad stets unabhängig von der Raumaufteilung des Dekompositionsmodells einheitlich definiert werden. Dadurch kann vermieden werden, dass eine Feuertreppe, die drei Stockwerken zugeordnet ist, unterbrochen wird, indem das mittlere Stockwerk in einer geringeren Detailstufe als das darüber- und darunterliegende repräsentiert wird. Ähnliche Überlegungen sind auch für Verbindungen zwischen verschiedenen Gebäuden anzustellen. Des Weiteren sollte zur Vermeidung einer unrealistischen Darstellung von Dachstrukturen darauf geachtet werden, dass nebeneinanderliegende Dachgeschosswohnungen in Bezug auf Fenster und Türen eine einheitliche Detaillierung aufweisen und dass das Dach entweder vollständig oder überhaupt nicht abgebildet wird. Sollte letzteres nicht berücksichtigt werden, so können beispielsweise Gebäudemodelle mit einem halben Dach entstehen.
Des Weiteren wäre eine automatische Generierung von Dekompositionsmodellen im Hinblick auf den Bearbeitungsaufwand wünschenswert, jedoch müssen dabei geometrische Eigenschaften eines Gebäudes zwangsläufig mitberücksichtigt werden.
Schließlich besteht auch noch weiterer Forschungsbedarf bezüglich der Rückgewinnungsmethode von Gebäudemodellen. So können die verwendeten komplexen Booleschen Operationen im 3D-Raum durch einfachere Klebeoperation, wie sie z.B. in MÄNTYLÄ (1988) vorgestellt wurden, ersetzt werden. Wie diese Klebeoperationen für das integrierte Modell umzusetzen sind, bleibt noch zu untersuchen.

6 Literaturverzeichnis

BENNER, J., GEIGER, A., GRÖGER, G., HÄFELE, K.H. & LÖWNER, M.O., 2013: Enhanced LOD Concepts for Virtual 3D City Models. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **II**(2/W1), 51-61.

BILJECKI, F., LEDOUX, H. & STOTER, J., 2016: An Improved LOD Specification for 3D Building Models. Computers, Environment and Urban Systems, **59**, 25-37.

BOTSCH, M., KOBBELT, L., PAULY, M., ALLIEZ, P. & LÉVY, B., 2010: Polygon Mesh Processing. CRC Press, Baca Raton, FL, USA.

COORS, V. & FLICK, S., 1998: Integrating Levels of Detail in a Web-based 3D-GIS. Proceedings of the 6th ACM International Symposium on Advances in Geographic Information Systems, 40-45.

GUÈZIEC, A., 1995: Surface Simplification with Variable Tolerance. Proceedings of the 2nd Annual International Symposium on Medical Robotics and Computer Assisted Surgery **'95**, 132-139.

HOPPE, H., 1996: Progressive Meshes. SIGGRAPH '96 Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, 99-108.

KADA, M., WICHMANN, A., FILIPPOVSKA Y. & HERMES T., 2016: Animation Strategies for Smooth Transformations Between Discrete LODs of 3D Building Models. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **41**(B2), 413-420.

KOLBE, T.H. & GRÖGER, G., 2003: Towards Unified 3D City Models. Proceedings of the Joint ISPRS Commission IV Workshop on Challenges in Geospatial Analysis.

KOLBE, T.H., GRÖGER, G. & PLÜMER L., 2005: CityGML- Interoperable Access to 3D City Models. Proceedings of the International Symposium on Geo-information for Disaster Management, 883-899.

LÖWNER, M.-O., GRÖGER, G., BENNER, J., BILJECKI, F. & NAGEL, C., 2016: Proposal for a New LOD and Multi-representation Concept for CityGML. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **IV**(2/W1), 3-12.

LUEBKE, D., 1997: A Survey of Polygonal Simplification Algorithms. UNC Chapel Hill Computer Science Technical Report, **TR97-045**.

MÄNTYLÄ, M., 1988: An Introduction to Solid Modeling. Computer Science Press, Rockville, Maryland.

RONFARD, R. & ROSSIGNAC, J., 1996: Full-range Approximation of Triangulated Polyhedra. Computer Graphics Forum, **15**, 67-76.

ROSSIGNAC, J. & BORREL, P., 1993: Multi-resolution 3D Approximations for Rendering. Modelling in Computer Graphics, 455-465.

SCHAEFFER, S.-E., 2007: Graph Clustering. Computer Science Review, **I**, 27-64.

STOTER, J., VOSSELMANN, G., GOSS, J., ZLATANOVA, S., VERBREE, E., KLOOSTER, R. & REUVERS, M., 2011: Towards a National 3D Spatial Data Infrastructure: Case of the Netherlands. Photogrammetrie - Fernerkundung - Geoinformation, 405-420.

VAN OOSTEROM, P., 2005: Variable-scale Topological Data Structure Suitable for Progressive Data Transfer: The GAP-face Tree and GAP-edge Forest. Cartography and Geographic Information Science, **32**(4), 331-346.

Mixed Reality Anwendungen mit 3D-Stadtmodellen

MARTIN CHRISTEN[1], URS CLEMENT[1] & ADRIAN MEYER[1]

Zusammenfassung: Mit einer Augmented Reality App soll der Bevölkerung der Stadt Basel einen neuen Einblick geboten werden. Dabei werden 2D-Karten und auch das physikalische Stadtmodell überlagert. In vier Modulen wird der offizielle Stadtplan mit interaktiven Grafiken bereichert. Zunächst kann das komplette 3D-Stadtmodell überlagert werden, dabei können unterschiedliche Layer wie beispielsweise Fahrradrouten oder Müllabfuhrzonen aktiviert werden. Auch die Visualisierung von Untergrunddaten, wie beispielsweise der Abwasserentsorgung, kann eingeblendet werden. Das zweite Modul beinhaltet die Adresssuche, verbunden mit Informationen zu Quartieren und Sehenswürdigkeiten. Ein weiteres Modul beinhaltet die Verkehrssimulation, bei der Züge, Trams und Schiffe auf der Karte überblendet werden können. Im letzten Modul werden aktuelle Wetterinformationen augmentiert. Dabei werden unter anderem auch Wolken und Regen mit volumetrischen Ansätzen visualisiert, dessen Datengrundlagen von Live-Wetterdaten stammen. In diesem Beitrag wird auch gezeigt, wie eine solche Applikation entwickelt wird und wie die Qualität der 3D-Visualisierung durch globale Beleuchtungsansätze verbessert werden kann.

1 Mixed Reality Anwendungen auf Smartphones

Augmented Realty (AR, erweiterte Realität) hat in den letzten Jahren dank Smartphones deutlich an Popularität gewonnen. Mit Augmented Reality werden virtuelle Welten mit Abbildungen realer Objekte ergänzt. Erste Entwicklungen in diesem Bereich waren noch mit aufwändiger Hardware realisiert (SUTHERLAND 1965). Dank der heutigen Rechenpower und der integrierten Sensonen bei Smartphones (COSTELLO 2016) ist Augmented Reality mit stabilen Tracking auf handelsüblichen Geräten möglich. Inzwischen liefern sogar die Smartphone Hersteller entsprechende APIs mit, so hat Apple das ARKit vorstellt (APPLE INC. 2017) und Google ARCore (KLING 2017). In der Definition Augmented Reality von AZUMA (1997) werden reale und virtuelle Inhalte kombiniert, und zwar interaktiv in Echtzeit und ist im dreidimensonalen Raum referenziert.

MILGRAM (1994) erwähnte den Begriff **Mixed Reality** im Rahmen des Konzepts des Virtualitätskontinuums. Mixed Reality umfasst das gesamte Realitäts-Virtualitäts Kontinuum, dass die virtuellen Objekte mit der realen Welt interagieren. Dies wurde auch bereits in zahlreichen Applikationen konkret umgesetzt wie beispielsweise in den kürzlich präsentierten Arbeiten von KERVEGANT et al. (2017), PIUMSOMBOON et al. (2017) und RODRIGUES et al. (2017). Weitere aktuelle Applikationen im Bereich zu AR finden sich auch in SCHMID (2017) und LEE (2017). Ein ausführlicher Bericht über Entwicklungen und Technologien im Bereich von Augmented Reality wurde kürzlich von BILLINGHURST (2015) erstellt.

Vuforia ist ein in vielfacher Hinsicht beliebtes Framework für Marker-basierte Augmented und Mixed Reality Anwendungen (LOESCH et al. 2015). Es umfasst ein Software Development Kit

[1] Fachhochschule Nordwestschweiz, Hochschule für Architektur, Bau und Geomatik, Institut Geomatik, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: [martin.christen, urs.clement, adrian.meyer]@fhnw.ch

(SDK) mit vollständiger Integration in Unity. Dadurch ergibt sich ein hohes Maß an Kompatibilität gegenüber iOS- und Android-Geräten, sowie mit der Wearable Platform Microsoft Hololens. Bei Vuforia werden jedoch in der Regel keine hardware-basierte Sensorinformationen wie GPS, Kompass und IMU/Accelerometer in die Lageberechnung miteinbezogen. Weil Vuforia durch fortgeschrittene Computer Vision Technik eine im Vergleich sehr robuste Markerdetektion für zweidimensionale Targets aufweist und eine Reihe von Konfigurationsmöglichkeiten für Entwickler bereitstellt, hat sich die Technologie (laut der Mutterfirma PTC) zur Zeit als AR-Marktführer positioniert (SHANDWICK AND WOWTSCHERK 2015). Auch für Ausstellungen und museale Settings, wie sie sich bei diesem Projekt finden, wurde Vuforia schon erfolgreich angewendet (CHRISTEN et al. 2016; LOESCH et al. 2015; PENG AND ZHAI 2017; WÜEST et al. 2016).
Vuforia wird derzeit von vielen Entwicklern als AR SDK bevorzugt, da es bis zu einem bestimmten Rahmen im nicht kommerziellen Bereich gratis verfügbar ist und eine große Bandbreite an Features enthält (SCHMIDT 2017). Außerdem bietet die Unterstützung von Unity3D für viele Entwickler eine bereits gewohnte Umgebung und relativ umfangreiche Möglichkeiten zur Portierung vorhandener Apps und Spiele. Der wohl stärkste Konkurrent war bis vor kurzem das Metaio SDK. Es bietet ein ähnlich stabiles Tracking mit vielen Features und kann durch die Unity Integration auf einer großen Bandbreite von Geräten angewendet werden. Ferner existieren ähnlich wie bei Vuforia eine große Community und ausreichend Dokumentation. Der Preis von 5500,- USD für Apps ohne «Watermarking» zusammen mit einer mehrfach beschriebenen «steep learning curve» ist jedoch relativ hoch angesetzt. Einige Entwickler bescheinigen Metaio eine bessere 3D-Marker-Erkennung im direkten Vergleich mit Vuforia. Seit der Übernahme von Metaio durch Apple im Jahr 2015 wird die eigene Software jedoch nicht mehr weiterentwickelt. Es ist anzunehmen, dass das Knowhow der Firma heute sowohl für ARKit, als auch für das FaceID-Verfahren verwendet wird (WIESEND 2015).
Ein weiterer Anbieter einer AR SDK ist Wikitude. Wikitude ermöglicht die Erkennung von 3D-Targets mittels vollintegrierter SLAM-Algorithmen und hat eine relativ große Nutzerbasis. Um Zugriff auf die Objekterkennungsbibliothek zu erhalten, ist jedoch ein Abonnement mit jährlichen Kosten von ca. 3000,- EUR notwendig (WIKITUDE 2017).
Weitere neue Marktteilnehmer mit einer Möglichkeit zur markerbasierten Objekterkennung sind «Total Immersion», «Arpa Solutions», «13th Lab», «ARToolkit» und «Orbotix». Aufgrund hoher Initialkosten, unzureichenden Tutorials/Community-Feedback, fehlender Unity-Unterstützung, instabilem Tracking oder einer schwierigen Einarbeitungsphase erschien eine Anwendung dieser Plattformen für die vorliegende Arbeit im gegebenen Rahmen als unrealistisch (MULA 2015).
Apples ARKit ist ein Software-Framework, das zur Verwendung in modernen iOS-Geräten (Chipsatz mindestens A9) mit dem Update auf iOS 11 ausgespielt wurde. Das Magazin «Apple Insider» ging im Oktober 2017 von einer Nutzerbasis (geschätzt an der Anzahl Installationen) von 350 Millionen Endgeräten aus (DILGER 2017). Das Verfahren der Umgebungsdetektion mittels der Fusion aus bewegungsphotogrammetrischen Bilddaten der rückseitigen Kamera und IMU Sensorik wird «Visual Inertial Odometry» genannt (APPLE INC. 2017). Erwähnenswert ist die Tatsache, dass das Framework ohne Kalibrierung auskommt, horizontale Flächen wie Tische und Böden erkennt, sowie eine Einfallswinkel-Einschätzung des Umgebungslichts ermöglicht. Bereits bei der erstmaligen Vorstellung der Software auf der WWDC2017 wurde außerdem eine

Unterstützung der Unity-Entwicklungsumgebung angekündigt. Derzeit wird die Unity-Integration noch als «experimentell» eingestuft; seit Juni 2017 gab es bereits 11 kleinere Releases (UNITY TECHNOLOGIES 2017). Eine komfortable, markerbasierte Mixed-Reality-Umsetzung außerhalb der frontseitigen Gesichtserkennung fehlt dem Framework jedoch noch. Eine Integration mit Vuforia oder eine andere Methode zur Objekterkennung, die über die Detektion horizontaler Oberflächen hinausgehen würde, ist derzeit noch nicht offiziell verfügbar. Apple forciert momentan noch die Entwicklung von ARKit-Apps mittels der eigenen Entwicklungsplattform Xcode, wobei markerbasierte AR derzeit nicht im Vordergrund zu stehen scheinen (LING 2017). Durch den Zukauf von Metaio im Jahr 2015 ist jedoch davon auszugehen, dass Apple das notwendige Knowhow für markerbasierte AR besitzt und noch an der Umsetzung arbeitet (WIESEND 2015). Einige Entwickler setzen sich jedoch seit der Lancierung von ARKit für die Programmierung von markerbasierter AR im Rahmen von Unity3D ein (STACK OVERFLOW 2017).

2 Problemstellung

Es soll eine Mixed-Reality Applikation entwickelt werden, welche auf zwei Arten mit der realen Welt interagieren kann. Zunächst soll diese Interaktion mit einem 2D-Stadtplan realisiert werden. In der zweiten Phase soll untersucht werden ob es möglich ist auch eine Interaktion mit einem physikalischen Stadtmodell zu realisieren.

Um der immer weiter etablierten Nutzung von Geodaten auf mobilen Endgeräten Rechnung zu tragen, hat sich das Grundbuch- und Vermessungsamt Basel-Stadt für die Entwicklung einer Augmented-Reality-App als Erweiterung ihres Geoviewers entschieden und das Institut Geomatik IGEO der Fachhochschule Nordwestschweiz FHNW mit der Realisierung einer Augmentierung des offiziellen 2D-Stadtplans Basels beauftragt. Mittels Smartphones und Tablets soll der um 3D-Daten (digitales 3D-Stadtmodell, Terrain), Vektordaten (u.a. ÖV-Linien, Abfuhrzonen, Zonenplan), Punktdaten (u.a. Points of Interest POI) und anderen Informationen, wie Wetter- und Verkehrsdaten erweiterte Stadtplan intuitiv erfahrbar sein und somit die frei verfügbaren Geodaten des Kantons in einem modernen Angebot der Öffentlichkeit präsentiert werden. Die Entwicklung der App ist in Arbeit und wird mit Unity3D unter Verwendung des Vuforia SDK realisiert. Gleichzeitig soll auch untersucht werden, inwiefern sich physikalische Stadtmodelle augmentieren lassen können. Das physikalische Stadtmodell von Basel-Stadt ist ein im Maßstab 1:1000 gefertigtes, realitätsnahes Architekturmodell (LOD2) der gesamten Basler Innenstadt, sowie einigen Vorstädten des umgebenden trinationalen Ballungsraumes. Der öffentlich zugängliche Ausstellungsraum verfügt über ein Beleuchtungssystem, welches einfache Schattenwurfsimulationen ermöglicht. Das dazugehörige Modellbauatelier verfolgt das Ziel, städtisch-raumplanerische Zusammenhänge im Rahmen von Architekturwettbewerben und Testplanungen in diesem musealen Kontext erleb- und erfahrbar zu machen (PLANUNGSAMT DES KANTONS BASEL-STADT 2017).

Es existieren derzeit verschiedene mögliche *Mixed/Augmented Reality* (AR)-Technologien, welche die Realisierung einer Mobile App zur Überlagerung des physikalischen Basler Stadtmodells mit zusätzlicher Information ermöglichen. Ziel er Vorstudie war die Identifikation einer technologischen Lösung, welche die dreidimensionale Lage- und Raumbeziehung von Modell und Betrachter abschätzen kann. Außerdem soll zusätzliche 3D-(Geo-)Informationen durch Bildüberla-

gerung auf handelsüblichen mobilen Endgeräten angezeigt werden können. Über die technische Machbarkeit und die Abwägung der verschiedenen Technologien hinaus, wurde eine Analyse vorgenommen, welche Daten sich zur Überlagerung mit dem Modell eignen, wie diese visualisiert werden und woher sie bezogen werden könnten.
Zusätzlich zur Vorstudie umfasst das Projekt einen Experimentalteil, der die direkte Anwendbarkeit der verschiedenen Technologien untersucht. Hierbei stand neben der Bewertung der verschiedenen Technologieträger vor allem die Entwicklung eines lauffähigen Prototyps im Vordergrund. Dieser stellt die technische Grundlage für die zukünftige Entwicklung (ab Mitte 2018) einer vollwertigen, downloadbaren, öffentlichen Mobile App dar. Der Fokus bei der Appentwicklung wurde in erster Linie auf die sichere Erkennung der Position von Modell und Betrachter, die Stabilität des Trackings und eine performante 3D-Laufzeitumgebung gelegt.
Als AR-Marker - im Sinne eines durch Computer Vision und Bildanalyse erkennbaren Objekts in der realen Welt - stellt das von Hand und aus Holz gefertigte Architekturmodell eine besondere Herausforderung dar: Die Größe des Modells von ca. 40 m^2 geht weit über die normalerweise zur Anwendung kommenden Ausmaße typischer AR-3D-Marker (z.B. Industrie- und Verbraucherprodukte wie Maschinenteile, Spielzeug, Fahrzeuge oder Möbel) hinaus. Auch die Lichtverhältnisse im musealen Kontext des Modells, sowie texturarme Oberflächen und repetitive Muster stellen teilweise hohe Anforderungen an Hard- und Software dar. Durch die (in den letzten Jahren massiv von verschiedensten Marktteilnehmern vorangetriebene) Technologieentwicklung ist die Realisierung von Zukunftsvisionen - wie der hier vorgestellten Echtzeit-Überlagerung von 3D-Information an einem Architekturmodell - nun möglich geworden.

3 Realisierung der Apps

3.1 Mixed Reality mit dem 2D-Stadtplan Basel-Stadt

Abb. 1: Augmented-Reality-App für den Stadtplan von Basel

3.1.1 Datenhandling

Der Umgang mit Geodaten in Unity3D ist mit großem Aufwand verbunden. Als Grundlagendaten liegen sämtliche verfügbaren Geodaten des Geoportals des Kantons Basel Stadt (http://www.geo.bs.ch) im Bezugssystem CH1903+ (LV95, EPSG:2056) vor. Unity verwendet ein lokales Koordinatensystem, wobei eine Unity-Einheit einem Meter entspricht. Sämtliche Geodaten wurden unter Beibehaltung der Skalierung zum Nullpunkt verschoben.
Unity kann mit Vektordaten nicht umgehen und erlaubt nur den Import von Solid Models in Formaten wie Wavefront OBJ oder Autodesk FBX. Volumenmodelle wie das 3D-Stadtmodell können verhältnismässig einfach konvertiert, importiert und in der AR-Szene platziert werden. Geländemodelle können mit Photoshop in .raw-Files gespeichert, in Unity als Heightmaps importiert und als Terrain-GameObjects mit Textur dargestellt werden.
Linien-, Flächen- und Punktdaten zu importieren erweist sich als aufwändiger. Für Shapefiles und andere Vektordatenformate bietet Unity keine direkte Möglichkeit. Aus den Liniengeometrien von Shapefiles wurden deshalb in FME Workbench mittels Buffering, Extruding und Surface Draping dreidimensionale Meshes generiert, ähnlich konnten die Modelle der Flächendaten erstellt werden. Damit allfällige Attribute nicht verloren gehen, können die Shapefiles zusätzlich in JSON konvertiert und die Attribute der Daten in C#-Klassen gespeichert werden. Die so erhaltenen Meshes der Linien- und Flächendaten werden direkt in Unity platziert. Auch aus den Shapefiles der Punktdaten werden C#-Klassen erstellt. Alle Attribute inklusive der Koordinaten werden gespeichert, was eine Platzierung aller POIs in der Szene mittels Funktionen aus C#-Scripts erlaubt.

3.1.2 Image Targets

Vuforia stellt ein Marker-basiertes Augmented Reality-Framework dar. Um den räumlichen Bezug zwischen Kamera und Projektionsfläche (in diesem Sinne der Stadtplan) herzustellen, werden Marker, sogenannte ImageTargets, erstellt.
2D Image Targets sind Bilder, welche durch die Vuforia SDK erkannt und getrackt werden können. Herkömmliche Referenzmarken wie QR Codes oder Data Matrix Codes benötigen spezielle Schwarz-Weiss-Muster oder Strich- und Zahlencodes für eine Detektion. Vuforia Target Manager erlaubt dagegen die Generierung einer Feature-Repräsentation beliebiger Bilder. Die generierten Feature-Repräsentationen der gewünschten Marker werden als Unity Package gespeichert, in Unity importiert und für das Tracking verwendet, indem im Kamerabild erkannte Features mit denjenigen der Image Targets verglichen werden. Marker können im PNG- und JPEG-Format hochgeladen werden, sollten mindestens 320px Seitenlänge aufweisen und dürfen nicht größer sein als 2MB (VUFORIA 2018).

Abb. 2: Stadtplan-Kachel als Image Target

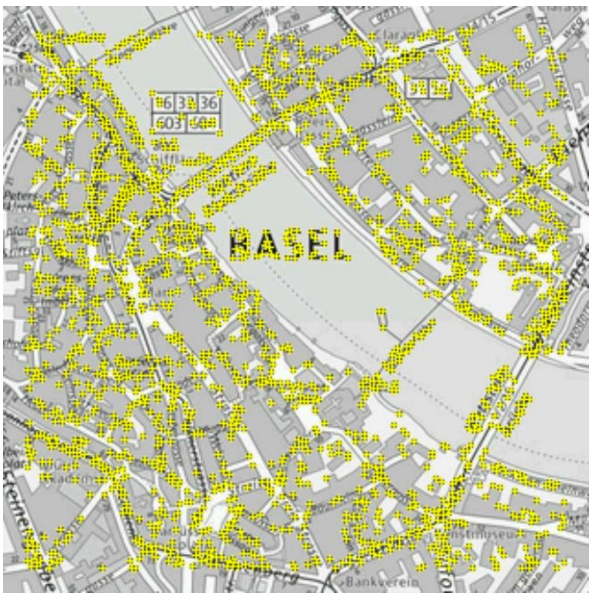


Abb. 3: von Vuforia SDK erkannte Features

2D Image Targets eignen sich speziell für die Augmentierung für Printmedien, weshalb der Stadtplan gekachelt und als Marker verwendet wird. Es werden drei Zoomstufen berücksichtigt, damit ein stabiles Tracking in allen realistischen Distanzen zwischen Kamera und Stadtplan gewährleistet ist. Neben der Eignung der Image Targets hängt ein erfolgreiches Tracking zusätzlich von einer ausreichenden und gleichmäßigen Belichtung des Markers ab.

3.1.3 User Interface UI

Abb. 4: User-Interface und Anzeige von Vektordaten in der 3D-Ansicht

Das User Interface der App ist noch in Entwicklung und erlaubt derzeit das Einstellen verschiedener Fokusmodi der Kamera, das Aufnehmen eines Screenshots, das Einblenden einer Info-Seite und natürlich die Augmentierung des Stadtplans mit Geodaten. Dafür steht zum Einen

eine Schaltfläche für das Umschalten zwischen 2D- und 3D-Modus und zum Anderen ein Button zum Aufrufen aller Geodaten-Inhalte zur Verfügung. So können bisher das 3D-Stadtmodell, das Verkehrsnetz und die Haltestellen des öffentlichen Verkehrs, der Zonenplan, die verschiedenen Abfuhrzonen, Velorouten und Point of Interests eingeblendet werden. Im 2D Modus wird alles flach auf den Stadtplan projiziert, der 3D-Modus liefert alle Geodaten unter Berücksichtigung der Höhenwerte. Mittels Double-Tap auf das Display können alle UI-Elemente für eine uneingeschränkte Betrachtung ausgeblendet werden.

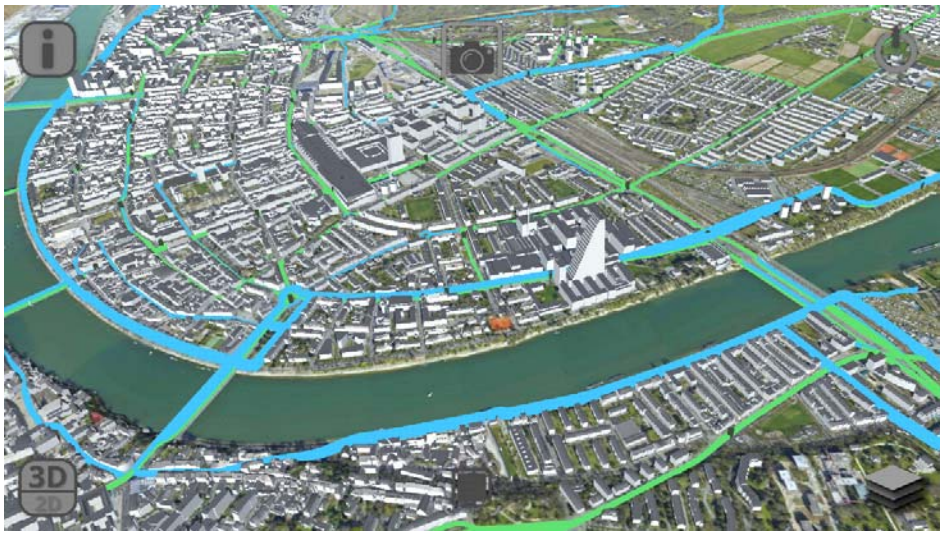


Abb. 5: Vektordaten (ÖV-Linien und Velorouten) als Meshes in der 3D-Ansicht

3.2 Mixed Reality am physikalischen Stadtmodell

3.2.1 Technologie

Nach einigen Vortests (Programmierung und Deployment von AR-Testapps im Rahmen der offiziellen Entwicklertutorials der jeweiligen Plattform) fiel die Wahl zunächst auf entweder ARKit oder Vuforia als aussichtsreichste Kandidaten für die Entwicklung eines Prototyps. Die Entscheidung fiel letztendlich gegen ARKit, weil kein Demonstrator und kein dem Rahmen dieses Projekts angemessener Programmierweg für die Integration eines 2D- oder 3D-Markers in die AR-Erfahrung gefunden werden konnte. Über den „Umweg“, den Benutzer das Modell zunächst umrunden zu lassen, und dabei bestimmte Keypoints manuell zu markieren, wäre eine Referenzierung des globalen Bezugsrahmens trotzdem möglich gewesen. Dies hätte jedoch einen zusätzlichen Aufwand für Endnutzer und Entwickler gleichermaßen bedeutet.

Abb. 6: Vuforia Object Scanner beim Digitalisierungstest an einem kontrastarmen Architekturmodell

Mit Vuforia hingegen konnten alle Vortests mit positivem Ergebnis durchlaufen werden. Auch auf die Gefahr hin, dass die Stabilität des Trackings mit der (nicht offiziell vorgesehenen) Dimensionierung des Stadtmodells gegenüber kleineren Objekten abnehmen könnte, lagen die Voraussetzungen für die Entwicklung eines Prototyps in dieser Umgebung gegenüber anderen Konkurrenten denkbar besser. Als Umsetzungsplattform wurde Microsoft Hololens verworfen, weil großflächige Marker noch sehr schwer handzuhaben waren und die Anschaffung mehrerer Geräte aufgrund des noch sehr hohen Preises für den Ausstellungsraum derzeit wohl unrealistisch ausfällt. Die Wahl fiel schließlich auf ein modernes Android-Smartphone statt auf ein iPad als Umsetzungsplattform, weil eine für dunkle Umgebungen besser geeignetere Kamera verbaut wurde und online auf eine aktivere Entwicklercommunity zurückgegriffen werden konnte.

3.2.2 Prototyp

Tracking und Orientierung von Modell und Betrachter funktionierten dabei grundsätzlich mittels Unity3D und Vuforia auf einem handelsüblichen Android-Smartphone. Die Stabilität des Trackings ist an etwa 80% des Modellumfangs gewährleistet, solange ein Abstand von mindestens 1m zum Modell eingehalten wird. Die Stabilität des Trackings erhöht sich mit zusätzlichem Abstand, sowie mit einer steileren Kameraperspektive auf das Modell.

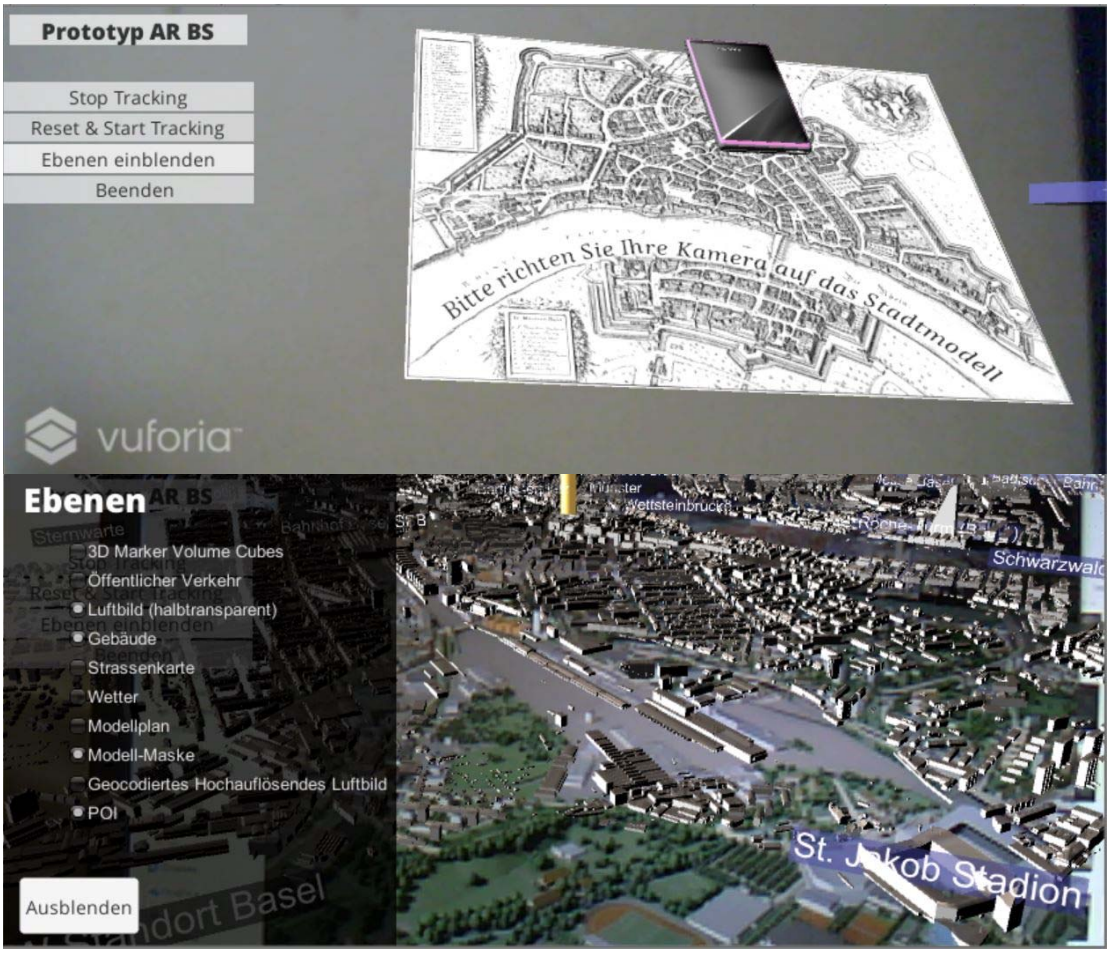


Abb. 7: Screenshots aus der Prototyp-Applikation

Durch das Scannen weiterer Target Regions und eine präzisere Feinjustierung der räumlichen Positionierung dieser Target Regions zueinander sollte eine Steigerung des trackbaren Modellumfangs auf >95%, sowie eine Unterstützung flacherer Kamerawinkel und niedrigerer Abstände möglich sein. Entscheidend war, herauszufinden, welche Faktoren das Vuforia-Framework bei der Detektion eines derart großen 3D-Markers beeinflussen. Als vorteilhafte Faktoren in der Skalierung stellten sich die Ortsstabilität und die kontrollierbaren Lichtverhältnisse im musealen Kontext dar. Nachteilig wirkten sich die enorme Größe und die schwierige, erhöhte Position des Modells, sowie die dunklen Lichtverhältnisse im Bereich der neben dem Modell anzusetzenden Object Scanning Targets (Abb. 8) aus. Das grundsätzliche Funktionieren des Prototyps beweist jedoch, dass prinzipiell die geforderte Funktionalität der AR-App durch Vuforia gewährleistet werden kann. Hinsichtlich der Stabilität des Trackings und der APK-Dateigröße besteht jedoch noch Optimierungsbedarf.

Punkt-, Linien-, Flächen- und 3D-Daten konnten erfolgreich an der richtigen Position im Modell augmentiert werden. Teilweise muss die Erkennung des Modells im Nahbereich noch verbessert werden, um eine hinreichende Überlagerungstreue herzustellen. Die 3D-Laufzeitumgebung von Unity bietet eine sehr performante und interaktive Darstellungsqualität. Selbst bei sehr umfang-

reichen Datensätzen wie dem digitalen 3D-Stadtmodell kam es zu keinem unangenehmen Ruckeln oder Abstürzen. Das schwächste getestete Gerät konnte unter Android 7.1.1 mit 3Gb RAM und Octacore-CPU eine flüssige Darstellung gewährleisten. Das Verhalten der 3D-Umgebung auf hardware-schwächeren Geräten muss noch getestet werden.

Abb. 8: Abfallende Lichtverhältnisse auf dem Object Scanning Targets erschweren die Digitalisierung einer spezifischen Region am physikalischen Stadtmodell

Die dargestellten Geodaten sind gut erkennbar, Beschriftungen wurden mittels C#-Script korrekt orientiert und das Menü gestattet das Ein- und Ausblenden von Informationslayern. Trotz der Integration nur relativ weniger, beispielhafter Geodaten ist die APK-Datei mit ca. 370 Mbyte derzeit noch sehr groß, was ein Herunterladen über eine datenlimitierte 4G-Verbindung unrealistisch macht. In Kombination mit dem kostenlos verfügbaren Gäste-WLAN des Ausstellungsraums sollte dies jedoch nur bedingt eine Hürde darstellen. Vielmehr ist der interne Speicherplatz vieler (günstiger) Android-Geräte jedoch häufig so knapp bemessen, dass einige Besucher möglicherweise gar nicht in der Lage wären, die App zu installieren. An der Größe der APK-Datei muss daher noch gearbeitet werden. Ein möglicher Ausweg wäre das Auslagern der Target-Erkennungsbibliothek und der Geodaten in die Cloud, was jedoch mit zusätzlichen Kosten auf Betreiberseite verbunden ist. Auch bei der Start-, bzw. Ladezeit der App bis zur Initialisierung des Trackings zeigt sich noch Optimierungsbedarf. Zusätzlich weist die Notwendigkeit, das Tracking derzeit noch manuell über einen Resetbutton steuern zu müssen, noch auf offenes Automatisierungspotential hin.

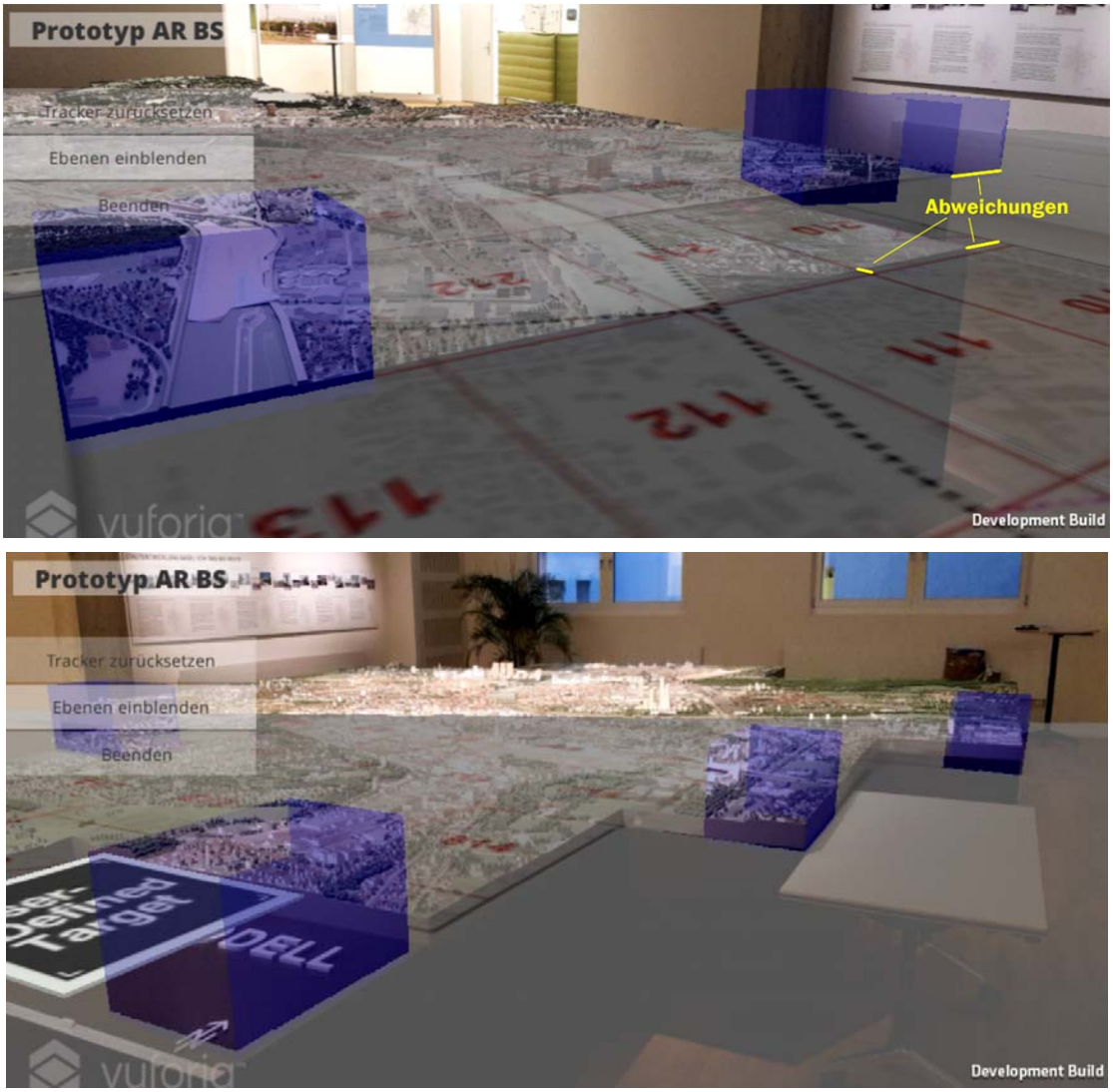


Abb. 9: Oben: Die virtuell skalierten Positionen zeigten zunächst Abweichungen (gelb). Unten: Erfolgreiche Reduktion der Abweichungen durch manuelle Referenzierung (dreidimensionales Einpassen der Object Target Regions, blaue Kuben).

4 Fazit & Ausblick

Es wurden einige Anwendungen im Bereich Mixed Reality mit 3D-Stadtmodellen präsentiert. Zukünftig können mehr Daten überlagert werden. Die Umsetzung des ersten Moduls ist weitgehend abgeschlossen. Erste Tests mit Untergrunddaten sowie Animation von Verkehrsmitteln sind erfolgt. Alle Module werden weiterentwickelt und implementiert bevor die App Ende März im Google Play Store und Apple App Store veröffentlicht wird.
Weiter wurde gezeigt, dass auch physische Stadtmodelle mit akzeptabler Genauigkeit überlagert werden können. Der vorliegende experimentale Prototyp der App demonstriert technisch eindrucksvoll, dass es heute möglich ist, unter überschaubarem Ressourcenaufwand innerhalb weniger Monate ein funktionierendes AR-Erlebnis zu entwickeln. In dieser App steckt noch viel

Potential - sowohl in verbesserten Methoden für das Tracking wie auch in der Anreicherung von Daten. In späteren Versionen der App wird es durchaus möglich sein auch das physische Stadtmodell mit geeigneten Geodaten zu überlagern. Zukünftig soll auch darauf geachtet werden, dass Geodaten effizienter komprimiert werden, um die App-Größe und auch das Download-Volumen so gering wie möglich zu halten.

5 Literaturverzeichnis

APPLE INC., 2017: ARKit - Apple Developer. https://developer.apple.com/arkit/, letzter Zugriff 2.10.17.

AZUMA, R.T., 1997: A Survey of augmented reality. In: Presence, **6**(4), 355-385.

BEHZADAN, A.H., DONG, S. & KAMAT, V.R., 2015: Augmented reality visualization: A review of civil infrastructure system applications. In: Advanced Engineering Informatics **29**/2 (April), 252-267. doi:10.1016/j.aei.2015.03.005.

BILLINGHURST, M., CLARK, A. & LEE, G., 2015: A Survey of Augmented Reality. (= Foundations and Trends in Human-Computer Interaction 8:2-3). https://www.orellfuessli.ch/shop/home/artikeldetails/a_survey_of_augmented_reality/mark_billinghurst/ISBN1-60198-920-2/ID42328590.html, letzter Zugriff 7.10.2017.

CHRISTEN, M., BLASER, S., LOESCH, B. & NEBIKER, S., 2016: Prozedurale Modellierung von historischen 3D-Stadtmodellen und deren Aufbereitung für 3D-Visualisierungen in einer Smartphone-Applikation am Beispiel der römischen Stadt Augusta Raurica. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 25, T. Kersten (Hrsg.), 217-229.

COLTEKIN, A., 2016: VISDOM: Vision to Visualization - Improving computational and human performance with highly realistic three-dimensional geographic visualizations by means of biomimicry. In: Forschungsdatenbank Universität Zürich. http://www.research-projects.uzh.ch/p20215.htm.

COSTELLO, S., 2016: The 6 Sensors in an iPhone. In: Lifewire, https://www.lifewire.com/sensors-that-make-iphone-so-cool-2000370, letzter Zugriff 22.10.2017.

DILGER, D.E., 2017: How Apple's iPhone X TrueDepth AR waltzed ahead of Google's Tango [WWW Document]. AppleInsider, appleinsider.com/articles/17/10/13/how-apples-iphone-x-truedepth-ar-waltzed-ahead-of-googles-tango, letzter Zugriff 22.11.17.

KLING, B., 2017: Augmented Reality: Samsung und Google schmieden Partnerschaft, http://www.zdnet.de/88316121/augmented-reality-samsung-und-google-schmieden-partnerschaft/

LEE, D., 2017: 9 cool AR apps you should download to try out iOS 11's ARKit. In: The Verge. https://www.theverge.com/2017/9/20/16329366/ios-11-apple-arkit-best-new-ar-apps, letzter Zugriff 22.10.2017.

LING, H., 2017: Augmented Reality in Reality. IEEE Multimedia 24, 10–15. DOI: https://doi.org/10.1109/MMUL.2017.3051517

LOESCH, B., CHRISTEN, M., WÜEST, R. & NEBIKER, S., 2015: Geospatial Augmented Reality– Lösungsansätze mit natürlichen Markern für die Kartographie und die Geoinformationsvi-

sualisierung im Außenraum. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 24, T. Kersten (Hrsg.), 89-97.

MULA, J., 2015: Augmented Reality: What to choose? In: Unt3Dmag, http://unt3dmag.com/augmented-reality-what-to-choose/; letzter Zugriff 22.11.17.

PENG, F. & ZHAI J., 2017: A mobile augmented reality system for exhibition hall based on Vuforia. 2nd International Conference on Image, Vision and Computing (ICIVC), 1049-1052. DOI: https://doi.org/10.1109/ICIVC.2017.7984714

PIUMSOMBOON, T., DAY, A., ENS, B., LEE, Y., LEE, G. & BILLINGHURST, M., 2017: Exploring enhancements for remote mixed reality collaboration. In SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications (SA '17). ACM, New York, NY, USA, Article 16, 5 pages. DOI: https://doi.org/10.1145/3132787.3139200

PLANUNGSAMT DES KANTONS BASEL-STADT, 2017: Das Modellbauatelier. http://www.planungsamt.bs.ch/ueber-uns/modellbauatelier.html; letzter Zugriff 26.10.17.

RODRIGUES, D.G., JAIN, A., RICK, S.R., SHANGLEY, L., SURESH, P. & WEIBEL, N., 2017: Exploring Mixed Reality in Specialized Surgical Environments. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17). ACM, New York, NY, USA, 2591-2598. DOI: https://doi.org/10.1145/3027063.3053273

SCHMID, J., 2017: 8 Best augmented reality SDK for AR development for iOS and Android in 2017. In: LinkedIn Pulse, https://www.linkedin.com/pulse/8-best-augmented-reality-sdk-ar-development-ios-android-schmidt, letzter Zugriff 22.11.17.

SUTHERLAND, I., 1965: The ultimate display. In: Multimedia: From Wagner to virtual reality.

STACK OVERFLOW, 2017: Is it possible to track objects with ARKIT like with Vuforia? https://stackoverflow.com/questions/44817416/arkit-is-it-possible-to-track-objects-like-with-vuforia, letzter Zugriff 22.11.17.

UNITY TECHNOLOGIES, 2017: Unity ARKit Plugin - Asset Store. https://www.assetstore.unity3d.com/en/#!/content/92515, letzter Zugriff 22.11.17.

VUFORIA, 2018: Vuforia Augmented Reality SDK, https://www.vuforia.com/

WIESEND S., 2015: Augmented Reality: Was Apple mit Metaio vorhat. In: Macwelt. https://www.macwelt.de/news/Augmented-Reality-Apple-Metaio-Faceshift-9894664.html, letzter Zugriff 22.11.17.

WIKITUDE, 2017: Store licenses for augmented reality products by Wikitude. https://www.wikitude.com/store/, letzter Zugriff 22.11.17.

WÜEST, R., ZWICK, M. & NEBIKER S., 2016: Geospatial Augmented Reality als interaktives Erlebnis im Museum am Beispiel der Swissarena App. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 25, T. Kersten (Hrsg.), 197-207.

Eine voxelbasierte Metadatenstruktur zur Änderungsdetektion in Punktwolken großflächiger urbaner Gebiete

JOACHIM GEHRUNG[1,2], MARCUS HEBEL[1], MICHAEL ARENS[1] & UWE STILLA[2]

Zusammenfassung: Im Rahmen dieser Arbeit wird eine Methode vorgestellt, mit welcher sich eine volumetrische Approximation einer Punktwolke erstellen lässt. Weiterhin wird ein Verfahren zur Strahlverfolgung vorgestellt, welches einen auf Ebenen basierenden Filter einsetzt, um auf Grund der Diskretisierung des Raums entstandenen Artefakten an planaren Oberflächen entgegenzuwirken. Mittels beider Verfahren wird eine Metadatenstruktur aufgebaut, bei welcher es sich um eine effiziente Repräsentation der Umgebung handelt, in der Zusatzinformationen mit Ortsbezug abgelegt sind. Eine Evaluation beider Verfahren mit einem öffentlich verfügbaren Datensatz hat ergeben, dass sich aus geometrischer Sicht komplexe Strukturen abbilden lassen. Die Metadatenstruktur ist frei von besagten Artefakten. Die Ergebnisse der Untersuchung deuten darauf hin, dass eine gesonderte Behandlung von bewegten Objekten und Vegetation notwendig ist.

1 Einleitung

Die zuverlässige dreidimensionale Erfassung der Umgebung ist ein zunehmend wichtiger Bestandteil bei heutigen Mobile Mapping Systemen, beispielsweise im Hinblick auf Anwendungen wie die 3D-Rekonstruktion von Häuserfassaden für 3D-Stadtmodelle. Immer häufiger wird hierfür mobiles Laserscanning bzw. LiDAR eingesetzt (engl. *Light Detection and Ranging*). Im Gegensatz zu kamerabasierten photogrammetrischen Ansätzen und luftgebundenen 3D-Sensoren ist mobiles LiDAR dazu in der Lage, dichte und zugleich detailreiche Punktwolken urbaner Umgebungen zu erzeugen. Die hohe Qualität dieser Daten ist insbesondere für die automatische Änderungsdetektion attraktiv. Durch diese kann die Entwicklung einer urbanen Umgebung dokumentiert, automatisch ausgewertet und quantifiziert werden. Gerade im Hinblick auf statische Strukturen lassen sich nicht nur offensichtliche Änderungen wie abgerissene oder neu erbaute Häuser identifizieren, sondern auch eher subtile Details wie beschädigte oder fehlende Stadtmöbel.

1.1 Forschungsfragen

Änderungsdetektion in großräumigen urbanen Gebieten wirft im Wesentlichen zwei Forschungsfragen auf. Die erste betrifft die Aufbereitung der gesammelten Daten. Da der Fokus in dieser Arbeit auf baulichen und damit statischen Strukturen liegt, ist zu klären, wie sich bewegte Objekte aus den Punktwolken oder anderen Formen der Umgebungsrepräsentation entfernen lassen. Dies betrifft vor allem bewegte, aber auch temporär unbewegte Objekte. Beispiele für letzteres sind geparkte Fahrzeuge oder stillstehende Personen. Die Entfernung beider Objektklassen ist

[1] Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB, Gutleuthausstraße 1, D-76275 Ettlingen, E-Mail: [joachim.gehrung, marcus.hebel, michael.arens]@iosb.fraunhofer.de
[2] Technische Universität München, Photogrammetrie und Fernerkundung, Arcisstr. 21, D-80333 München, E-Mail: stilla@tum.de

notwendig, da diese im Fall einer Änderungsdetektion als Änderungen erkannt werden, aber erwartbare Änderungen und damit ungewollte Detektionen darstellen.
Die zweite Forschungsfrage beschäftigt sich mit der Skalierung der Änderungsdetektion auf großflächige Gebiete. Aus methodischer Sicht ist zu klären, inwiefern Algorithmen und Datenstrukturen auch für große Areale skalieren. Datenstrukturen sind so zu konzipieren, dass sie möglichst speichereffizient sind und wenn möglich sogar eine fixe Speicherobergrenze aufweisen. Aus algorithmischer Sicht sind Methoden so zu entwerfen, dass sie möglichst laufzeiteffizient sind. Eine wichtige Rolle spielt hierbei auch die Möglichkeit zur Parallelisierung der Verfahren.

1.2 Lösungsansatz

Bei der Umgebungsrepräsentation durch Voxel handelt es sich um eine Datenstruktur, welche sowohl der Änderungsdetektion an sich als auch deren Anwendung auf große urbane Gebieten zuträglich ist. Eine solche volumetrische Repräsentation skaliert gut für große Umgebungen, da sie auf Grund der Diskretisierung des Raums nicht nur ähnliche Messdaten zusammenfasst, sondern auch eine fixe Speicherobergrenze aufweist. Für die Änderungsdetektion haben Volumen den Vorteil, dass hier neben reinen Oberflächeninformationen auch freie und bisher unbeobachtete Raumbereiche repräsentiert werden können. Diese Informationen spielen eine wichtige Rolle wenn es darum geht zu ermitteln, ob tatsächlich eine Änderung stattgefunden hat oder ob die potentielle Änderung daher rührt, dass der betroffene Raumbereich zum ersten Mal mit den Sensoren erfasst wurde. Im ersten Fall findet ein Zustandswechsel der betroffenen Voxel von frei nach belegt bzw. umgekehrt statt; im zweiten Fall erfolgt der Wechsel von ungesehen in einen beliebigen anderen Zustand. Für die volumetrische Repräsentation wird auf Methoden und Datenstrukturen aus unseren früheren Arbeiten zurückgegriffen, in denen eine ähnliche Datenstruktur bereits eingesetzt wurde (GEHRUNG et al. 2016).

1.3 Probleme volumenbasierter Repräsentationen

Der Nachteil eines auf Volumen basierten Ansatzes ist die Anfälligkeit für Artefakte, die durch die Diskretisierung des Raums bedingt sind (vergleiche Abb. 1). Durch diese kann es dazu kommen, dass Raumbereiche unterschiedlicher Natur wie z.B. Freiraum und belegter Raum im selben Voxel zusammengefasst werden. Ein Artefakt entsteht also, wenn ein Voxel von Sichtstrahlen durchlaufen wird, obwohl es auf Grund von anderen Observationen als belegt anzusehen ist. Hierdurch kommt es zu einem Aufweichen des Voxelzustands, da diesem durch die widersprüchlichen Beobachtungen nicht mehr eindeutig ein Label wie *frei* oder *belegt* zugeordnet werden kann. Diese sogenannten Traversionsartefakte treten an Grenzen zwischen Objekten und Freiraum auf und wurden in einer unserer vorangegangenen Arbeiten bereits thematisiert (GEHRUNG et al. 2017). Es ist zwischen zwei Arten dieser Artefakte zu unterscheiden. Die erste Kategorie tritt vereinzelt an Objektgrenzen auf. Eine stärker ausgeprägte Variante findet man vor allem an planaren Oberflächen. In flachem Winkel auftreffende Sichtstrahlen traversieren hier meist mehrere Voxel der Oberfläche, so dass es zu einer großflächigen Zustandsverfälschung kommt (siehe Abb. 1 (B)).

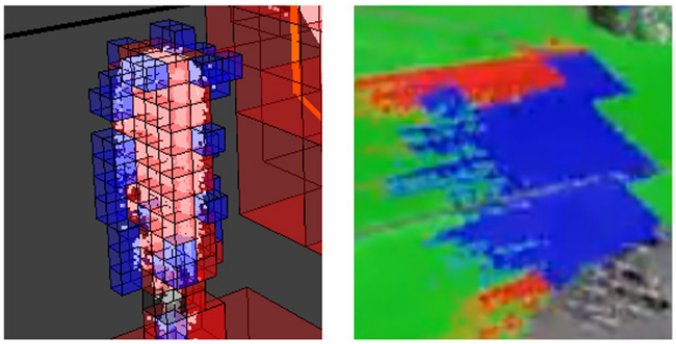


Abb. 1: Durch Diskretisierung bedingte Artefakte (A) an Objektgrenzen und (B) entlang planarer Flächen.

1.4 Beitrag dieser Arbeit

Diese Arbeit stellt einen weiteren Schritt in Richtung einer volumetrischen Repräsentation dar, die für die Änderungsdetektion in großen urbanen Gebieten geeignet ist. Es wird ein methodischer Ansatz vorgestellt, welcher eine volumetrische Approximation einer Punktwolke bzw. einer Menge an Sichtstrahlen ermöglicht, wobei sich die Auflösung der Volumina dynamisch an die lokale Geometrie anpasst. Weiterhin wird ein für diese Methode eingesetzter Algorithmus zur Strahlverfolgung (engl.: Raycasting) vorgestellt, welcher einen auf Ebenen basierten Filterschritt nutzt, der das Auftreten der oben beschriebenen Artefakte an planaren Oberflächen verhindert.

2 Stand der Forschung

Änderungsdetektion ist ein eigenständiger Forschungszweig, eine klare Abgrenzung zu anderen Bereichen ist jedoch nicht immer möglich. In eine ähnliche Richtung, jedoch mit einem weitaus kürzeren betrachteten Zeitfenster, geht beispielsweise *Detection and Tracking of Moving Objects (DATMO)*. Das Erkennen bewegter Objekte spielt vor allem im Forschungsumfeld von Robotik, autonomem Fahren und SLAM (*Simultaneous Localization and Mapping*) eine große Rolle, wenn auch mit unterschiedlichem Fokus. So entfernen z.B. LITOMISKY & BHANU (2013) mit einem clusterbasierten Ansatz bewegte Objekte aus Punktwolken, um damit die Performanz eines SLAM-Verfahrens zu verbessern. Viele Verfahren, die sich im weitesten Sinne mit Änderungsdetektion befassen, sind im Bereich des 2D-Maschinensehens angesiedelt. Da der Fokus dieser Arbeit jedoch auf der Verarbeitung von Punktwolken liegt, wird hier nur auf Verfahren mit eben diesem Schwerpunkt eingegangen.

2.1 Änderungsdetektion in Punktwolken

Auf gemessenen 3D-Punktwolken basierende Ansätze haben gegenüber der Verwendung von Bilddaten den Vorteil, dass sie schnell und effizient sind, da keine geometrische 3D-Rekonstruktion stattfinden muss und keine Verfälschung durch unterschiedliche Blickrichtungen eintritt. GIRARDEAU-MONTAUT et al. (2005) etwa organisieren Punktwolken räumlich in einem

Octree und stellen darauf aufbauend einige einfache Vergleichsverfahren für Teilpunktwolken vor. ZEIBAK & FILIN (2008) haben einen Ansatz zum Vergleich von Tiefenbildern vorgestellt, welche mittels eines stationären terrestrischen Laserscanners erzeugt wurden. Auf Grund der Einfachheit des Verfahrens ist nur der Vergleich von Aufnahmen möglich, die vom selben Standpunkt aus aufgenommen wurden. Beide Ansätze berücksichtigen jedoch keine Informationen über freien oder ungesehenen Raum.

2.2 Änderungsdetektion basierend auf Sichtstrahlen

Eine alternative Interpretation von Punktwolken sind Sichtstrahlen. Hierbei wird die Sensorposition als Ausgangspunkt des Strahls angenommen, der Messpunkt als Endpunkt. Durch die Verfolgung des Strahls lassen sich Informationen über freie und ungesehene Raumbereiche ableiten. UNDERWOOD et al. (2013) nutzen Strahlverfolgung in sphärischen Koordinaten, um einzelne Sensoraufnahmen im Sinne der Änderungsdetektion miteinander zu vergleichen. HEBEL et al. (2013) haben einen ebenfalls auf Sichtstrahlen basierten Ansatz zur Änderungsdetektion in LiDAR-Messungen fliegender Sensorträger vorgestellt. Einzelne Sichtstrahlen werden in einer voxelbasierten Index-Datenstruktur organisiert und mittels eines Evidenzmodells auf Basis der Dempster-Shafer Theorie miteinander verglichen. XIAO et al. (2015) nutzen einen ähnlichen Ansatz für das Auffinden von Änderungen in mobilen LiDAR-Daten. Ein Vorteil dieser Kategorie von Ansätzen ist es, dass die ursprünglichen Messinformationen nicht durch eine Diskretisierung reduziert werden. Dies hat jedoch zugleich den Nachteil, dass für eine Auswertung alle oder zumindest ein Teil der Sichtstrahlen herangezogen werden müssen. Für Änderungsdetektion in mobilen Laserdaten, die Millionen von Einzelmessungen pro Sekunde umfassen, ist dies im Hinblick auf Speicherbedarf und Laufzeit ein wenig praktikables Vorgehen.

2.3 Änderungsdetektion basierend auf Segmenten oder Objekten

Anstatt eine Änderungsdetektion auf den Messdaten an sich durchzuführen, kann auch eine vorgeschobene Vorverarbeitung der Punktwolken durchgeführt werden, beispielsweise durch Segmentierung oder Klassifikation. SCHACHTSCHNEIDER et al. (2017) nutzen eine Segmentierung basierend auf Region-Growing, um eine Punktwolke in Cluster zu unterteilen. Deren zeitliches Verhalten wird mittels eines Belegungsgitters ermittelt und ausgewertet. AIJAZI et al. (2013) setzen ebenfalls Segmentierung ein, klassifizieren dann allerdings die Cluster in Instanzen bekannter Objektklassen. Einige davon sind temporärer, andere permanenter Natur. Eine Ähnlichkeitskarte basierend auf einem Belegungsgitter wird eingesetzt, um Messdaten mehrerer Epochen zu fusionieren und eine dreidimensionale urbane Karte abzuleiten bzw. Änderungsdetektion durchzuführen. Diese Kategorie von Ansätzen setzt eine zuverlässig funktionierende Segmentierung bzw. Klassifikation voraus. Gerade letzteres stellt eine Einschränkung auf bekannte Objektklassen dar und somit auch eine Zwangsspezialisierung auf das jeweilige Anwendungsgebiet.

2.4 Änderungsdetektion in Belegungsgittern

Belegungs- bzw. Evidenzgitter approximieren einen Raum, indem dieser durch ein zwei- oder dreidimensionales Gitter in einzelne Zellen unterteilt wird. In den einzelnen Gitterzellen sind Informationen mit einem räumlichen Bezug abgelegt. PAGAC et al. (1996) setzen ein 2D-Belegungsgitter basierend auf der Dempster-Shafer Theorie zur Repräsentation des Umfeldes

eines autonomen Fahrzeugs ein. WOLF & SUKHATME (2004) stellen einen ähnlichen Ansatz vor, allerdings werden hier zwei Belegungsgitter eingesetzt, um statische und dynamische Umgebungselemente zu repräsentieren. Andere Ansätze nutzen einen populären dreidimensionalen Belegungs-Octree, welches unter der Bezeichnung OctoMap bekannt ist (HORNUNG et al. 2013). AZIM & AYCARD (2012) nutzen OctoMap, um bewegte Objekte in Echtzeit zu identifizieren. Hierzu wird eine Voxelliste kontinuierlich aktualisiert, indem eine Konfliktsuche zwischen der Umgebungsrepräsentation und den aktuellen LiDAR-Messungen durchgeführt wird. Die Voxel werden geclustert und gefiltert, um Objektkandidaten zu generieren. Das theoretische Rahmenwerk von OctoMaps wurde auch bei unseren vorangegangenen Arbeiten eingesetzt. Es wurde gezeigt, dass das Framework zur Identifikation bewegter Objekte eingesetzt werden kann, es hierbei jedoch zu den oben beschriebenen Artefakten kommt (GEHRUNG et al. 2016; GEHRUNG et al. 2017).

3 Volumenbasierte Approximation von Punktwolken

In diesem Abschnitt wird der Beitrag dieser Arbeit vorgestellt. Zunächst wird kurz auf das theoretische Rahmenwerk eingegangen, welches als Basis für die Umgebungsrepräsentation eingesetzt wird. Danach wird eine Methode zur Strahlverfolgung vorgestellt, welche einen auf Ebenen basierten Filterschritt nutzt, um die oben beschriebenen Artefakte an planaren Oberflächen zu verhindern. Darauf aufbauend wird ein Verfahren vorgeschlagen, mit dem sich eine volumetrische Approximation einer gegebenen Punktwolke aufbauen lässt.

3.1 Volumetrische Repräsentation

Die als Basis dieser Arbeit verwendete Metadatenstruktur basiert auf der in unseren vorangegangenen Arbeiten eingesetzten volumetrische Repräsentation (GEHRUNG et al. 2016). Räumliche Informationen werden in einem globalen kartesischen Koordinatensystem wie beispielsweise ECEF abgelegt. Mit Hilfe eines Gitters wird der Raum in würfelförmige Zellen zerlegt. Eine Zelle wird als *Kachel* bezeichnet und enthält neben den zugehörigen Sichtstrahlen auch eine volumetrische Repräsentation des eingeschlossenen Raums. Hierbei handelt es sich um einen Octree mit variabler Tiefe, der die effiziente Ablage von Volumen unterschiedlicher Auflösung ermöglicht. Jede Octree-Zelle beinhaltet Statistiken wie die Anzahl und Reihenfolge der Sichtstrahlen, die ein Voxel durchlaufen bzw. darin enden. Ebenfalls pro Zelle enthalten ist eine Tabelle mit Wahrscheinlichkeiten, welche die Zugehörigkeit eines Voxels zu einer Klasse wie *frei*, *belegt* oder *Restklasse* beschreibt. Aus Gründen von Datenorganisation und Visualisierung wurde in der praktischen Umsetzung die Kantenlänge einer Kachel auf 32 Meter festgelegt.

3.2 Strahlverfolgung mit Ebenenfilter

Die Absicht des vorgestellten Algorithmus zur Strahlverfolgung ist es, Diskretisierungsartefakte entlang planarer Strukturen zu vermeiden. Diese werden wie oben beschrieben durch Messstrahlen verursacht, die im flachen Winkel zur Oberfläche auftreffen. Die zu Grunde liegende Idee ist, parallel zu jeder planaren Struktur eine virtuelle Ebene zu definieren, an welcher der Sichtstrahl abgeschnitten wird. Unter Abschneiden ist hier zu verstehen, dass nur der zu traversierende Teil des Strahls gekürzt wird, jedoch nicht der Endpunkt entfernt wird. Um sicherzustellen, dass kein

Voxel traversiert wird, welches zu der planaren Struktur gehört, ist die Distanz zwischen beiden Ebenen die Diagonale eines Voxels. Da eine Ebene über eine unendliche Ausdehnung verfügt, ist sie nicht geeignet, um eine räumlich begrenzte Struktur zu repräsentieren. Aus diesem Grund wird anstelle einer Ebene eine orientierte Bounding Box (OBB) eingesetzt, welche als lokale Annäherung der Ebene fungiert. Besagte orientierte Bounding Box wird aus der Punktwolke berechnet. Durch Expansion der Box um die Voxeldiagonale kann der Schnitt mit der o.g. virtuellen Ebene als Schnittberechnung zwischen Sichtstrahl und der orientierten Bounding Box formuliert werden.

3.2.1 Extraktion planarer Oberflächen

Die Extraktion der planaren Oberflächen aus der Punktwolke erfolgt bei unserem Ansatz durch einen auf Normalen basierten Region-Growing-Algorithmus, der von RABBANI et al. (2006) vorgestellt wurde. Hierbei werden benachbarte Punkte zu Clustern zusammengefasst, wenn sie sich in Bezug auf ein Ähnlichkeitsmaß gleichen. Bei diesem handelt es sich um den Winkel zwischen den Oberflächennormalen an den Positionen beider Punkte. Ausgangspunkte für den Algorithmus werden in Regionen mit minimaler Krümmung ausgewählt, da so die Anzahl möglicher Cluster reduziert wird. Da ebenfalls eine Vielzahl kleiner Oberflächen gefunden wird, wurde ein zusätzlicher Filterschritt angefügt. Hierzu wird eine Liste der Cluster erstellt, die nach der Clustergröße – also der Anzahl der Messpunkte im Cluster – sortiert wird. Das untere Viertel der Liste wird verworfen, nur die verbliebenen Cluster werden für das Ebenenfilter verwendet.

3.2.2 Berechnung der orientierten Bounding Boxen

Eine orientierte Bounding Box approximiert eine Punktwolke lokal besser als eine achsenparallele Bounding Box (AABB), ist jedoch aufwendiger zu berechnen. OBBs werden hier erzeugt, indem zunächst mit Hilfe der Hauptkomponentenanalyse das lokale Koordinatensystem der Punktwolke berechnet wird. Die Eigenvektoren sowie der Schwerpunkt des Clusters werden dazu benutzt, die Transformation in das Koordinatensystem zu berechnen und alle Punkte zu transformieren. Der letzte verbliebene Schritt ist nun die Bestimmung der achsenparallelen Bounding Box. Da die Punktwolke lokal nicht immer vollständig planar ist, wird die Achse der Box mit dem kleinsten Eigenwert hier (vorläufig) auf null gesetzt. Die achsenparallele Bounding Box wird zusammen mit der Transformation gespeichert, um später den Schnitttest durchzuführen.

3.2.3 Strahlverfolgung mit Ebenenfilter

Die Transformation einer Punktwolke in die in dieser Arbeit verwendete volumetrische Repräsentation erfolgt durch Strahlverfolgung. Hierzu wird ein Sichtstrahl von der Sensorposition bis hin zum gemessenen Oberflächenpunkt traversiert. Der durchlaufene Bereich wird als Freiraum vermerkt, der Endpunkt des Strahls als belegter Raum (oder ebenfalls als Freiraum, wenn der Strahl an der Kachelgrenze abgeschnitten wurde). Die entsprechenden Voxel werden aktualisiert, indem für jeden der beiden Fälle ein Zähler inkrementiert wird. Die Strahlverfolgung mit Ebenenfilter ist ein zweistufiger Prozess:

1. Kürzen des Sichtstrahls durch Schnitt mit allen Bounding Boxen.
2. Traversion des Freiraums und behandeln der Endpunkte.

Um nun die Traversion von Voxeln zu verhindern, die Teil von planaren Oberflächen sind, muss jeder Sichtstrahl zunächst mit allen Bounding Boxen geschnitten werden. Dies setzt voraus, dass der Anfangs- und Endpunkt des Strahls in das Koordinatensystem der orientierten Bounding Box transformiert werden. Hier findet dann ein Schnitttest mit einer achsenparallelen Bounding Box statt, beispielsweise der von WILLIAMS et al. (2003) vorgestellte. Bevor der Schnitttest durchgeführt wird, wird zunächst die (zuvor noch flache) Box um die Diagonale eines Voxels erweitert. Dies kann nicht bei der Extraktion der Box geschehen, da die jeweilige Voxelgröße und damit die benötigte Erweiterung der Bounding Box erst zum Zeitpunkt des Schnitttests bekannt sind. Liegt ein Schnitt vor, so liefert das Verfahren einen Offset t entlang des Strahls, welcher den Schnitt mit der Box angibt. Dieser wird dazu verwendet, den neuen Endpunkt entlang des untransformierten Strahls zu bestimmen.

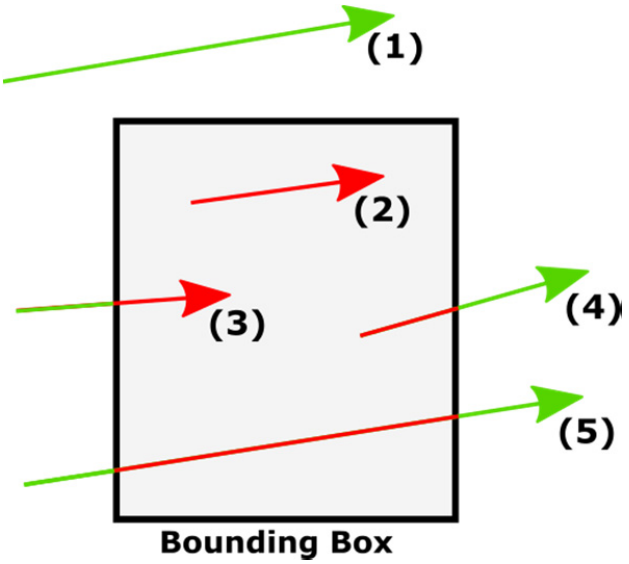


Abb. 2: Sonderfälle beim Schnitt zwischen Strahl und Box (1) Kein Schnitt. (2) Strahl komplett innerhalb. (3) Strahl beginnt außerhalb, endet innerhalb. (4) Strahl beginnt innerhalb, endet außerhalb. (5) Strahl startet und endet außerhalb, traversiert die Box.

Beim Schnitt zwischen Strahl und Bounding Box kann es zu mehreren Sonderfällen kommen (vergleiche Abb. 2). Die folgende Aufzählung erläutert die einzelnen Fälle und gibt an, wie mit diesen umzugehen ist.

1. Es liegt kein Schnitt zwischen Sichtstrahl und Box vor. Leerraum wird traversiert, der Endpunkt wird eingefügt.
2. Der Sichtstrahl liegt komplett innerhalb der Box. Keine Traversion des Freiraums, der Endpunkt wird nur eingefügt, wenn er als belegt gilt.
3. Der Sichtstrahl startet außerhalb der Box und endet innerhalb. Nur der Teil des Strahls, der sich außerhalb der Box befindet, wird traversiert. Nur ein als belegt geltender Endpunkt wird eingefügt.
4. Der Strahl startet in der Box und endet außerhalb. Nur der Teil des Sichtstrahls außerhalb der Box wird traversiert. Der Endpunkt wird eingefügt; hierbei spielt es keine Rolle, ob er frei oder belegt ist.

5. Der Strahl startet und endet außerhalb der Box, durchläuft diese jedoch. Der erste Teil des Strahls wird traversiert, es wird jedoch kein Endpunkt eingefügt. Der Rest des Strahls wird der Liste noch zu bearbeitender Sichtstrahlen hinzugefügt.

Die Schnitte werden fusioniert, indem die Offsets $t_{start} = 0.0$ und $t_{end} = 1.0$ für Anfang und Ende des neuen Sichtstrahls angepasst werden. Dies erfolgt über die beiden Offsets t_{near} und t_{far}, welche den Schnitt des Strahls mit der nahen und fernen Seite der Bounding Box beschreiben. Beide Variablen werden vom Schnitttest geliefert und dienen unter anderem dazu, die oben genannte Fallunterscheidung durchzuführen. Folgende Aufzählung gibt wieder, wie der Strahl in den oben genannten Fällen angepasst wird:

1. Kein Schnitt, daher keine Anpassung notwendig.
2. Der Sichtstrahl ist komplett in der Bounding Box und wird nicht traversiert. Dies ist ein Abbruchkriterium für das Prüfen gegen weitere Ebenen.
3. Wenn $t_{near} < t_{end}$, dann $t_{end} = t_{near}$. Hierdurch wird sichergestellt, dass ein Sichtstrahl nur zum Start hin verkürzt wird.
4. Wenn $t_{far} > t_{start}$, dann $t_{start} = t_{far}$. Hierdurch wird sichergestellt, dass ein Sichtstrahl nur zum Endpunkt hin verkürzt wird.
5. Siehe Fall 3.

Nach der Kürzung des Sichtstrahls findet die eigentliche Strahlverfolgung statt. Der verbliebene Freiraum wird durchlaufen und die Volumen werden aktualisiert. Im Anschluss wird der freie oder belegte Endpunkt eingefügt; je nachdem, ob die Fallunterscheidung dies für den jeweiligen Typ des Endpunktes zulässt oder nicht.

3.3 Iterative Verfeinerung

Die im vorherigen Abschnitt beschriebene Methode zur Strahlverfolgung mit Ebenenfilter wird hier dazu eingesetzt, eine volumetrische Approximation einer gegebenen Punktwolke aufzubauen. Diese ist frei von Traversionsartefakten an planaren Oberflächen. Das Verfahren ist in eine Initialisierungsphase und eine iterative Verfeinerungsphase unterteilt.

Im Rahmen der Initialisierung wird eine Volumenrepräsentation der Umgebung mit grober, einheitlicher Auflösung aufgebaut. Für ein urbanes Umfeld hat sich eine Voxelgröße von einem Meter als praktikabel erwiesen, da diese Auflösung zum einen grob genug ist, um eine schnelle Laufzeit zu gewähren, zum anderen exakt genug, um die Anzahl der nötigen Verfeinerungsschritte zu minimieren. Die Minimierung der benötigten Verfeinerungsschritte ist auch der Grund, warum eine solche Initialisierung stattfindet. Eine Auflösung von einem Meter lässt sich erzielen, indem bei einer Kachel-Kantenlänge von 32 m die Tiefe 5 im Octree gewählt wird.

Die iterative Verfeinerungsphase findet für jedes Voxel einzeln statt. Hierzu wird das zu verfeinernde Voxel einer Queue zugewiesen. Diese wird solange abgearbeitet, bis sie leer ist; dann gilt das Voxel als maximal verfeinert. Zu Beginn einer Iteration wird das nächste Voxel aus der Queue entnommen. Sind weniger als eine Mindestanzahl an Beobachtungen (d.h. traversierende und darin endende Messstrahlen) darin enthalten, so wird dem Voxel das Label *Restklasse* zugewiesen und keine weitere Verfeinerung durchgeführt. Dieser Schwellwert ist notwendig, da der Voxelzustand umso robuster geschätzt werden kann, je mehr Beobachtungen vorliegen.

Liegt eine Mindestanzahl an Beobachtungen vor und die maximale Tiefe im Octree wurde noch nicht erreicht, dann wird für das Voxel eine naive Zustandsschätzung durchgeführt. Wird das

Voxel nur von Sichtstrahlen durchlaufen, dann wird ihm das Label *frei* zugewiesen. Enden hingegen nur Sichtstrahlen im Voxel, dann kann man davon ausgehen, dass es belegt ist. Da belegte Voxel jedoch weiter zu unterteilen sind, da es sonst nur zu einer sehr groben Annäherung der Punktwolke kommt, wird das Label *aufteilen-belegt* zugewiesen. Tritt der Fall auf, dass ein Voxel von Sichtstrahlen durchlaufen wird und zugleich andere Sichtstrahlen darin enden, dann gehört das Voxel entweder zum Rand einer Struktur oder eines Objekts. Es wird mit dem Label *aufteilen* versehen. Hat das Voxel nach Abschluss der naiven Zustandsschätzung entweder das Label *aufteilen* oder *aufteilen-belegt*, wird es in seine Kindsknoten aufgespaltet. Für jeden dieser Knoten wird eine ebenen-gefilterte Strahlverfolgung durchgeführt. Die Laufzeit des Verfahrens lässt sich rapide senken, indem dafür nur die Sichtstrahlen herangezogen werden, welche bereits das übergeordnete Voxel durchlaufen haben. Anschließend werden die Kindsknoten zwecks weiterer Verfeinerung der Queue hinzugefügt.
Wird ein Voxel der Liste entnommen und dessen Tiefe im Octree entspricht der maximalen Tiefe, dann wird ein finaler Labeling-Schritt vorgenommen. Voxel mit dem Label *aufteilen-belegt* wird das Label *belegt* zugewiesen. Bei Voxeln mit dem Label *aufteilen* kann es sich entweder um stationäre Strukturen handeln, um temporär stationäre Objekte wie geparkte Fahrzeuge, um von einem bewegten Objekt durchlaufenen Freiraum oder aber um Voxel an Objekträndern. Da die korrekte Zuweisung eines Labels keinesfalls trivial ist, wird diese Aufgabe als Gegenstand zukünftiger Arbeiten betrachtet.

4 Experimente

4.1 Verwendeter Datensatz

Die Evaluation der hier beschriebenen Verfahren wird auf Basis des *MLS1 – TUM City Campus* Datensatzes durchgeführt. Dieser wurde im Jahr 2016 mit dem Messfahrzeug MODISSA des Fraunhofer-Instituts für Optronik, Systemtechnik und Bildauswertung (IOSB) aufgezeichnet. Der Datensatz umfasst ca. 2.25 Milliarden Einzelmessungen zweier mobiler LiDAR-Sensoren, verteilt über eine Fläche von 0,17 km². Aufgenommen wurden das Stammgelände der TU München sowie der umliegende Straßenraum.
Die Aufnahmen wurden mit zwei Velodyne HDL-64E LiDAR-Sensoren durchgeführt, die in einem Winkel von 35° auf dem Dach im vorderen Bereich des Fahrzeugs montiert sind. Die Georeferenzierung der Daten basiert auf den Messungen eines Applanix POS LV Navigationssystems, welches zwei GPS-Antennen, sowie eine IMU und einen Raddrehgeber einsetzt. Die Navigationsdaten wurden nachprozessiert, um die Genauigkeit zu verbessern. Der Datensatz ist unter einer Creative Commons Lizenz[3] frei verfügbar.
Für die Evaluation wurde die Kreuzung von Theresienstraße und Luisenstraße nordwestlich des TUM Stammgeländes ausgewählt. Die Kreuzung grenzt an mehrere mehrstöckige Gebäude, Vegetation in Form von Bäumen ist ebenfalls vorhanden. Im Straßenraum befinden sich Verkehrszeichen und Ampeln, Fußgänger und PKWs nehmen am Verkehr teil. Das Messfahrzeug kommt

[3] http://s.fhg.de/mls1

von der Theresienstraße, wartet einige Sekunden an der Kreuzung und biegt dann in die Luisenstraße ab.

4.2 Qualitative Evaluation

Die qualitative Evaluation wird mittels manueller Überprüfung durch einen Beobachter anhand zuvor festgelegter Kriterien durchgeführt. Für dieses Vorgehen gibt es mehrere Gründe. Im Fall der volumetrischen Approximation ist die praktikabelste Möglichkeit nur der Vergleich mit einer bereits bestehenden volumetrischen Repräsentation. Zwar lässt sich prüfen, ob Messpunkte sich innerhalb eines als belegt gekennzeichneten Voxels befinden, für Freiraum jedoch ist dies nicht so ohne weiteres möglich. Beim Vergleich mit einer anderen volumetrischen Repräsentation tut sich zudem die Frage nach dem Grad der Diskretisierung auf, da es für dessen Wahl kein objektives Kriterium gibt. Im Fall der Strahlverfolgung mit Ebenenfilter ist eine Grundwahrheit notwendig, in der planare Flächen oder zumindest Gebäude und ebene Flächen gekennzeichnet sind. Fraglich ist hier jedoch, ab welcher Größe Flächen von Relevanz sind. In beiden Fällen ist es nicht so ohne weiteres möglich, eine objektive Grundwahrheit zu erstellen. Am praktikabelsten erscheint es also, beide Verfahren an einer möglichst komplexen Szene zu testen und dann eine manuelle Überprüfung anhand zuvor festgelegter Kriterien durchzuführen. Für die volumetrische Interpretation wurde folgendes Kriterium definiert:

- Wie gut ist die Approximation der Punktwolke durch die Volumen?

Die Evaluation der Strahlverfolgung mit Ebenenfilter wurde anhand der folgenden Kriterien durchgeführt:

- Welchen Einfluss hat die Strahlverfolgung mit Ebenenfilter auf die Repräsentation?
- Wurden alle relevanten planaren Oberflächen erkannt; d.h. fehlen Oberflächen oder wurden unerwünschte Oberflächen erkannt?
- Wurden planare Oberflächen im Hinblick auf die Segmentierung der Ebenen und Schätzung der Bounding Boxen richtig approximiert?

Die obigen Fragen werden beantwortet durch die Betrachtung einer farbkodierten Volumenrepräsentation in Kombination mit der akkumulierten Punktwolke, aus welcher erstere erstellt wurde. Neben den oben genannten Kriterien wurde auf sonstige Besonderheiten geachtet.

5 Diskussion der Ergebnisse

Qualität der Approximation – Das Ergebnis der volumetrischen Approximation für die im letzten Abschnitt beschriebene Beispielszene wird in Abb. 3 (C) gezeigt. Die Szene enthält viele Objekte, Vegetation und Strukturen und kann damit aus geometrischer Sicht also als komplex angesehen werden. Wie sich erkennen lässt sind auch feine geometrische Details gut erkennbar aufgelöst. Selbst Objekte wie beispielsweise Ampeln und Straßenschilder lassen sich in der volumetrischen Darstellung klar erkennen. Ebenfalls als solche zu erkennen sind Bäume und bewegte Verkehrsteilnehmer, wie etwa der Radfahrer in der Mitte der Kreuzung.

Einfluss der Strahlverfolgung – Um die Vorzüge der Strahlverfolgung mit Ebenenfilter zu demonstrieren wurde dieselbe Szene einmal mit und einmal ohne das Verfahren aufgebaut. Das Resultat kann in Abb. 3 (A, B) betrachtet werden. Im ersten Fall ist zu erkennen, dass ein Groß-

teil der belegten Voxel durch Messstrahlen traversiert wurde. Entsprechende Voxel sind gelb gekennzeichnet. Hierbei handelt es sich um die zu Beginn beschriebenen Traversionsartefakte an planaren Oberflächen. Die Anwendung der Strahlverfolgung mit Ebenenfilter führt dazu, dass alle Voxel einer korrekt identifizierten planaren Oberfläche als belegt erkannt werden. Eine Ausnahme bilden die grau gekennzeichneten Voxel der Restklasse, für die keine ausreichende Menge an Messpunkten vorliegt.

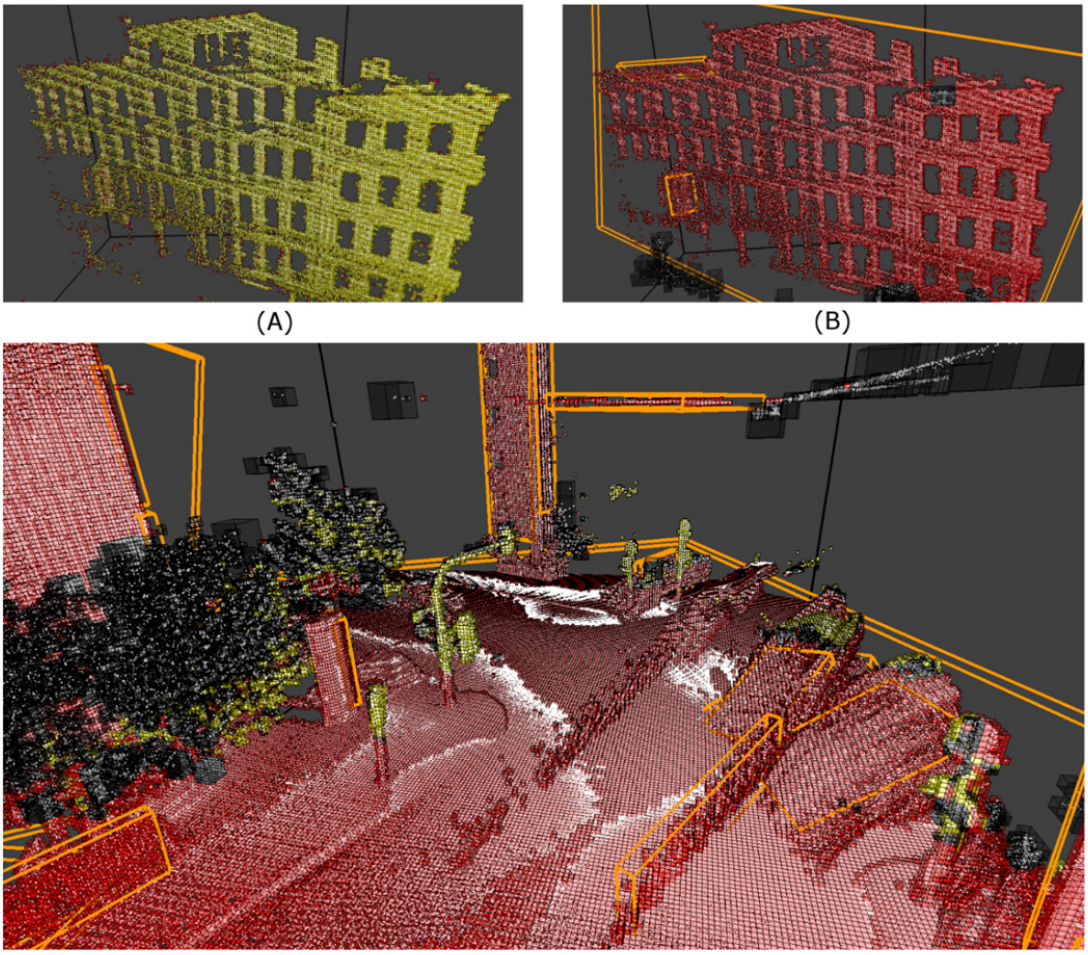


Abb. 3: Volumetrische Approximationen. (A) Ohne und (B) mit Strahlverfolgung mit Ebenenfilter. (C) Eine Beispielszene (rot=belegt, gelb=zu klassifizieren, grau=Restklasse)

Segmentierung und Approximation planarer Oberflächen – Die Methode zur Erkennung planarer Oberflächen ist dazu in der Lage, die für diese Arbeit relevanten dominanten Ebenen wie die Bodenebene und Gebäudefassaden zu identifizieren (vergleiche Abb. 4). Anomalien zeigen sich jedoch an sehr kleinen Flächen und an Ebenen, die nur aus wenigen Punkte bestehen. In einigen seltenen Fällen kommt es auch vor, dass in beliebigen Punktclustern fälschlicherweise Ebenenkandidaten identifiziert werden. Gelegentlich kommt es vor, dass Ebenen in bewegten Objekten erkannt werden. Da es sich bei der zu Grunde liegenden Datenbasis um akkumulierte

Einzelmessungen handelt, ziehen bewegte Objekte eine Art Schweif hinter sich her. Gerade bei PKWs kann die planare Oberfläche des Dachs dazu führen, dass diese auf Grund ihrer großen räumlichen Ausdehnung und des Schweifs als vielversprechender Kandidat erkannt wird.
In einigen wenigen Fällen sind Bounding Boxen nicht korrekt orientiert. Dies lässt sich nicht mit dem populären Fehler zu begründen, dass die Eigenvektoren auf Grund von Eigenarten des Eigenwert-Lösers kein rechtshändiges Koordinatensystem formen. Es wird angenommen, dass sich dieses Verhalten durch die Hauptkomponentenanalyse in Kombination mit der vorliegenden Punktwolke begründen lässt. Das Verhalten kann korrigiert werden, indem weitere Vorverarbeitungsschritte eingefügt werden oder eine Alternative zur Hauptkomponentenanalyse gefunden wird.

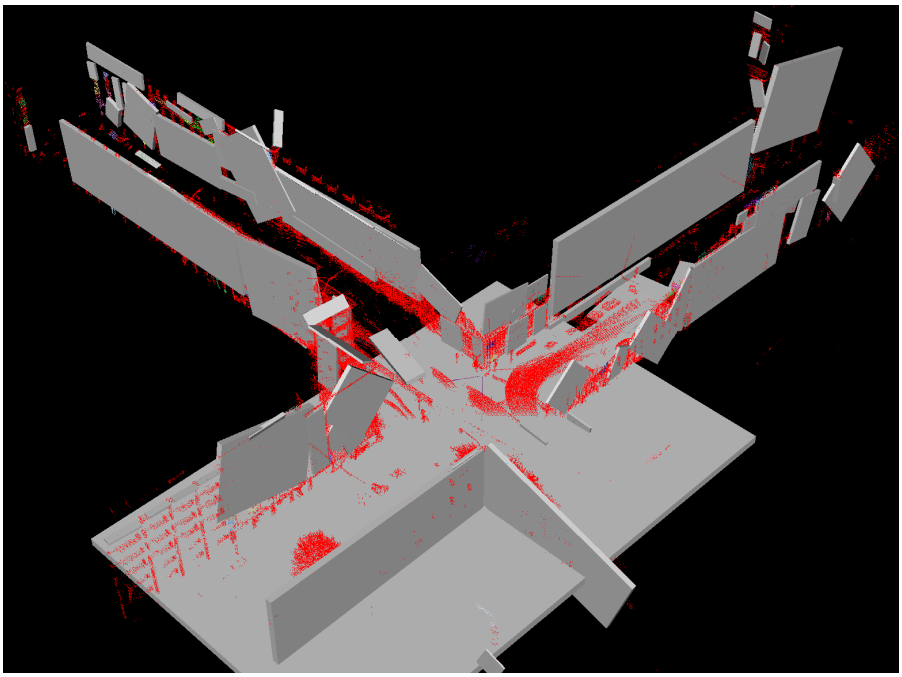


Abb. 4: Planare Oberflächen (grau), extrahiert aus der Punktwolke (rot). Ein Großteil aller planaren Oberflächen ist vorhanden und korrekt approximiert. Einige Bounding Boxen sind nicht korrekt ausgerichtet.

Sonstige Anomalien – Abb. 3 (C) zeigt, dass Traversionsartefakte an Objektgrenzen nach wie vor ein Problem darstellen. Ähnliches gilt auch für Bäume. Derselben Abbildung lässt sich entnehmen, dass diese entweder vom selben Artefakt betroffen sind oder auf Grund zu weniger Einzelmessungen der Restklasse zugewiesen wurden. Weiterhin ist anzumerken, dass bewegte Objekte auf Grund ihrer Nähe zur Bodenebene komplett oder teilweise als solide Struktur erkannt werden. Dies legt den Schluss nahe, dass für Vegetation und bewegte Objekte eine Sonderbehandlung notwendig ist.

6 Fazit & Ausblick

In dieser Arbeit wurde ein Verfahren zum Aufbau einer zur Änderungsdetektion geeigneten, auf Voxeln basierten Metadatenstruktur vorgestellt, bei dem eine volumetrische Approximation einer Punktwolke unter Nutzung einer Methode zur Strahlverfolgung mit Ebenenfilter durchgeführt wird. Letztere ermöglicht es, durch die Diskretisierung des Raums verursachte Artefakte an planaren Oberflächen zu verhindern. Eine Auswertung der Ergebnisse beider Verfahren hat ergeben, dass sich aus geometrischer Sicht komplexe Strukturen durch Volumen annähern lassen. Wichtige planare Oberflächen in der volumetrischen Repräsentation wie die Bodenebene und Häuserfassaden sind frei von Diskretisierungsartefakten. An Objektgrenzen zeigen sich wie erwartet weiterhin Artefakte. Dies und auch die Tatsache, dass gelegentlich planare Oberflächen in bewegten Objekten erkannt werden, legt eine Sonderbehandlung für diesen Fall nahe. Selbiges gilt für Vegetation. Detektiert man beide Objektklassen bereits vor Aufbau der Volumenrepräsentation in den Messdaten, so kann man diese gesondert einpflegen. Dies vereinfacht den Umgang mit den oben beschriebenen Artefakten deutlich.

7 Literaturverzeichnis

AIJAZI, A.K., CHECCHIN, P. & TRASSOUDAINE, L., 2013: Detecting and Updating Changes in Lidar Point Clouds for Automatic 3D Urban Cartography. ISPRS - Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, **II**-5/W2, 7-12.

AZIM, A. & AYCARD, O., 2012: Detection, classification and tracking of moving objects in a 3D environment. Proceedings of the IEEE Intelligent Vehicles Symposium (IV), 802-807.

GEHRUNG, J., HEBEL, M., ARENS, M. & STILLA, U., 2016: A Framework for Voxel-based Global Scale Modeling of Urban Environments. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **42**(2/W1), 45-51.

GEHRUNG, J., HEBEL, M., ARENS, M. & STILLA, U., 2017: An approach to extract moving objects from MLS data using a volumetric background representation. ISPRS - Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, **IV**-1/W1, 107-114.

GIRARDEAU-MONTAUT, D., ROUX, M., MARC, R. & THIBAULT, G., 2005: Change detection on point cloud data acquired with a ground laser scanner. ISPRS - International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, **36**(3/W19), 30-35.

HEBEL, M., ARENS, M. & STILLA, U., 2013: Change detection in urban areas by object-based analysis and on-the-fly comparison of multi-view ALS data. ISPRS Journal of Photogrammetry and Remote Sensing, **86**, 52-64.

HORNUNG, A., WURM, K.M., BENNEWITZ, M., STACHNISS, C. & BURGARD, W., 2013: OctoMap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees. Autonomous Robots, **34**, 189-206.

LITOMISKY, K. & BHANU, B., 2013: Removing Moving Objects from Point Cloud Scenes. Advances in Depth Image Analysis and Applications, Springer Berlin Heidelberg, 50-58.

PAGAC, D., NEBOT, E.M. & DURRANT-WHYTE, H., 1996: An evidential approach to probabilistic map-building. Proceedings of IEEE International Conference on Robotics and Automation, **1**, 745-750.

RABBANI, T., VAN DEN HEUVEL, F.A. & VOSSELMAN, G., 2006: Segmentation of Point Clouds Using Smoothness Constraint. ISPRS: Proceedings of the ISPRS Commission V Symposium, **35**, 248-253.

RABINER, L. R., 1989: A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, **77**, 257-286.

SCHACHTSCHNEIDER, J., SCHLICHTING, A. & BRENNER, C., 2017: Assessing temporal behavior in LIDAR point clouds of urban environments. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **42**(1/W1), 543-55.

UNDERWOOD, J. P., GILLSJÖ, D., BAILEY, T. & VLASKINE, V., 2013: Explicit 3D change detection using ray-tracing in spherical coordinates. IEEE International Conference on Robotics and Automation, 4735-4741.

WILLIAMS, A., BARRUS, S., KEITH, R. & SHIRLEY, M.P., 2003: An efficient and robust ray-box intersection algorithm. Journal of Graphics Tools, **10**, 54.

WOLF, D. & SUKHATME, G.S., 2004: Online simultaneous localization and mapping in dynamic environments. 2004 IEEE International Conference on Robotics and Automation, **2**, 1301-1307.

XIAO, W., VALLET, B., BRÉDIF, M. & PAPARODITIS, N., 2015: Street environment change detection from mobile laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, Elsevier, **38**, 38-49.

ZEIBAK, R. & FILIN, S., 2008: Change detection via terrestrial laser scanning. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, **36**, 430-435.

Traffic sign detection and recognition with convolutional neural networks

ALEXANDER HANEL[1] & UWE STILLA[1]

Abstract: Traffic sign detection and recognition in street scene images of a vehicle camera allow to localize the traffic signs in a street scene image and to classify them with regard to their semantic meaning for the car driver. In this contribution, a method is described to detect traffic signs in a street scene image by evaluating image patches, sampled by a sliding window approach, with a convolutional neural network detector. Robust shape fitting is performed on the image patch of a positive detection to obtain the exact position of the traffic sign shape in the patch. A second convolutional neural network is applied to the image patches centred on the fitted shapes to classify the meaning of these traffic signs. The networks are trained and tested with samples of traffic signs and other street scene objects from the GTSRB and GTSDB datasets. The results have shown that models for traffic sign detection and recognition can be trained with an overall accuracy of more than 90 % obtained for a test set. The position of a traffic sign, known from shape fitting, has been shown to be an important a-priori knowledge to select the appropriate image patch to ensure a high accuracy of the subsequent traffic sign recognition.

1 Safer roads by traffic sign warnings

Modern advanced driver assistance systems are an important contribution to increase the safety on roads. An assistance system using traffic sign recognition can warn a car driver against a speed limit by means of optical hints (EURO NCAP 2013), for example. As well as a human driver, an assistance system needs to extract information from the environment around the car. Cars can therefore be equipped with sensors to perceive the environment. As the upcoming drive way and the traffic signs applicable for that way are in front of the vehicle, a front-looking sensor is suitable for traffic sign detection and recognition (e.g. LARSSON & FELSBERG 2011). Due to small costs compared to a Radar or LiDAR sensor (ETKBMW.COM), often an optical mono-camera is the first choice for this task.

Such a camera observing the environment acquires typically an image of the street scene including the street, traffic signs, buildings, traffic participants and further static and dynamic objects (Fig. 1). There exist algorithms for object detection which use the complete image of the street scene as input (e.g. REDMON et al. 2016). Other algorithms require an image patch centered on the object which should be detected (red rectangle in Fig. 1) as input (e.g. SERMANET & LECUN 2011). In the second case, the input image patch can be obtained by a sliding window approach sampling from the street scene image, for example.

Traffic sign detection and traffic sign recognition have different objectives (ZHU et al. 2016): the objective of traffic sign detection is to decide whether the input image, or the image patch, shows

[1] Technical University of Munich, Photogrammetry & Remote Sensing, Arcisstraße 21, D-80333 Munich, email: [alexander.hanel, stilla]@tum.de

a traffic sign or another arbitrary object. In contrast, the objective of traffic sign recognition is typically to classify the meaning of a traffic sign, given an image patch from which it is known to show a traffic sign (e.g. German Traffic Sign Recognition Benchmark dataset, STALLKAMP et al. 2012).

Fig. 1: Example image of a street scene recorded by a front-looking mono-camera. The image is an excerpt from the Cityscapes dataset (CORDTS et al. 2016), which can be used to train machine learning algorithms in the domain of advanced driver assistance systems, for example for traffic sign recognition

In a recent work, ZHU et al. (2016) propose a method for simultaneous detection and recognition of traffic signs using a neural network-based deep learning approach. Their method provides the bounding box, a pixel-wise mask and the class label for each detection. Several examples (in supplementary material provided by these authors) show that the bounding boxes of partly occluded traffic signs tend to be smaller than the sign in the image is. Using the pixels of the mask to extract the shape contour of the traffic sign will therefore be prone to errors.
HANEL & STILLA (2018) have recently proposed a method to calibrate a vehicle camera on public roads, where the exact shape of a traffic sign in the image has to be known precisely. Their method derives image points along the shape of a sign and the corresponding metric 3D position, for example known from governmental regulations, as reference points for calibration.
In this contribution, a method for traffic sign detection and recognition is proposed, which delivers the correct shape of the sign in the image simultaneously. A sliding window approach is used to sample image patches from a street scene image. The samples are evaluated by a traffic sign detector, which is realized by a neural network-based deep learning approach. Deep approaches have shown to provide a higher accuracy than shallow learning approaches, as for example ZHU et al. 2016 mention. Mean shift clustering is used to reduce multiple detections of the same traffic. Such multiple detections with slightly different positions and sizes are typical for a sliding window approach (COMASCHI et al. 2013). The shape of the traffic sign in the image is obtained by fitting an appropriate geometric primitive (e.g. an ellipse for circular traffic signs) with RANSAC followed by a least-squares adjustment. The extracted shape is used to refine the image position of the patch, which is then used by another deep learning approach to classify the meaning of the traffic sign.

2 Stepwise traffic sign detection and recognition

By the sliding window approach, image patches are taken from an image with a higher geometric resolution (called 'complete image' in the following). In the scope of this contribution, the complete image is typically a street scene image. Each patch shows a different part of the complete image; the patch is shifted over the complete image and the complete image is resized to consider traffic signs at different image positions and with different sizes in the image. The deep network architecture (Fig. 2) of the traffic sign detector uses such an image patch as input sample. In contrast to architectures, which use the complete image as input, the sliding window approach allows to reduce the computational effort when tracking traffic signs in an image series by evaluating only the local neighbourhood around the predicted image position of a traffic sign in a subsequent image in contrast to evaluating the complete subsequent image.

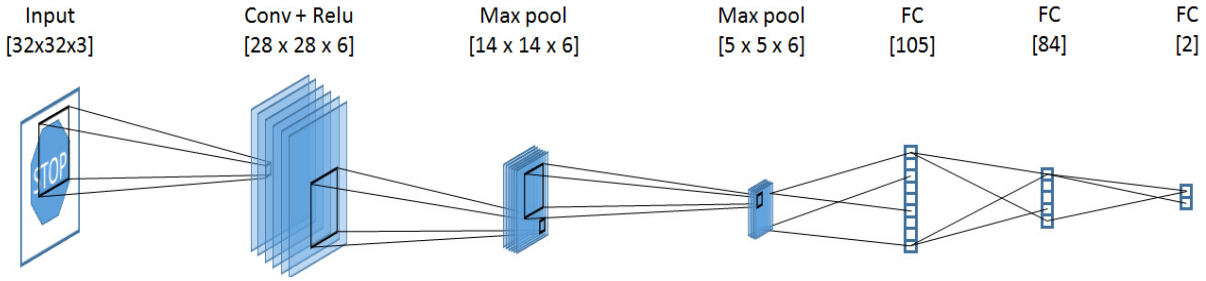


Fig. 2: Architecture of the convolutional neural network used to distinguish between *traffic signs* and *other objects* shown in an image patch taken from a street scene image.

The traffic sign detector classifies each sample in either the class *traffic sign* or *other objects*. The architecture of our network as described in the following is a variation of the traffic sign detector architecture described by WU et al. (2013): the input sample with a fixed resolution of 32 x 32 pixels and RGB color channels is first convolved with different kernels to extract features. The "ReLu" activation function ensures the non-linearity property of the network to prevent that the network behaves just like a single-layer perceptron. Max pooling in the second and third layer reduces the resolution of the feature maps created by the previous convolutions. Dropout is applied to the third layer to avoid overfitting. The following layers are designed to classify the input patch based on the extracted features: the output of the feature extraction layers is flattened, followed by three fully-connected (FC) layers in which the number of neurons is reduced step-by-step. The last FC layer has two output neurons, one for each of the two classes to distinguish. Softmax loss is used. In contrast to by WU et al. (2013), we use no branching (for details on branching see the description of the traffic sign recognition architecture), but increase the depth of the classification part. ZHU et al. (2016) have mentioned that deeper networks perform better. According to these authors, branching could also increase the performance, but at the cost of a higher computational effort.

The detector model is trained with a set of positive and negative samples. Image patches showing one traffic sign are called positive samples, while image patches showing arbitrary objects like buildings or vegetation are called negative samples. The class of an image patch with unknown content is predicted after training by feeding the patch together with the trained model into the

neural network. Output of the prediction are the probabilities for both classes; the image patch is assigned to the class with the higher probability.
Image patches from slightly different positions and sizes in the complete image are likely to show the same object. Therefore, it can be expected that several of such neighboring patches will be detected as *traffic sign*. To link all detections belonging to the same traffic sign, the detections are grouped to a single detection by mean shift clustering (FUKUNAGA & HOSTETLER 1975). Each detection is hereby represented by the central point of the image patch. Mean shift clustering groups the central points of all detections into several groups using the Euclidean distance as criterion to distinguish different groups. The size of the image patch belonging to a single detection is determined by the mean size of the patches of all contributing detections. An advantageous property of mean shift clustering is that it does not require the number of clusters to be determined manually, being therefore able to handle a varying number of traffic signs in different complete images.

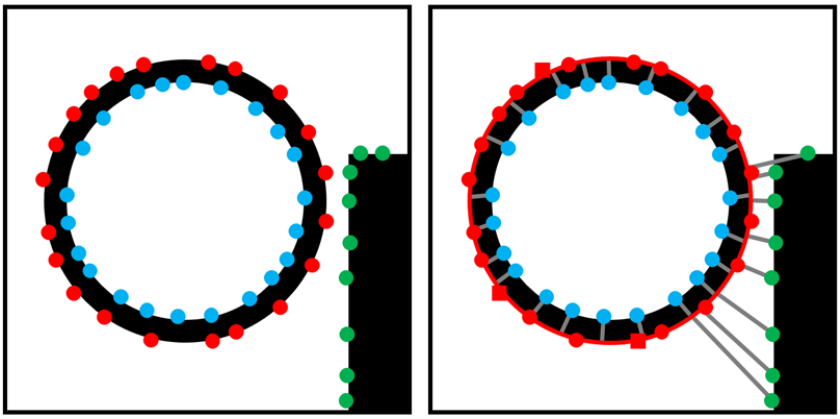


Fig. 3: Left: Edge of a traffic sign (black circle) and a background object (black rectangle) in a binary image patch. Contour points (red, blue, green balls) are along the edges of these objects. Right: Ellipse (red circle) fitted with sample points (red squares) selected from the contour points. The distance (grey lines) threshold for each contour point from the ellipse defines the inlier set used for the final-least-squares ellipse fit

As next step, the shape of the traffic sign in the image patch of each single detection is extracted by ellipse fitting. Result of the shape extraction is the position, size and orientation of the ellipse in the complete image used as precise location of a traffic sign prior to classifying the meaning of the sign. Further, mean shift clustering and shape fitting could be used to recognize false positive detections (e.g. by a small consensus set, high eccentricity of the ellipse, ...). The method for shape extraction is in the following described for circular traffic signs to allow analyses on the potential of this approach before further research is done. The method assumes that an image patch shows a traffic sign completely. To ensure this, the part of the complete image the patch covers can be increased. The ellipse is selected as geometric primitive as a real-world circle is projected to an ellipse in an image in general (e.g. AHN et al. 1999). The ellipse is fitted to groups of contour points extracted from the binary image patch. The image patch is binarized by applying an absolute global intensity threshold to it to separate objects with different unique intensities. The algorithm of SUZUKI & ABE (1985) is used to extract a group of contour points

(Fig. 3 left) along the edges of each image region with unique intensity in the binary image. Groups with a small number of points are neglected, assuming that the traffic sign is the dominant region with a high number of points compared to other groups in the patch. The RANSAC algorithm (FISCHLER & BOLLES 1981) is applied to randomly chosen contour points to select iteratively the consensus set of contour points belonging to the edge of the traffic sign shape. The inliers are defined by a maximal distance of the contour points to the ellipse in each iteration (Fig. 3 right). The largest set of inliers in all iterations is used as consensus set. The number of iterations is chosen to have chosen at least one set of sample points without outliers with a probability of 99 %. The final ellipse parameters are estimated in a least-squares-adjustment using the consensus set.

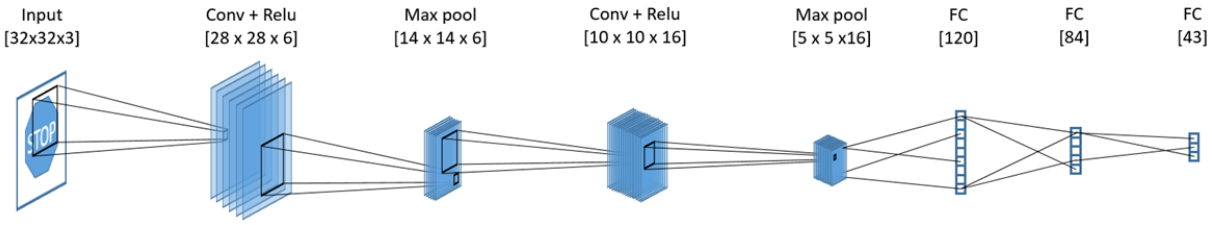


Fig. 4: Architecture of the neural network used to recognize the semantic meaning of a traffic sign shown in an image patch. In total, 43 different meanings are considered.

To recognize the semantic meaning of a traffic sign shown in an image patch, another machine learning method based on a convolutional neural network is used (Fig. 4). Our network is adapted from SERMANET & LECUN (2011). Its architecture can be divided in a feature extraction and a classification part as well. In the feature extraction part of our network, the so-called *skip architecture* used by these authors is left out. Thereby, the output of a layer would be branched and fed to a later layer by skipping some layers in between, while another branch uses all layers in between. More global features of traffic signs could be extracted by using branching. Instead, we increase the depth in the classification part, i.e. we increase the number of FC layers to three. The number of output neurons in the last layer is determined by the number of classes of traffic sign meanings in *The German Traffic Sign Recognition Benchmark* dataset (STALLKAMP et al. 2012).

Fig. 5: Positive samples (taken from the GRSRB traffic sign recognition dataset). Left: Triangle-shaped construction warning sign in the shadow. Center: Circle-shaped go straight sign in sunny light. Right: Negative sample showing vegetation (extracted from street images from the GTSDB dataset)

3 Dataset and experiments

The models for the detector are trained with a set of sample patches showing one traffic sign or other objects, respectively. Separate models are trained for the following groups of traffic signs with different shapes and dominant colours: prohibitive signs (circle, red), mandatory signs (circle, blue), danger (triangle, red) and others (e.g. right of way, give way). One joint model for traffic sign recognition is trained with samples showing traffic signs with various meanings. The traffic sign samples (example in Fig. 5 left and centre) are taken from *The German Traffic Sign Recognition Benchmark* dataset (STALLKAMP et al. 2012). This dataset provides traffic sign image patches in various daylight illumination situations (e.g. sunny, shadowy) with a roughly frontal view on them, which would be also typical for a front-looking vehicle camera. The negative samples (example in Fig. 5 right) are extracted by randomly sampling image patches from street scene images from *The German Traffic Sign Detection Benchmark* (GTSDB) dataset (HOUBEN et al. 2013). The order of the samples is randomly shuffled, a subset of 80% of the samples is used for training and validation. The remaining subset of 20% is used for tests. The training is performed with a learning rate of 0.001 and a dropout rate of 0.5. 80 epochs are used in training, until there is no remarkable decrease in the loss anymore. The hyperparameter values have been obtained by hand-tuning. The proposed method is tested with 900 street scene images of the GTSDB dataset, for which the ground truth position and meaning of traffic signs are known.

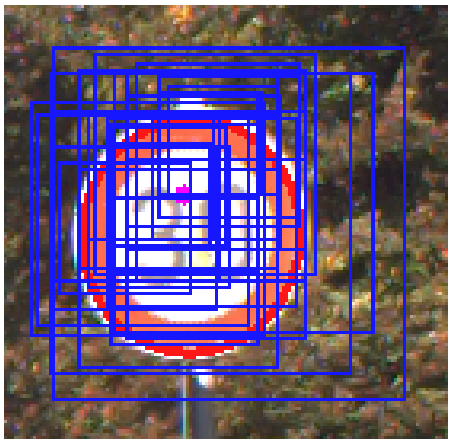


Fig. 6: Multiple detections (blue rectangles) for one traffic sign. The central point (pink dot) obtained by mean shift clustering lies above the middle of the sign, leading to a possibly wrong position of the image patch for the following traffic sign detection. The fitted ellipse (sharp red circle) matches with the shape of the sign and allows therefore to extract the image patch for traffic sign recognition at the appropriate position

4 Results of traffic sign detection and recognition

The trained models for both traffic sign detection and recognition can obtain high overall accuracy values on the respective test set (97.2%, 93.4%). The approximately balanced number of samples in all classes ensures that the obtained overall accuracies do not result from models deciding always for the same class. It can be observed for the test set, that traffic signs with a very similar

appearance (e.g. different speed limit signs) are the most challenging ones for traffic sign recognition. Further experiments using the skip architecture mentioned above in the neural networks have not shown to result in a higher overall accuracy.
As expected, traffic sign detection yields typically multiple detections for one traffic sign shown in a street scene image (Fig. 6). The larger the traffic sign in the image is, the higher the number of detections is. Mean shift clustering is able to reduce the detections to one cluster for large traffic signs (pink dot in Fig. 6). Critical attention has to be payed for remote traffic signs (i.e. small in the image), as observations (done using visualisations like in Fig. 6) have shown that adjacent small traffic signs might be grouped in one cluster by mistake. Similar observations can be made for ellipse fitting, which can be seen as robust for close traffic signs (i.e. large in the image).
Mean shift clustering and shape fitting as intermediate steps can be seen as useful to reduce the number of false positives resulting from traffic sign detection. Applied to the street scene images of the GTSDB dataset, the number of false positives would be 20 % higher if these intermediate steps are skipped. Hereby, a detection is seen as true positive if its image patch is overlapping with more than 50 % of the ground truth bounding box of a traffic sign. The comparison is done using the number of false positives instead of the false positive rate, as the high number of image patches resulting from the sliding window approach leads to a high number of true negatives and would lead to very similar false positive rates, which could be misinterpreted. Further, using the intermediate steps leads to a slightly higher number of true positives detections, which could be drawn back that adjusting the position and size of the image patch of a single detection, resulting from the intermediate steps, leads to a greater overlap with the ground truth label. Consequently, the number of false negatives is slightly reduced.
A further experiment underlines the demand that the image patch of a detection has to be centred precisely on the traffic sign for traffic sign recognition: the overall accuracy of traffic sign recognition decreases by 20 percentage points, if a test set of image patches is used, which are not centred on a traffic sign (randomly shifted by up to $\pm$50 % of the patch size), compared to using a set of patches, which are centred.

5 Conclusion

In this contribution, a method for traffic sign detection and recognition in street scene images using two separate convolutional neural networks has been proposed. Detection with different models for different groups of traffic signs (e.g. red prohibitive signs, blue mandatory signs) and recognition with one joint model for all traffic signs can achieve an overall accuracy of around 95 % in average on the tested set of image patches. Applying the traffic sign detector to image patches extracted from a street scene image using a sliding window approach yields typically multiple detections for the same traffic sign. Without clustering multiple detections to a single detection and adjusting the position of the image patch by fitting an ellipse (for circular traffic signs) to the shape of a traffic sign detection, the number of false positive detections is 20 % higher for the test images. Future extensions to the proposed method can be made by integrating a traffic sign tracker to reduce the computational effort fort detection by reducing the search space for the detector in subsequent images of an image series. Furthermore, a method to fit ad-

ditional shapes (e.g. rectangle, triangle) to the image patches of positive detections could be integrated to make the method better usable for varying types of traffic signs.

6 References

AHN, S. J., WARNECKE, H.-J. & KOTOWSKI, R., 1999. Systematic geometric image measurement errors of circular object targets: mathematical formulation and correction. Photogrammetric Record, **16**(93), 485-502.

COMASCHI, F., STUIJK, S., BASTEN, T. & CORPORAAL, H., 2013. RASW: a Run-time Adaptive Sliding Window to Improve Viola-Jones Object Detection. 2013 Seventh International Conference on Distributed Smart Cameras (ICDSC), 1-6.

CORDTS, M., OMRAN, M., RAMOS, S., REHFELD, T., ENZWEILER, M., BENENSON, R., FRANKE, U., ROTH, S. & SCHIELE, B., 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3213-3223.

ETKBMW.COM, 2018. Distance Systems, Cruise Control BMW G11 sedan 57553. Website, https://www.etkbmw.com/bmw/EN/search/selectCar/G11/Lim/BMW+750i/ECE/66. Accessed 2018-01-16.

EURO NCAP, 2013: Speed Assistance Systems. Website, https://www.euroncap.com/en/vehicle-safety/the-ratings-explained/safety-assist/speed-assistance/. Accessed 2018-01-16.

FUKUNAGA, K. & HOSTETLER, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Transactions on Information Theory, **21**(1), 32-40.

HANEL, A. & STILLA, U., 2018. Iterative calibration of a vehicle camera using traffic signs detected by a convolutional neural network. Proceedings of the 4th International Conference on Vehicle Technology and Intelligent Transport Systems. Paper accepted; to be published

LARSSON, F. & FELSBERG, M., 2011. Using Fourier Descriptors and Spatial Models for Traffic Sign Recognition. Image Analysis: Proceedings of 17th Scandinavian Conference, SCIA 2011, Anders, H. & Kahl, F. (eds.), Springer Berlin / Heidelberg, 238-249.

REDMON, J., DIVVALA, S., GIRSHICK, R. & FARHADI, A., 2016. You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788.

SERMANET, P. & LECUN, Y., 2011. Traffic sign recognition with multi-scale Convolutional Networks. The 2011 International Joint Conference on Neural Networks, 2809-2813.

STALLKAMP, J., SCHLIPSING, M., SALMEN, J. & IGEL, C., 2012. Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition. Neural Networks **32** (Supplement C), 323-332.

WU, Y., LIU, Y., LI, J., LIU, H. & HU, X., 2013. Traffic sign detection based on convolutional neural networks. The International Joint Conference on Neural Networks (IJCNN), 1-7.

ZHU, Z., LIANG, D., ZHANG, S., HUANG, X. & HU, S., 2016. Traffic-Sign Detection and Classification in the Wild. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2110-2118.

Fußgängerbezogene Informationsgewinnung zur Situationsanalyse mit einem mobilen Multisensorsystem

BJÖRN BORGMANN[1,2], MARCUS HEBEL[1], MICHAEL ARENS[1] & UWE STILLA[2]

Zusammenfassung: Fußgänger gehören im städtischen Verkehrsraum nicht nur zu den verwundbarsten Verkehrsteilnehmern, sondern sind auch aufgrund ihrer Größe häufig nur schwer wahrzunehmen. Technische Systeme wie z.B. autonome Fahrzeuge oder Fahrerassistenzsysteme sind daher besonders gefordert diese zu erkennen. Solche Systeme verfügen häufig über eine Vielzahl an Sensoren um eine Abdeckung des kompletten Umfelds zu erreichen. In solchen Fällen ist eine Form von Datenfusion zwischen den Sensoren notwendig. Der vorliegende Beitrag stellt ein Verfahren zur Personendetektion in 3D-LiDAR Daten eines MLS-Systems vor, welches mehrere LiDAR parallel nutzt, um Personen zu detektieren. Es geht dabei mit Verzerrungseffekten um, die durch eine unvollständige zeitliche Synchronisation der eingesetzten Sensoren in deren Überschneidungsbereichen auftreten können. Solche Effekte lassen sich insbesondere bei scannenden Systemen kaum vermeiden. In einer experimentellen Untersuchung der vorgestellten Methode wird gezeigt, dass sie beim vermehrten Auftreten der beschriebenen Verzerrungseffekte bessere Ergebnisse liefert als einfachere Ansätze der Datenfusion.

1 Einleitung

Fußgänger sind eine bedeutende Gruppe von Verkehrsteilnehmern im städtischen Umfeld. Da sie im Vergleich mit anderen Verkehrsteilnehmern nicht nur verwundbarer sind, sondern außerdem aufgrund ihrer kleineren Größe leichter übersehen werden, sollte ihnen ein besonderes Augenmerk zuteilwerden. Hier besteht auch bei technischen Systemen, z.B. bei Fahrerassistenzsystemen oder autonomen Fahrzeugen, ein Bedarf, Fußgänger automatisch und in Echtzeit zu detektieren. Für solche oft mobilen technischen Systeme, die im Umfeld von Personen verwendet werden, hat eine spezifische Detektion von Personen gegenüber einer unspezifischen Hindernis- bzw. Bewegtobjektdetektion verschiedene Vorteile: Es wird zum einen so möglich, auf die Sicherheit von Fußgängern besonderen Wert zu legen, was zum Beispiel in Form eines größeren Sicherheitsabstands erfolgen kann. Außerdem kann so das besondere Bewegungsverhalten von Personen berücksichtigt werden. So kann eine Person z.B.

Abb. 1: Eine Person aufgenommen von zwei LiDAR-Sensoren

[1] Fraunhofer IOSB, Abteilung Objekterkennung, Gutleuthausstr.1, D-76275 Ettlingen, E-Mail: [bjoern.borgmann, marcus.hebel, michael.arens]@iosb.fraunhofer.de
[2] Technische Universität München, Photogrammetrie und Fernerkundung, Arcisstraße 21, D-80333 München, E-Mail: stilla@tum.de

viel abrupter die Richtung wechseln, anhalten oder beschleunigen als ein Fahrzeug.
Oben beschriebene technische Systeme verfügen oft über eine Vielzahl unterschiedlicher Sensoren. Dazu gehören neben Kameras für das sichtbare und infrarote Licht vielfach auch Radar und LiDAR-Sensoren (engl. *Light Detection and Ranging*, z.B. Laserscanner). In der Praxis sind solche Systeme oft auch mit mehreren Instanzen der gleichen Sensorart ausgestattet, um z.B. eine bessere Rundumabdeckung zu erzielen, die sonst konstruktionsbedingt oft nicht erreicht werden kann. Im Überschneidungsbereich mehrerer Sensoren wird zudem eine höhere Datendichte erzielt. Durch die Verwendung verschiedener Sensorarten kann von den individuellen Vorteilen der verschiedenen Sensoren profitiert werden. Die Verwendung mehrerer Sensoren stellt aber auch eine zusätzliche Herausforderung dar. So muss eine Form von Datenfusion erfolgen. Selbst bei der Fusion gleichartiger Sensoren kann es aufgrund einer nicht perfekten geometrischen Registrierung oder einer unvollständigen zeitlichen Synchronisation der Sensoren zu Verzerrungseffekten kommen. Wenn sich solche nicht vermeiden lassen, sollte die weitere Datenverarbeitung mit ihnen umgehen können. Abb. 1 stellt das Ergebnis einer direkten Datenfusion zwischen zwei scannenden LiDAR-Sensoren dar, d.h. eine räumliche Überlagerung der aktuellen Punktwolken. Ein Teil der Person wurde vom ersten Sensor erfasst, ein anderer Teil leicht zeitlich versetzt vom zweiten. Da scannende Sensoren ihre Umgebung nach und nach abtasten ist eine wirkliche zeitliche Synchronisation zwischen zweien solcher Sensoren in der Praxis nur schwer zu erreichen.
Die vorliegende Arbeit beschäftigt sich mit einem Verfahren zur Detektion von Fußgängern in MLS-Daten (engl. *Mobile Laser Scanning*) eines mobilen Sensorträgers, der über mehrere LiDAR-Sensoren verfügt. Das Verfahren versucht dabei, die gerade beschriebenen Probleme der Verzerrungseffekte zu lösen.

2 Verwandte Arbeiten

Im Bereich der Objekt- bzw. Personendetektion gibt es eine Vielzahl verschiedener Verfahren. Häufig wird auf eine Verarbeitungskette zurückgegriffen, in der die Daten zunächst segmentiert und die Segmente dann klassifiziert werden. Eine Gruppe von Verfahren verwenden als Klassifikator *Support Vector Machines* (SVM). Bei diesen wird in einem Trainingsprozess eine Hyperebene ermittelt, die die Trainingsdaten in einem Merkmalsraum in verschiedene Klassen trennt. Die ermittelte Hyperebene wird dann verwendet um unbekannte Daten zu klassifizieren. NAVARRO-SERMENT et al. (2010) verwenden zwei hintereinander angewendete SVMs zur Klassifizierung von Punktwolkensegmenten. Die erste SVM verwendet dabei eine Reihe geometrischer Merkmale als Eingabe. Ihre Ausgabe wird dann um verschiedene Merkmale aus einem Trackingverfahren ergänzt und als Eingabe der zweiten SVM verwendet um Personen zu erkennen. PREMEBIDA et al. (2014) verwenden eine Datenfusion zwischen LiDAR- und RGB-Sensoren. Bei dieser werden die LiDAR-Sensoren genutzt um für die RGB-Bilder Tiefeninformationen zu liefern. Die SVM wird dann auf diese Tiefenbilder angewendet.
Eine weitere Gruppe von Klassifikatoren greifen auf „Wörterbücher“ zurück, die im Rahmen eines Trainingsprozesses generiert werden. Die „Wörter“ dieser Wörterbücher sind üblicherweise über einen Merkmalsvektor beschrieben und Stimmen über die Klassifizierung der verarbeiteten Daten ab. BEHLEY et al. (2013) verwenden eine Reihe von Wörterbüchern um Segmente

von LiDAR-Daten zu klassifizieren. Die Wörterbücher verwenden dabei unterschiedliche Merkmalstypen bzw. unterschiedlich parametrisierte Merkmale. Dies erlaubt es, mit verschiedenen Eigenschaften der verarbeiteten Segmente umzugehen, z.B. mit unterschiedlicher Datendichte. Es ist denkbar, ein solches Verfahren für die Verarbeitung von verzerrten Daten, wie sie z.B. aus der Fusion mehrerer Sensoren entstehen, zu erweitern und auch hierfür Wörterbücher speziell zu trainieren.
Eine Erweiterung der wörterbuchbasierten Methoden stellen *Implicit Shape Models* (ISM) dar. Diese berücksichtigen nicht nur das Vorhandensein bestimmter Merkmale, sondern auch deren geometrische Verteilung. Sie tun dies, indem die Wörter nicht mehr nur für eine Klasse stimmen, sondern auch für eine Position der Klasse relativ zu der Position des Wortes. Eine Klassifizierung findet dann statt, wenn vermehrt für eine ähnliche Position gestimmt wird. ISM wurden ursprünglich für die Verarbeitung von Bilddaten verwendet (LEIBE et al. 2008), finden aber auch für die Verarbeitung von 3D-Daten Verwendung. KNOPP et al. (2010) verwenden ein 3D-ISM und 3D-SURF Merkmale zur generellen Klassifikation von Objekten in 3D-Daten. VELIZHEV et al. (2012) verwenden eine Verarbeitungskette für Punktwolken, die über eine Bodenextraktion und eine auf dem *Region Growing* (engl. Regionenwachstum) basierende Segmentierung verfügt, um anschließend die Punktwolkensegmente mithilfe von ISM in Objektklassen wie Fahrzeug oder Straßenlaternen zu klassifizieren. Bei den von ihnen verarbeiteten Daten handelt es sich um eine Fusion von mehreren zeitlich aufeinanderfolgend aufgenommenen Punktwolken aus einem urbanen Gebiet. Bewegte Objekte werden daher von ihnen nicht erkannt. Wir selbst haben einen ähnlichen Ansatz verwendet um Personen in Einzelscans eines MLS-Systems zu erkennen (BORGMANN et al. 2017).
SHOTTON et al. (2011 & 2013) verwenden *Random Decision Forests* um einzelne Pixel eines Tiefenbildes zu klassifizieren. Sie nutzen dies um sowohl Personen als auch deren Körperteile zu detektieren und zu verfolgen. Die Tiefenbilder sind dabei verhältnismäßig hoch aufgelöst. Durch die Klassifizierung einzelner Pixel entfällt die Notwendigkeit zuvor eine Segmentierung durchzuführen. Bei Random Decision Forests handelt es sich um eine Ansammlung von mit unterschiedlichem Zufallselement trainierten Entscheidungsbäumen. Durch das Zufallselement und die Verwendung mehrerer Bäume wird das Problem der Überspezialisierung einzelner Entscheidungsbäume gelöst.
In den letzten Jahren werden vermehrt *Convolutional Neural Networks* (CNN) zur Objekterkennung sowohl in RGB-Bildern als auch in Tiefenbildern (SOCHER et al. 2012) und in Volumenrepräsentationen von 3D-Daten (MATURANA & SCHERER 2015, GARCIA-GARCIA et al. 2016) eingesetzt.

3 Methode

In diesem Abschnitt stellen wir unsere Methode zur Detektion von Personen (oder anderen Objekten) in LiDAR-Daten vor. Wir gehen dabei davon aus, dass die entsprechenden Daten durch mehrere LiDAR-Sensoren parallel erfasst werden und für die Sensoren getrennt als zeitlich aufeinanderfolgende Punktwolken vorliegen. Eine Punktwolke umfasst dabei jeweils einen Einzelscan (z.B. eine Rotation) des jeweiligen Sensors. Die Sensoren sind zueinander geometrisch kalibriert, aber wie in der Einleitung dieser Arbeit beschrieben nicht exakt zeitlich synchronisiert

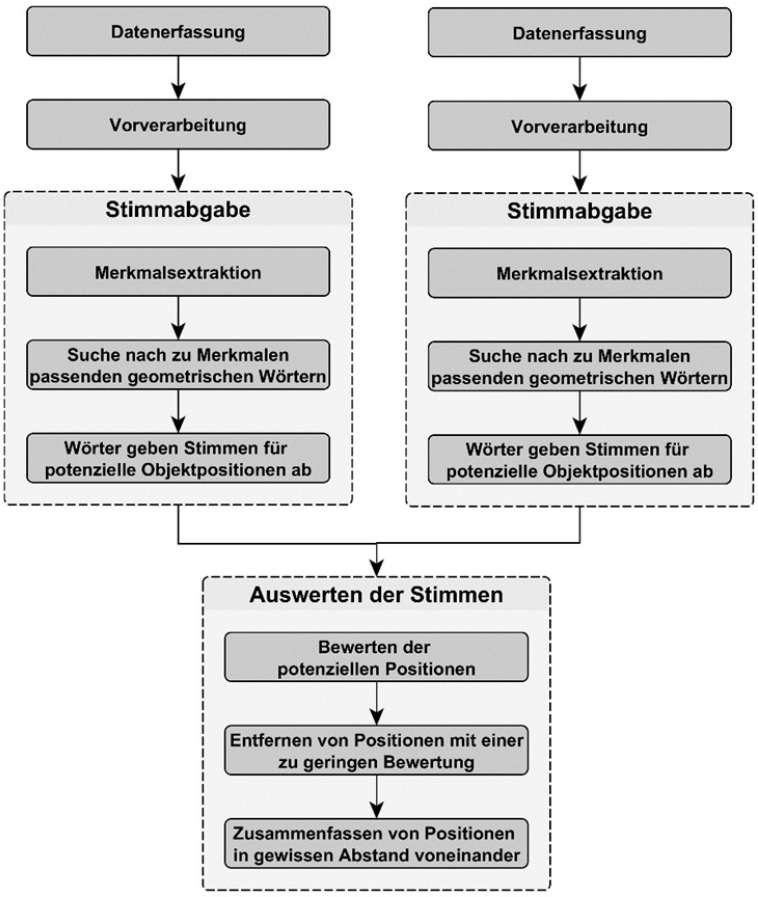


Abb. 2: Schematische Darstellung unseres Verfahrens zur Detektion von Personen in den Daten mehrerer LiDAR Sensoren. Hier dargestellt für zwei Sensoren.

bzw. synchronisierbar. Die Methode muss daher mit den auftretenden temporalen Verzerrungseffekten umgehen.

Der Ansatz stellt eine Weiterentwicklung eines Ansatzes dar, den wir bereits in der Vergangenheit vorgestellt haben (BORGMANN et al. 2017). Wir haben ihn modifiziert und ergänzt. Es handelt sich um ein stimmbasiertes Verfahren welches an 3D *Implicit Shape Models* (ISM) angelehnt ist. Es geht bereits davon aus, dass die abgegebenen Stimmen nicht komplett exakt sind, es also gewisse Unsicherheiten im Stimmraum des Verfahrens gibt. Dies erlaubt es uns eine Datenfusion zwischen mehreren Sensoren in diesem Raum vorzunehmen, ohne dass evtl. auftretende zeitlich bedingte Verzerrungseffekte zwischen den unterschiedlichen Sensoren das Ergebnis stark beeinträchtigen. Eine solche Fusion für die tatsächlichen Punktwolken durchzuführen birgt die Gefahr, dass die Extraktion von Punktmerkmalen durch die Verzerrungseffekte beeinträchtigt wird.

Unser Verfahren ist in Abb. 2 dargestellt und besteht aus drei Komponenten: Einer Vorverarbeitung zur Datenreduktion, der Abgabe von Stimmen und dem Auswerten der abgegebenen Stimmen. Während wir die ersten beiden Komponenten für jeden Sensor getrennt durchführen, erfolgt die dritte Komponente für alle Sensoren gemeinsam. Das Verfahren basiert auf einem Wörterbuch bestehend aus geometrischen Wörtern und verwendet Merkmalsvektoren, um sowohl die

zu verarbeitenden Daten als auch die Wörter im Wörterbuch zu beschreiben. Im Folgendem werden wichtige Komponenten unseres Verfahrens näher beschrieben.

3.1 Merkmalsextraktion

Wir verwenden eine Merkmalsextraktion um Merkmale zu erhalten, welche die lokale Nachbarschaft eines Punktes der gerade verarbeiteten Daten beschreiben. Die Annahme hierbei ist, dass die jeweilige lokale Nachbarschaft Rückschlüsse auf das Objekt zulässt, zu dem der Punkt gehört und es so erlaubt dieses zu klassifizieren bzw. zu detektieren. Für 3D-ISM Verfahren werden im Hinblick auf die Merkmalsextraktion zwei Strategien verfolgt. Eine ist es, komplexere Merkmale für eine kleine Gruppe von wohlausgewählten Punkten zu bestimmen (KNOPP et al. 2010). Die andere Strategie ist es, „weniger leistungsfähige“ Merkmale für eine größere Anzahl zufällig ausgewählter (VELIZHEV et al. 2012) oder gar aller Punkte zu ermitteln. Diese Variante ist robuster gegen verschiedene Störeffekte und Verdeckungen. In unserem Verfahren verwenden wir die zweite Strategie und extrahieren *Fast Point Feature Histograms* (FPFH) (RUSU et al. 2009) für alle Punkte die wir verarbeiten. Wir erhoffen uns von den FPFH Merkmalen, dass diese auch bei verhältnismäßig geringen Datendichten gute Ergebnisse liefern.

3.2 Wörterbuch

Das Wörterbuch enthält geometrische Wörter, welche die Stimmen abgeben die zur Detektion von Objekten verwendet werden. Es ist das Ergebnis des Trainingsprozesses unserer Methode. In Abb. 3 wird die Struktur des Wörterbuchs dargestellt. Es besteht aus mehreren Wörtern, die jeweils durch einen Merkmalsvektor beschrieben sind und mehrere Stimmen abgeben können. Jede Stimme wiederum stimmt für das Vorhandensein eines Objektes einer bestimmten Klasse an einer bestimmten Position (relativ zum Wort) ab. Als Position ist hierbei der Mittelpunkt des Objektes definiert. Jede Stimme hat außerdem ein Stimmgewicht, welches im Trainingsprozess ermittelt wird.

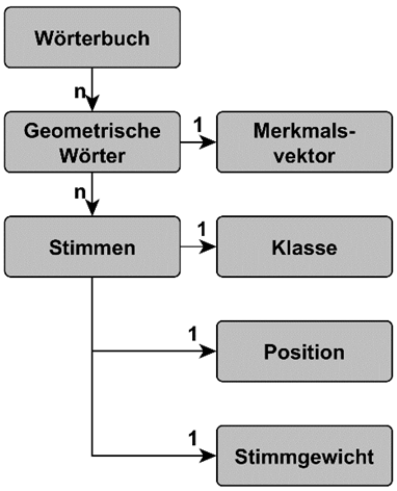


Abb. 3: Struktur des verwendeten Wörterbuchs

Das Training unserer Methode erfolgt in zwei Schritten und erhält als Eingabe Trainingsdaten, bei denen es sich um vorklassifizierte Punktwolkensegmente handelt, die jeweils ein Objekt bzw.

eine Person umfassen. Im Fokus dieses Beitrags haben wir Trainingsdaten der Klassen „Person“ und „keine Person“ verwendet. Das Training mit solchen Negativbeispielen hat sich als vorteilhaft gegenüber einem reinen Training mit Positivbeispielen herausgestellt, da es hierdurch möglich wird, personentypische Merkmale von solchen zu unterscheiden, die bei unterschiedlichsten Objektklassen auftreten. Im ersten Schritt des Trainings erfolgt für jeden Punkt jedes Trainingsdatensatzes eine Merkmalsextraktion. Die dabei ermittelten Merkmale sowie die relative Position des jeweiligen Punktes zum Objektmittelpunkt des Trainingsdatensatzes werden dann genutzt um neue Wörter mit jeweils einer Stimme zu initialisieren.
Nachdem alle Trainingsdaten verarbeitet wurden erfolgt der zweite Schritt des Trainings. Hierbei werden im Merkmalsraum ähnliche Wörter mithilfe eines *k*-means Clustering zusammengefasst und der Umfang des Wörterbuchs so reduziert. Bei diesem Prozess erben die zusammengefassten Wörter zunächst alle Stimmen der ursprünglichen Wörter. Anschließend werden für jedes Wort und jede Klasse getrennt voneinander ähnliche Stimmen, also Stimmen bei denen die Position ähnlich ist, ebenfalls zusammengefasst. Zuletzt wird das Stimmgewicht jeder Stimme ermittelt, wofür folgende Formel Anwendung findet:

$$G_S(w_i) = \frac{1}{N_S(w_i)}$$

$G_S(w_i)$: Gewicht jeder Stimme vom Wort w_i
$N_S(w_i)$: Anzahl der Stimmen des Wortes w_i

Hierdurch wird das Gesamtgewicht der Stimmen eines Wortes normiert. Außerdem bevorzugen wir so Stimmen von aussagekräftigen Wörtern (Wörter, die nur wenige unterschiedliche Stimmen abgeben) gegenüber denen von weniger aussagekräftigen Wörtern (Wörter, die viele verschiedene und widersprüchliche Stimmen abgeben).

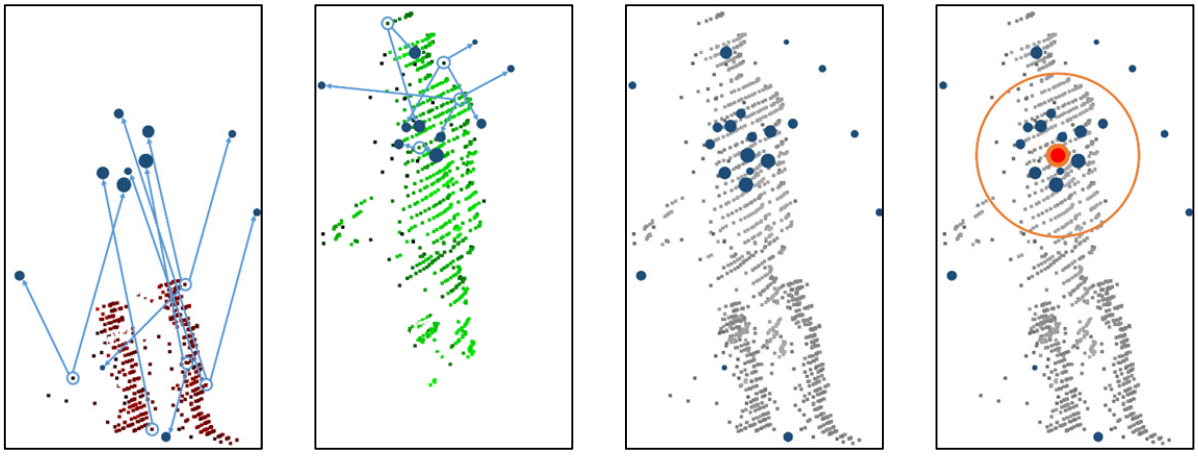
3.3 Vorverarbeitung

Die Vorverarbeitung stellt den ersten Schritt der Datenverarbeitung unserer Methode dar und dient dazu, das Laufzeitverhalten der Methode durch eine Datenreduktion zu verbessern. In der Vergangenheit bestand unsere Vorverarbeitung aus mehreren Schritten und umfasste eine Bodenextraktion, eine Segmentierung und ein Ausfiltern von Segmenten, die aufgrund einfacher geometrischer Merkmale keine Person sein konnten (BORGMANN et al. 2017). Ähnliche Vorverarbeitungsketten finden sich auch in vergleichbaren Methoden (VELIZHEV et al. 2012).
Obwohl dies meist gut funktioniert, kann es insbesondere bei Punktwolken aus Einzelscans eines MLS-Systems zu Problemen kommen. Solche Punktwolken weisen eine verhältnismäßig geringe Datendichte und viele Verdeckungen auf. Dies erschwert die korrekte Segmentierung der Daten, sodass es vermehrt zu einer Über- oder Untersegmentierung kommt. Solche Segmentierungsfehler führen nachfolgend zu Schwierigkeiten beim Filtern. Insbesondere im Falle einer Übersegmentierung wird zudem die eigentliche Detektion erschwert, da der Detektor ebenfalls auf den Punktwolkensegmenten arbeitet. Im Falle einer zu starken Segmentierung einer Person sind dann im einzelnem Segment nicht mehr genügend Informationen für eine erfolgreiche Detektion vorhanden.

Im Rahmen dieser Arbeit haben wir daher die Vorverarbeitung auf eine Bodenextraktion beschränkt. Durch diese lässt sich für unsere Anwendungsfälle bereits eine Datenreduktion von ca. 45% erreichen (BORGMANN et al. 2017). Damit der Detektor auch bei unsegmentierten Daten korrekt funktioniert sind einige Anpassungen nötig, welche im Abschnitt 3.5 erläutert werden.

3.4 Stimmabgabe

Zur Abgabe von Stimmen erfolgt zunächst die Merkmalsextraktion. Die ermittelten Merkmale werden dann verwendet um im Wörterbuch das zum jeweiligen Merkmal am besten passende Wort zu suchen. Für diese Suche im Wörterbuch greifen wir auf einen Verfahren zur schnellen approximativen Nachbarschaftssuche zurück welches von MUJA & LOWE (2009) vorgestellt wurde. Dieses wenden wir auf den Merkmalsraum der Wörter des Wörterbuchs an. Die Stimmen der gefundenen Wörter werden dann basierend auf der Stimmposition und der Position des Wortes im 3D-Raum in diesen projiziert.
In Abb. 4a) und b) wird die Stimmabgabe beispielhaft für die Verarbeitung von Messdaten zweier Sensoren dargestellt. Es ist zu beachten, dass anders als in der Darstellung in unserer Methode tatsächlich für alle Punkte Stimmen abgegeben werden. Wie bereits erläutert erfolgt die Stimmabgabe für beide Sensoren getrennt. Dies verhindert gegenüber der direkten Fusion der Punktwolken Probleme bei der Merkmalsextraktion, die durch zeitliche Verzerrungseffekte verursacht werden. Der Nachteil hierbei ist, dass wir bei der Merkmalsextraktion nicht von einer potenziell höheren Datendichte im Überlappungsbereich profitieren können. Wir gehen jedoch davon aus, dass die Vorteile hier die Nachteile überwiegen.



a) Beispiele von Stimmen für Daten von Sensor 1

b) Beispiele von Stimmen für Daten von Sensor 2

c) Fusion der Stimmen beider Sensoren

d) Bewerten einer Position basierend auf Stimmen der Umgebung

Abb. 4: Illustration unseres Detektors für die Verarbeitung von Daten zweier LiDAR-Sensoren

3.5 Bewerten der Stimmen

Nach der Stimmabgabe erfolgt die weitere Verarbeitung aller abgegebenen Stimmen aller Sensoren gemeinsam. Der Zustand zu Beginn dieses Schritts ist beispielhaft in Abb. 4c) dargestellt. Wir versuchen nun zu entscheiden, bei welchen der abgegebenen Stimmen es sich tatsächlich um

eine Person handelt. Hierfür betrachten wir jede Stimme als eine potenzielle Position eines Objektes und nehmen eine Bewertung von diesen vor, und zwar basierend auf ihrem eigenen Gewicht und dem Gewicht von anderen potenziellen Positionen derselben Klasse in der Nachbarschaft. Als Nachbarschaft definieren wir dabei einen gewissen Radius um die gerade bewertete Position. Die Bewertung ist in Abb. 4d) illustriert und basiert auf der Gaußschen Normalverteilung. Sie addiert einen Teil des Gewichts benachbarter Positionen zu dem Gewicht der gerade bewerteten Position. Die Annahme hierbei ist, dass ein großes Stimmgewicht auf das Gebiet um die tatsächliche Position eines Objektes fallen wird. Wir verwenden für die Bewertung folgende Formel:

$$B_p = \sum_{k \in K} W_k \times RW_{\text{norm}} \times e^{-\frac{D_{pk}^2}{2 \times \sigma^2}}$$

B_p : Bewertetes Gewicht von Position p
K : Alle potenziellen Positionen mit derselben Klasse wie Position p im Bewertungsradius
W_k : Nicht bewertetes Gewicht der Position k
RW_{norm} : Normierungsfaktor für den Bewertungsradius R
D_{pk} : Euklidische Distanz zwischen Positionen p und k
σ : Breite der Normalverteilung.

Die Formel verfügt über einen Normierungsfaktor, welcher es uns erlaubt mit unterschiedlichen Datendichten umzugehen und dabei trotzdem später einen einfachen Schwellwert als Detektionsentscheidung anwenden zu können. Anders als in unseren früheren Arbeiten basiert der Normierungsfaktor in dieser Arbeit auf der Anzahl an potenziellen Positionen im Bewertungsradius. Früher basierte er auf der Anzahl an potenziellen Positionen in dem gerade verarbeiteten Segment der Daten. Dieser Änderung wurde notwendig um auf eine Segmentierung verzichten und um mit Stimmen aus unterschiedlichen Punktwolken umgehen zu können. Diese ergeben sich aus der parallelen Verarbeitung für mehrere Sensoren. Der Normierungsfaktor wird wie folgt berechnet:

$$RW_{\text{norm}} = \frac{1}{N(P_R)}$$

RW_{norm} : Normierungsfaktor für den Bewertungsradius
$N(P_R)$: Anzahl Positionen im Bewertungsradius R

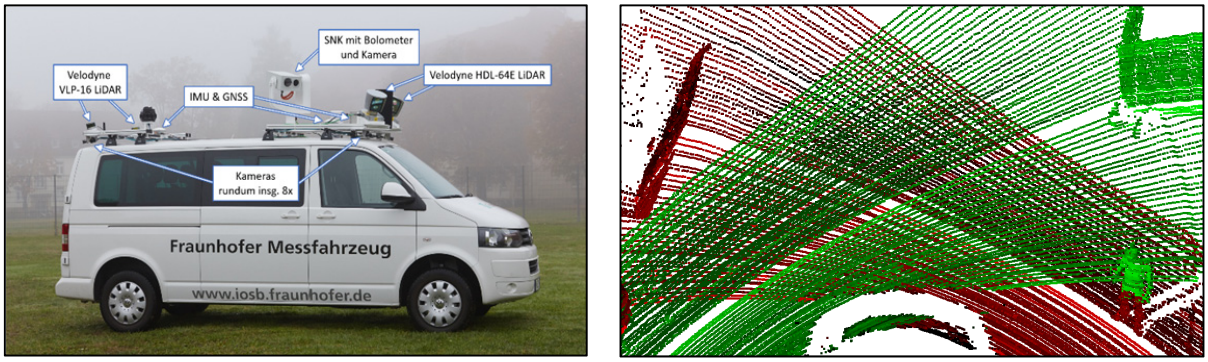
Diese Variante der Normierung ist anfällig in Regionen, in denen es nur sehr wenige Stimmen in der Nachbarschaft gibt. Daher gibt es einen Schwellwert für die Mindestanzahl an Nachbarn, die eine potenzielle Position in ihrem Bewertungsradius haben muss, um weiter verarbeitet zu werden.

4 Experimentelle Untersuchung

In diesem Abschnitt beschreiben wir die von uns durchgeführten Experimente. Der Fokus liegt dabei auf der Verarbeitung von Daten mehrerer LiDAR-Sensoren. Wir beschreiben zunächst den Aufbau und Ablauf der durchgeführten Experimente. Danach stellen wir deren Ergebnisse vor und analysieren diese.

4.1 Ablauf

Für unsere Experimente benötigten wir 3D-Messdaten von Personen, welche von mehreren LiDAR-Sensoren erfasst werden. Derartige Daten haben wir mit unserem am Institut vorhandenen Messfahrzeug *MODISSA* erstellt. Dieses ist in Abb. 5a) dargestellt und verfügt unter anderem über zwei Velodyne HDL-64E Laserscanner, welche wir für unsere Experimente verwendet haben. Bei den Sensoren handelt es sich um scannende LiDAR-Sensoren mit einem rotierenden Aufnahmekopf. Sie sind jeweils in der Lage 1,3 Millionen Messungen pro Sekunde durchzuführen, welche sich auf 64 Scanzeilen verteilen. Durch den rotierenden Kopf ist der horizontale Sichtbereich der Sensoren 360°, der vertikale ist 26,9°. Für die Experimente wurden die LiDAR-Sensoren mit 10 Umdrehungen pro Sekunde betrieben und für jeden Sensor alle 0,1 Sekunden eine Punktwolke generiert. Jede Punktwolke umfasst also ca. 130.000 Messungen. Die tatsächliche Anzahl der Punkte ist jedoch geringer, da Teile der Messungen zu keinem Ergebnis führen. Dies ist z.B. der Fall, wenn der Laserpuls auf kein Objekt trifft („Messungen in den Himmel“). Durch das am Fahrzeug vorhandene GNSS/IMU-System ist eine direkte Georeferenzierung der aufgenommenen Daten möglich. Diese erlaubt es die Eigenbewegungen des Fahrzeugs zu kompensieren. Die verbleibenden Verzerrungen in den Daten sind daher auf die Bewegungen der aufgenommenen Objekte und Personen zurückzuführen.



a) Das verwendete Messfahrzeug MODISSA b) Beispiel für die verwendeten Daten

Abb. 5: Experimentalsystem und Ablauf der Experimente

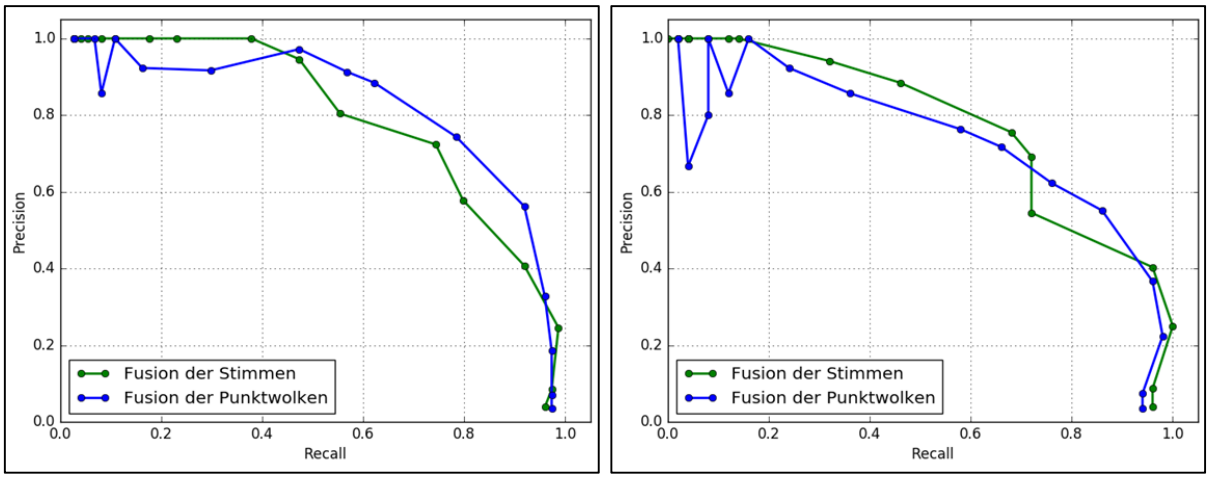
Für die Experimente wurde zunächst ein Wörterbuch trainiert. Für dieses Training wurden 993 per Hand annotierte Punktwolken verwendet. Die Annotationen umfassen dabei sowohl Positivbeispiele der Klasse „Person“ als auch Negativbeispiele, bei denen es sich um keine Personen handelt. Die für das Training verwendeten Daten stammen aus verschiedenen Messkampagnen bzw. gestellten Szenen.

Für die Experimente wurde zwei Sequenzen aufeinanderfolgender Punktwolken beider Sensoren verwendet. Die erste Sequenz umfasst 74 Punktwolken pro Sensor, insgesamt also 148. Die zweite Sequenz umfasst 50 Punktwolken pro Sensor. Beide Sequenzen wurden ebenfalls annotiert um eine Grundwahrheit zu generieren. In Abb. 5b) wird ein Beispiel dieser Daten gezeigt, wobei die einzelnen Sensoren farblich hervorgehoben sind. Die Sequenzen unterscheiden sich darin, wie sehr sich die aufgenommene Person im Überschneidungsbereich der beiden Sensoren bewegte. Bei Sequenz 1 gibt es nur wenige solcher Bewegungen, bei Sequenz 2 jedoch mehr. Es

kommt daher in dieser öfter zu den beschriebenen Verzerrungseffekten. Außerdem gibt es in der Sequenz 2 einen längeren Zeitraum, in dem die Person trotz gemeinsamer Verwendung zweier Sensoren nur teilweise zu sehen ist.
Um unser Verfahren der Datenfusion zwischen den Sensoren im Stimmraum mit einem einfachen Ansatz der Fusion der Punktwolken zu vergleichen, haben wir die Testdaten mit unserer Methode in beiden möglichen Varianten mit unterschiedlichen Detektionsschwellwerten verarbeitet. Es wurde ausgewertet, ob Detektionen in den verarbeiteten Daten mit Detektionen in der Grundwahrheit übereinstimmen. Dies wurde dann genutzt um Richtig-Positiv, Falsch-Positiv und Falsch-Negativ zu ermitteln. Diese Kennzahlen nutzen wir um *Precision* und *Recall* zu ermitteln und um damit die Leistung der Methode zu quantifizieren. Precision gibt den Anteil der korrekten Detektionen bezogen auf alle Detektionen an. Recall den Anteil der detektierten Personen bezogen auf die in der Grundwahrheit tatsächlich vorhandenen Personen.

4.2 Ergebnisse

In Abb. 6 werden die Ergebnisse unserer Experimente als Precision-Recall-Kurve dargestellt. a) stellt dabei die Ergebnisse für die Verarbeitung der ersten Testsequenz dar und b) für die zweite. Es wird deutlich, dass die Fusion im Stimmraum verglichen mit der Fusion der Punktwolken in der zweiten Sequenz bessere Leistungen liefert als in der ersten Sequenz. Dies ist mit dem größeren Anteil an Verzerrungseffekten zwischen den Sensoren in der zweiten Sequenz zu erklären. Treten solche Effekte kaum auf, profitiert die Fusion der Punktwolken davon, dass sie eine größere Datendichte für die Merkmalsextraktion verwenden kann. Umgekehrt hat sie jedoch mehr Fehler bei der Merkmalsextraktion, wenn solche Verzerrungseffekte häufiger vorkommen.

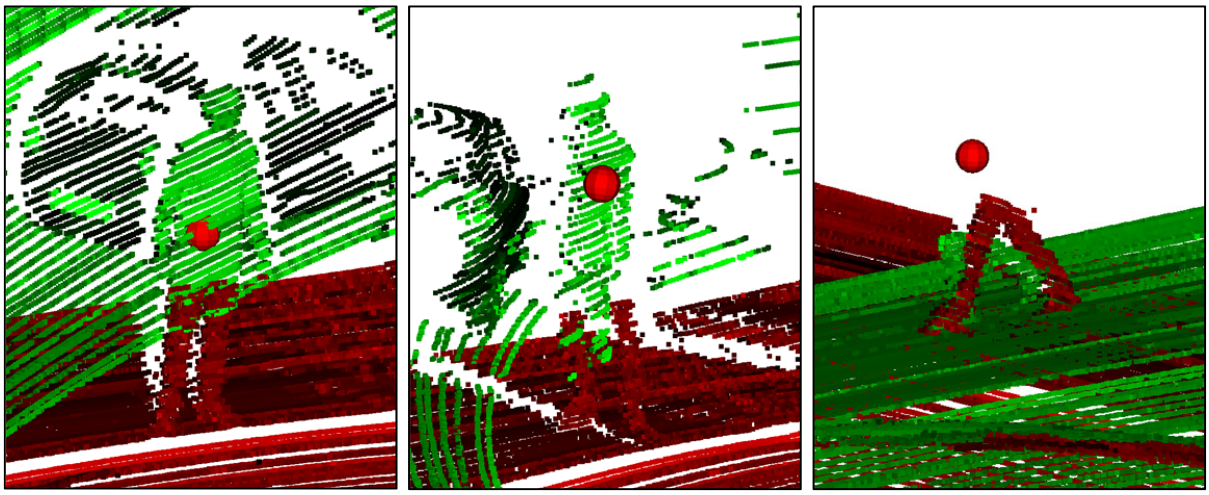


a) Precision-Recall Kurve für die erste Testsequenz b) Precision-Recall Kurve für die zweite Testsequenz

Abb. 6: Auswertung der Ergebnisse

Die Leistung ist insgesamt in der zweiten Sequenz geringer als in der ersten. Dies liegt am höheren Schwierigkeitsgrad der zweiten Sequenz. Sie hat nicht nur einen größeren Anteil Verzerrungseffekte, sondern auch Bereiche in denen die Person überhaupt nur teilweise zu sehen ist.

Abb. 7 zeigt Beispiele für die Ausgabe unserer Methode ohne Verzerrungen a) und mit Verzerrungen b). In Abb. 7c) ist zu sehen, wie unsere Methode auch noch eine korrekte Personenposition liefert, wenn der Mittelpunkt der Person gar nicht in den Daten enthalten ist.



a) Person ohne Verzerrungseffekte b) Person mit Verzerrungseffekten c) Person nur teilweise zu sehen

Abb. 7: Beispiele für die Ausgabe unserer Methode

5 Fazit & Ausblick

In diesem Beitrag haben wir unseren auf 3D-ISM basierten Ansatz auf eine Art und Weise erweitert, die ihn Robust gegenüber Verzerrungseffekte macht, wie sie auftreten, wenn bewegte Objekte von mehreren nicht komplett zeitlich synchronisierten LiDAR-Sensoren erfasst werden. Wir verwenden dafür eine Datenfusion im Stimmraum des ISM Verfahrens. Es wurde experimentell untersucht, dass diese Methode bei steigender Anzahl der beschriebenen Verzerrungseffekte gegenüber einer einfachen Fusion der Eingangsdaten bessere Ergebnisse liefert.
Die vorgestellte Methode ist für die Detektion von Personen bzw. mit entsprechendem Training auch anderer Objektklassen nicht auf eine vorherige Segmentierung der Daten angewiesen und unterscheidet sich darin von vielen anderen Methoden mit gleicher Zielsetzung. Sie ist dadurch robust gegen Segmentierungsfehler.
In der Zukunft planen wir den Ansatz der Datenfusion im Stimmraum weiter auszubauen und ihn auch für die Integration eines Trackingverfahrens und ggf. anderer Sensortypen zu nutzen. So kann die Ausgabe eines Trackingverfahrens ebenfalls als potenzielle Objekt Position mit einem gewissen Gewicht in den Stimmraum überführt werden. Die Ausgabe eines Kamerabasierten Detektors könnte als Strahl in diesen überführt werden.

6 Literaturverzeichnis

BEHLEY, J., STEINHAGE, V. & CREMERS, A.B., 2013: Laser-based segment classification using a mixture of bag-of-words. 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 4195-4200.

BORGMANN, B., HEBEL, M., ARENS, M. & STILLA, U., 2017: Detection of persons in MLS point clouds using implicit shape models. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W7, 203-210.

GARCIA-GARCIA, A., GOMEZ-DONOSO, F., GARCIA-RODRIGUEZ, J., ORTS-ESCOLANO, S., CAZORLA, M. & AZORIN-LOPEZ, J., 2016: PointNet: A 3D Convolutional Neural Network for real-time object class recognition. 2016 International Joint Conference on Neural Networks (IJCNN), 1578-1584.

KNOPP, J., PRASAD, M., WILLEMS, G., TIMOFTE, R. & VAN GOOL, L., 2010: Hough Transform and 3D SURF for Robust Three Dimensional Classification. Proceedings of the 11th European Conference on Computer Vision: Part VI, Springer, Heraklion, Crete, Greece, 589-602.

LEIBE, B., LEONARDIS, A. & SCHIELE, B., 2008: Robust Object Detection with Interleaved Categorization and Segmentation. International Journal of Computer Vision **77**(1), 259-289.

MATURANA, D. & SCHERER, S., 2015: VoxNet: A 3D Convolutional Neural Network for real-time object recognition. 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 922-928.

MUJA, M. & LOWE, D. G., 2009: Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. International Conference on Computer Vision Theory and Application VISSAPP'09, INSTICC Press, 331-340.

NAVARRO-SERMENT, L. E., MERTZ, C. & MARTIAL, H., 2010: Pedestrian Detection and Tracking Using Three-dimensional LADAR Data. The International Journal of Robotics Research **29**(12), 1516-1528.

PREMEBIDA, C., CARREIRA, J., BATISTA, J. & NUNES, U., 2014: Pedestrian detection combining RGB and dense LIDAR data. 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, 4112-4117.

RUSU, R.B., BLODOW, N., & BEETZ, M., 2009: Fast Point Feature Histograms (FPFH) for 3D Registration. Proceedings of the 2009 IEEE International Conference on Robotics and Automation, IEEE Press, Piscataway, NJ, USA, 1848-1853.

SHOTTON, J., FITZGIBBON, A., COOK, M., SHARP T., FINOCCHIO M., MOORE, R., KIPMAN, A. & BLAKE, A., 2011: Real-Time Human Pose Recognition in Parts from a Single Depth Image. CVPR 2011, IEEE.

SHOTTON, J., GIRSHICK, R., FITZGIBBON, A., SHARP, T., COOK, M., FINOCCHIO, M., MOORE, R. KOHLI, P., CRIMINISI, A., KIPMAN, A. & BLAKE, A., 2013: Efficient Human Pose Estimation from Single Depth Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, **35**(12), 2821-2840.

SOCHER, R., HUVAL, B., BATH, B., MANNING, C.D. & NG, A.Y., 2012: Convolutional-recursive deep learning for 3d object classification. Advances in Neural Information Processing Systems, 656-664.

VELIZHEV, A., SHAPOVALOV, R. & SCHINDLER, K., 2012: Implicit shape models for object detection in 3D point clouds. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, I-3, 179-184.

High Definition Mapping Using LiDAR Traced Trajectories

STEFFEN BUSCH[1], JANNIK QUEHL[2] & CLAUS BRENNER[1]

Abstract: In this paper, we automatically reconstruct a high definition (HD), lane accurate model through trajectory analysis. The fundamental idea behind our approach is that in the future, many, if not all, of the vehicles will carry sensors, which can be used to keep HD maps up to date. In order to explore this idea, we used a static 3D laser scanner placed at an intersection. In the continuously generated scans, we tracked objects and obtained their trajectories. From these, a map was derived based on trajectory clustering and least squares adjustment. We evaluate the results by comparing them to a ground truth map labeled manually from a mobile mapping LiDAR point cloud.

1 Introduction

Today there are more and more use cases for digital maps. Car navigation systems, land use, vacation trip planning, market analysis and many other web-based services are widespread cases with different requirements with regard to reliability, precision and accuracy. Especially autonomous vehicles, virtual reality (VR) and smart cities will increase the requirement for future maps dramatically. Maps will become more detailed to provide features for automatic localization and behavior recognition. The high dynamics of the environment makes frequent map updates necessary. Features may appear, vanish or change rapidly because of road construction work, accidents, or other unpredictable changes. These changes need to be mapped as soon as they appear to update the virtual representation of the scene. For some applications, like VR or augmented reality for decision making support, there are only moderate requirements with respect to the accuracy and timeliness of the contained information. However, for others, like autonomous vehicles, the map must

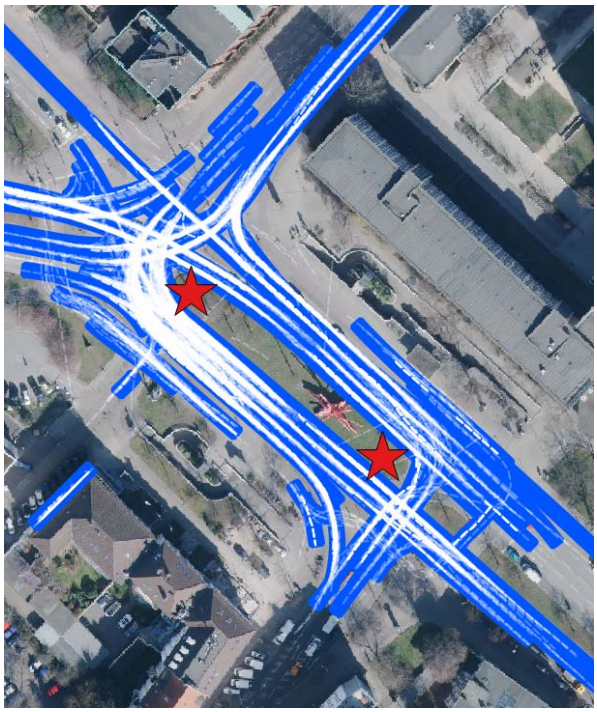


Fig. 1: Data and results of our experiment: trajectories (white), lane segments (blue), scanner positions (red). Ortho image: (HANNOVER 2017)

[1] Leibniz Universität Hannover, Institute of Cartography and Geoinformatics, Appelstraße 9a, D-30167 Hannover, E-Mail: [Steffen.Busch, Claus.Brenner]@ikg.uni-hannover.de
[2] Karlsruhe Institute of Technology, Institute of Measurement and Control Systems, Engler-Bunte-Ring 21, D-76131 Karlsruhe, E-Mail: Jannik.Quehl@kit.edu

be an accurate, up-to-date representation of the environment, which is constantly compared to the perception of the "smart machines" which use it.
We propose to generate and maintain a representation of the environment in the form of a dynamic high definition map, in a crowd sensing manner through sensor information gathered by the daily traffic. For our analysis, we use trajectories, because highly automated driving will at least provide information about the movement of other traffic participants to assist the drivers. Moreover, trajectories are easily collectable in a crowd-sensing manner from many different sources. In this approach, we focused on a lane accurate map for autonomous vehicles (Fig. 1).
We present our first experiment to derive lanes with an accuracy of a few centimeters by analyzing vehicle trajectories. Since the acquisition of a large number of vehicle trajectories would require us to fit a fleet of vehicles with very accurate positioning sensors, which is beyond our scope, we used a 3D laser scanner instead to observe a specific junction. In more detail, we placed a Velodyne HDL-64 S2 3D laser scanner at a complex junction to track vehicles with a scan frame frequency of 10 Hz. We focused on trajectory analysis because trajectories provide more detailed and dynamic information about the actual traffic behavior than static infrastructure. The resulting trajectories were clustered in order to find the lanes. The lanes were then approximated using a cubic polynomial, fitted by least squares adjustment. We accomplished the geo-referencing by aligning poles, automatically detected in the laser scans, to a reference pole map obtained using a total station. The geo-referenced HD map was evaluated against a manually generated lane model, based on Riegl VMX-250 mobile mapping system measurements.
The paper is structured as follows: In section 2, we give an overview over actual mapping procedures and trajectory based road network construction in detail. In section 3, we will present our approach: firstly, we present our pole geo-referencing procedure. Secondly, we will introduce our clustering and lane alignment procedure. In section 4, we present our dataset. Section 5 evaluates our result by comparing the lane model to a manually generated model. Finally, we will discuss our results and give an outlook in section 6.

2 Related Work

This section gives an overview over the state of the art of mapping approaches. The approaches are divided into *mobile mapping* (MM) and *remote sensing* (RS), which are considered to be associated with a relatively large acquisition cost, versus crowd sensing, which is considered to be a low cost technique. The overview illustrates the need for crowd sensing to achieve high frequent map updates and define the accuracy for high definition maps. The section concludes with the overview of crowd sensed trajectory analysis approaches.
There are plenty of different mapping approaches for different kinds of maps. We focus on the state of the art mapping procedures for commercial navigation systems. Today, most maps are acquired using measurement campaigns, where mobile mapping systems or airplanes gather the required data (TOTH & JÓŹKÓW 2016). Thus, the spectrum of data useful for road network reconstruction reaches from aerial imagery, via ego camera images, to high-resolution LiDAR point clouds. On the one hand, RS is more complex but also more efficient for measuring large areas in contrast to the high effort for frequent, area-wide measurements via a fleet of mobile mapping systems. On the other hand, even if UAV holds the key to simplify RS (NEITZEL & KLONOWSKI

2011; TOTH & JÓŹKÓW 2016), mobile mapping vehicles generate much more detailed information. MM reaches a resolution down to a few millimeters (PUENTE et al. 2013; GUAN et al. 2016) in contrast to meter or decimeter resolution (TOTH & JÓŹKÓW 2016) for airborne measurements. With respect to geo-referencing, all mobile mapping systems depend on their localization accuracy. Geo-referencing is distinguishable into direct (GNSS/IMU based), indirect (control point based) and a combination – the integrated sensor orientation (ISO) approaches. The direct approaches vary between 10 and 0.02 meters accuracy, depending on the sensor quality and correction approach (TOTH & JÓŹKÓW, 2016). Although ISO methods are more robust against occlusion, we decided to use full indirect geo-referencing in our first experiment by ensuring the visibility of control points from different static measurement positions. The presented approaches have high data acquisition costs in common and the large amounts of acquired sensor data still need a large amount of manual work in post processing to transform them into maps like the widely used Google or Here maps. An alternative is the use of crowd sensing, discussed in the following.

One of the most famous examples for crowd sensing in the geographic area is Open Street Map (OSM) (M. HAKLAY & WEBER 2008). The members of the OSM community are both users and volunteers. GOODCHILD (2007) characterizes the user responsibility for completeness and quality of geographic information with the term "volunteered geographic information" (VGI). HAKLAY et al. (2013) analyze the commercial use of crowd sourcing and distinguish between passive and active community input. OSM uses the active community input by allowing its community members to change every information, for the price of temporally inconsistent data. In contrast to OSM, commercial providers usually only use data of their users to take care of the data accuracy at every moment. However, widely known providers use their community knowledge to benefit from the self-healing effect of active input. Google produces maps by allowing users to upload their "expert knowledge", for example, filling gaps or correcting mapping errors. Some commercial providers use a hybrid approach by checking the active community information for consistence before they integrate the knowledge into their maps. Other companies work with passive community information to avoid the complex consistency-check of user expert knowledge. They cherish the users as sensor probes and update their database with traditional tools. We will follow this approach of passive community information by combining sensor data from different independent poses and by analyzing trajectories, which could easily be collected by the crowd (TANG et al. 2016).

Trajectory data generated by crowd sensing have been used in the past to create digital navigation maps (ROETH et al. 2017; RUHHAMMER et al. 2017; DURAN et al. 2016). The analysis methods can be broadly classified into three different approaches: *intersection linking*, *incremental track insertion* and *point clustering* (AHMED et al. 2015). *Intersection linking* first finds a set of intersections by analysis of movement characteristics or point density and afterwards connects these nodes by interpolation. These approaches work best for 'Manhattan style' networks. The incremental track insertion tries to match tracks successively to a map, starting with the first track and an empty map. The map becomes more detailed by adding nodes and edges based on the tracks. The point clustering approaches can be distinguished into kernel density and intersection clustering based approaches. Much research was done in the field of network graph generation, whereas relatively few approaches address the generation of highly detailed information,

e. g. about the lane curvature, width or stop line position. Autonomous vehicles and even advanced driver assistance systems rely on highly detailed map information. BENDER et al. (2014) introduce so-called lanelets as map representation for autonomous vehicles. All three approaches, *intersection linking, incremental track insertion* and *point clustering,* have the potential to provide high definition maps, if the input data is detailed enough. However, the strict requirements of the navigation data standard (NDS) with respect to accuracy and detail make it necessary to use crowd sensing data; JUNG et al. (2016) use dash cam images for an efficient lane detection. The acquisition of lane-accurate maps from mobile mapping data approaches cannot satisfy the up-to-date requirement for the NDS, *active crowd-sensing* with an error of up to 1.5 m (LANDSIEDEL AND WOLLHERR, 2017) cannot maintain the accuracy to ensure a save navigation and *hybrid* solutions are too complex and too expensive for area-wide mapping. Several approaches have been proposed to automatically analyze passive crowd-sensing data, like ego-camera images (JUNG et al., 2016; BENDER et al., 2014) and trajectories (ROETH et al., 2017; RUHHAMMER et al., 2017; DURAN et al.,2016) to generate a high definition lane accurate maps. In contrast to ROETH et al. (2017) and RUHHAMMER et al. (2017), JUNG et al. (2016) and RABE et al. (2016) rely on the identification of road markings in sensor data. Nevertheless, the accuracy of trajectory analysis as well as road marking analysis depend to the quality of geo-referencing, even if the alignment methods can improve the accuracy. Our approach uses trajectory data from two different viewpoints, with the advantage that trajectories from one source guarantee a relative precision by LiDAR tracking beyond other geo-referencing tracks (RAZA AND ZHONG 2017; TREIBER & KESTING 2013; MORDECHAI HAKLAY & WEBER 2008; HU et al. 2017; VIVACQUA et al. 2017; KNOOP et al. 2017).

3 Our HD Mapping Approach

Our fully automatic trajectory high definition mapping combines 3D laser scanner measurements from different positions by geo-referencing. Therefore, we track dynamic objects in each scan and map them to their global position in the UTM coordinate frame. After we cluster the tracks by DB-SCAN clustering, we generate lanes by least square alignment; see Fig. 2.

3.1 Geo-referencing

We use a map of poles in the city of Hanover (SCHLICHTING & BRENNER 2014) and a coarse localization to restrict the area for an automatic geo-referencing by an iterative closest point (ICP) algorithm. We determine the position of the poles in the scans by transforming the scan data into a depth image. Therefore, we define a zero degree azimuthal angle (φ) and sort the measurement points into a grid with the resolution of approximately 0.17° for the scanner rotation. The vertical resolution corresponds the 64 vertical rays of the scanner, which is why the resolution of the grid rows is irregular. Finally, we fill in the distance for each measurement into the grid and analyze the resulting depth image. The analysis starts with finding depth jumps row-wise to define azimuthal angles of pole candidates. Next, candidates of different rows are connected by region growing. In the end the candidates are filtered by geometric features; objects wider than 0.5 m and shorter than 2 m are discarded. The resulting poles are matched to the pole map by using the ICP algorithm with various initial setups.

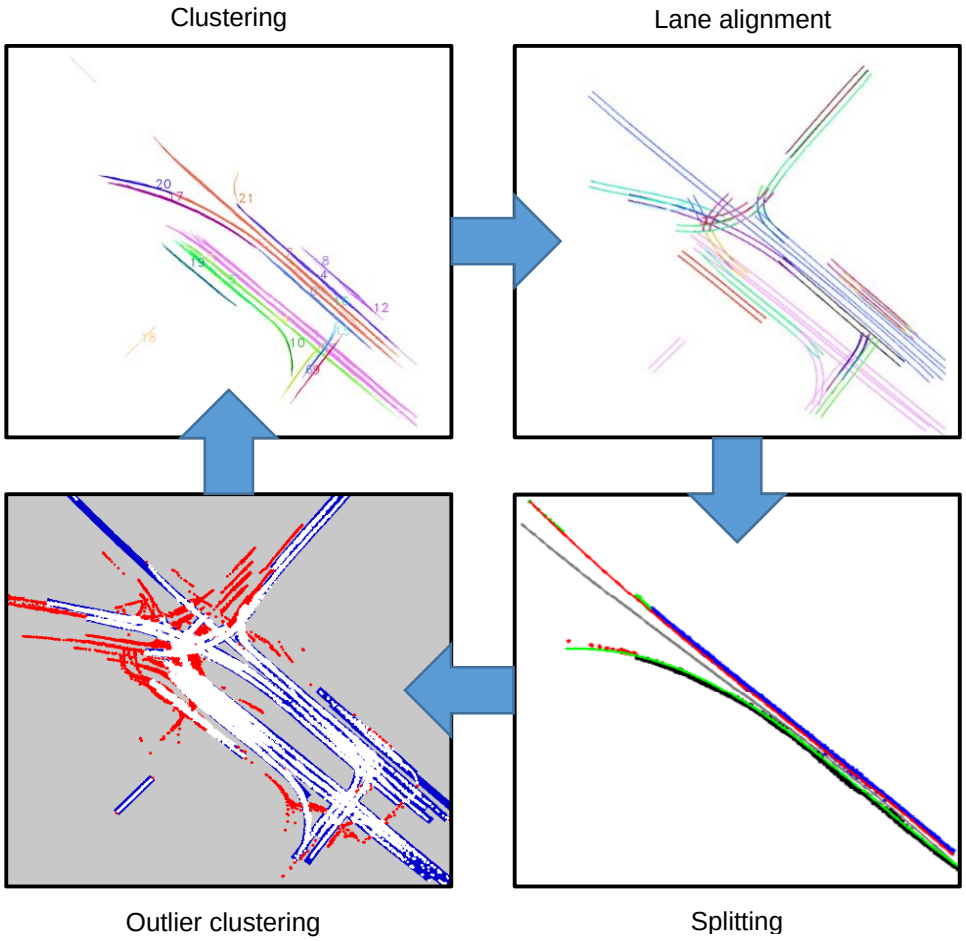


Fig. 2: Main steps of our mapping approach

3.2 Tracking

First, we detected objects in the point clouds from the 3D laser scanner; by removing the ground plane and clustering the remaining points. We estimated the plane by using RANSAC (FISCHLER & BOLLES 1981). The objects were identified by using an Euclidean cluster extraction provided by the point cloud library (RUSU & COUSINS, 2011). The cluster centers of gravity of the convex hulls determined the position of the object. We generated the trajectories (see Fig. 3) by using a *Unscented Kalman Filter* with a CYRA (HOUENOU et al. 2013; SCHUBERT et al. 2008) motion model and associated objects of successive time steps with the nearest neighbor

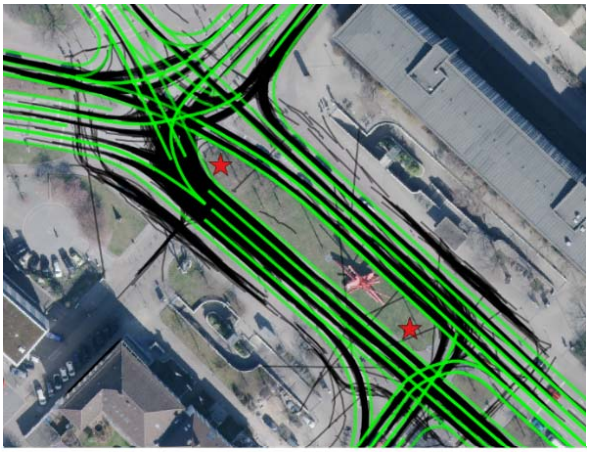


Fig. 3: Trajectories (black) and reference lane markings (green). Ortho image: (HANNOVER 2017)

approach.

3.3 Lane Clustering

We assume that each track represents part of a lane and can be approximated by a Hermite Cubic Spline (*HCS*) (CATMULL & ROM, 1974). This assumption allows for an efficient distance metric and clustering. Our approach uses two successive clustering steps for road segments that differ a lot in density and frequency, e.g. straight lanes and curves. First, we cluster the clear, less curvy and frequently used road segments with a cluster distance of 0.3 m. Secondly, we cluster the outlier trajectories, which do not fit the previously generated lane model with a cluster distance of 0.5 m (see Fig. 2**Fehler! Verweisquelle konnte nicht gefunden werden.**, *outlier clustering*). We calculate the cluster distance for the DB-SCAN clustering by transforming each trajectory from the world frame into a *HCS* frame by translating the origin into the middle of the trajectory and rotate the axis by the angle of a vector v. This vector v is calculated by the difference between the end and start point of the trajectory. This transformation is used to transform the start and end points as well as their gradient of trajectories to determine the *HCSs.* The cluster distance is calculated from 5 points from one *HCS* by transforming the world coordinates into the frame of the other *HCS* replacing the y coordinate by its function value and transforming the point back into the world to calculate the Euclidean distance. In case one point is outside the definition area of the *HCS* we extrapolate in a linear manner to calculate the new y value.

3.4 Lane Alignment

The lane segment geometry is estimated using a cubic polynomial by a least squares adjustment. First, the start and end point for the trajectory cluster are determined by analyzing the cluster direction via principal component analysis of all trajectory points. Secondly, the temporary direction vector is used to define the first start and last end point. Finally, these start and end points define vector *v* (1), which determines the trajectory cluster center *o* (2).

$$v_i = e_i - s_i \tag{1}$$

$$o = s_i + \frac{v_i}{2} \tag{2}$$

With *v, u* (3) *and o* each trajectory point of one cluster is transformed into the local ("polynomial") frame by (4).

$$u_i = \begin{pmatrix} -v_{2,i} \\ v_{1_i} \end{pmatrix} \tag{3}$$

$$x' = \frac{v^T}{\|v^T\|}(p - o)$$

$$y' = \frac{u^T}{\|u^T\|}(p - o) \tag{4}$$

The least squares adjustment (5) estimates the lane polynomial for a cluster.

$$\hat{s} = (A^T A)^{-1} A^T l \tag{5}$$

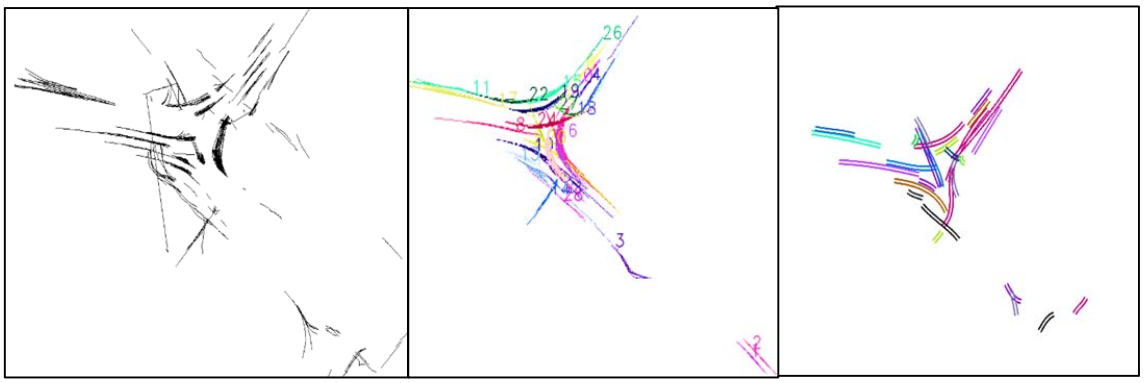
With:

$$\hat{s} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}, A = \begin{pmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_i^3 & x_i^2 & x_i & 1 \end{pmatrix} \text{ and } l = \begin{pmatrix} y_1 \\ \vdots \\ y_i \end{pmatrix}$$

Clustering errors are detected by filtering each aligned lane segment with an RMSE () above 0.5 m.

$$rmse = \sqrt{(Ax - l)^T (Ax - l)}/i \tag{6}$$

The clustered trajectory points of incorrect lane segments are split into points below and above or points left and right. Thus, four lanes are aligned by allocating each point to two of the four sets left, right, above and below if the x value belongs to the first or second part or the y value is above or below the function value. The pair (left/right or above/below) with the lower RMSE is replacing the splitting segment. Fig. 2, s*plitting,* illustrates the splitting step for one of the clusters. This procedure is repeated until no more cluster error is detected. Our splitting procedure can generate some duplicates of the same lane segments in some cases, for example, if more than one split is required to distinguish neighboring lanes. Thus, four segments describe two real world lanes. All generated duplicates are identified by an overlap of over 50% and direction difference below 20° and get merged by estimating the merged lane through the union of the pseudo lane trajectory points. Next, we assume a general lane width of 3 m and look for outlier trajectories from our preliminary road lane model. Therefore, we split each trajectory and the entrance and leaving points of our lanes and filter the trajectory snippets outside our model, see Fig. 4, *outlier clustering*.



Outlier trajectories Outlier cluster Outlier lane segments

Fig. 4: Outlier trajectory clustering

Fig. 4 shows the outlier trajectory clustering with less strictly set cluster parameters, cluster distance of 0.5 m in contrast to 0.3 m and the corresponding aligned lane segments. Finally, our lane accurate HD-Map is generated by merging both sets of lane segments and removing the overlapping segments.

4 Experiment

We used a total station, a mobile mapping system (Riegl VMX-250), and a 3D laser scanner (Velodyne HDL-64 S2) to map a complex junction, *Königsworther Platz* in Hannover. First, we measured poles around the area of the *Königsworther Platz* to generate a map of poles for the geo referencing procedure (BRENNER & HOFMANN, 2012). Secondly, we did several mobile mapping measurements from different directions along the junction. Finally, we recorded the traffic at the junction for two hours with the 3D laser scanner; see Fig. 5. We used two different positions for the measurements to generate trajectories that are more comprehensive and in order to evaluate our automatic geo-referencing approach. Thus, we generated 72000 (36000 per scan pose) point clouds and 4621 trajectories with a length of more than 10 m.

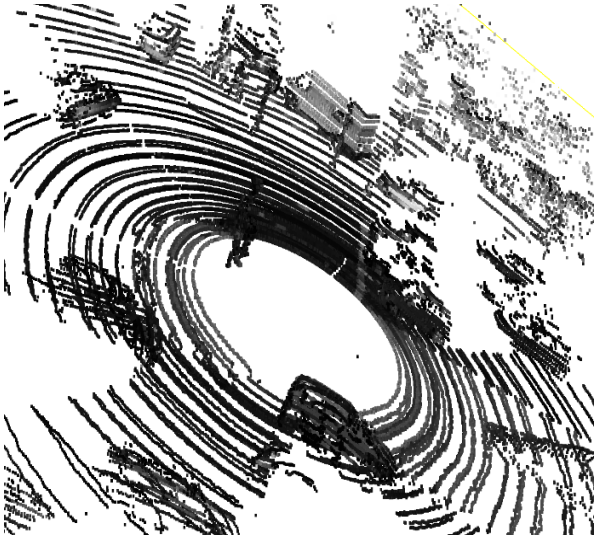


Fig. 5: Velodyne data

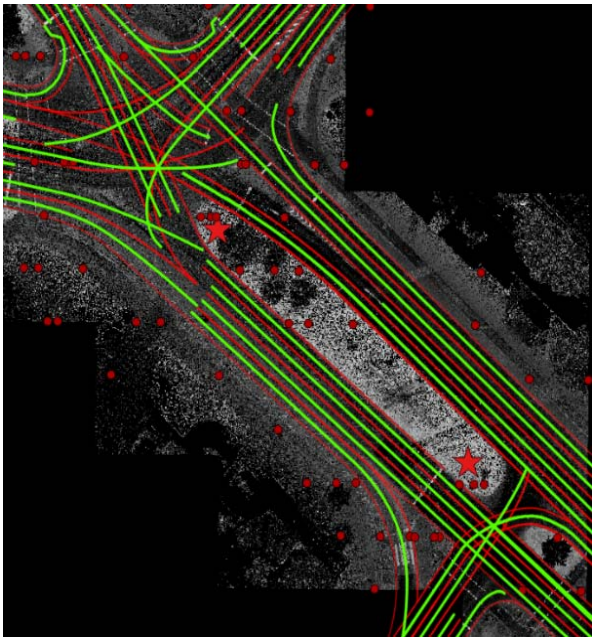


Fig. 6: Reference data lane borders (red lines), middle axis (green lines) and poles (red points)

4.1 Static reference data

Fig. 6 shows the 91 poles (red dots) which were measured by a total station with an accuracy better than 10 cm (BRENNER & HOFMANN 2012). Several mobile measurement drives were used to generate a highly detailed point cloud with 165.6 million points. We picked 1956 points from this point cloud at lane marking and curbs to estimate 35 lane segments with their borderlines (red) and middle axes (green). In the two cases where one lane border was not clearly defined by markings or curbs, we interpolated the lane border with a width of 3 m. Afterwards, we calculated the middle axis of each lane segment to evaluate our automatically detected middle axis.

5 Evaluation

First, we evaluated the geo-referencing of the scanner positions by calculating the root mean square error between the automatically detected poles and the manually measured poles. Secondly, we calculated the error between the automatically detected and the manually determined lane middle axis.

5.1 Geo-referencing

Fig. 7 shows the scanner positions (red stars) which were generated by calculating the transformation of the 21 automatically detected pole positions (blue circles) to the 91 manually measured pole positions (white circles). The RMSE errors for positions are 0.05 m and 0.07 m.

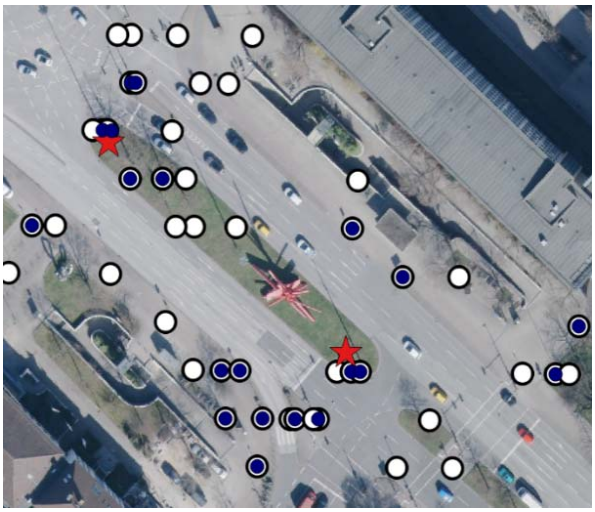


Fig. 7: Geo-referencing, manually measured poles (white circle), automatically detected poles (blue circle) and scanner position (red stars). Ortho image: (HANNOVER 2017)

5.2 Map

We evaluated the quality of our lane accurate map by comparing the lane middle axis based on the manually detected points (mp) at lane markings and curbs to the automatically detected middle axis grid in 10 cm steps. Fig. 8 shows the histogram of the distance for each manually detected point to the nearest automatically detected lane (left) and the distance for the automatically detected middle axis point to the nearest manually detected middle axis lane (right). The large distances reflect the challenge of defining the center of vehicles tracked by a laser scanner. Depending on scanning a vehicle frontally or from the side, the center varies up to half the vehicle width. Fig. 9 shows that the biggest error occurs at the border of the evaluation or tracking area and at areas with missing tracks and additional tracks from pedestrians and cyclists. Moreover, the direct alignment of the

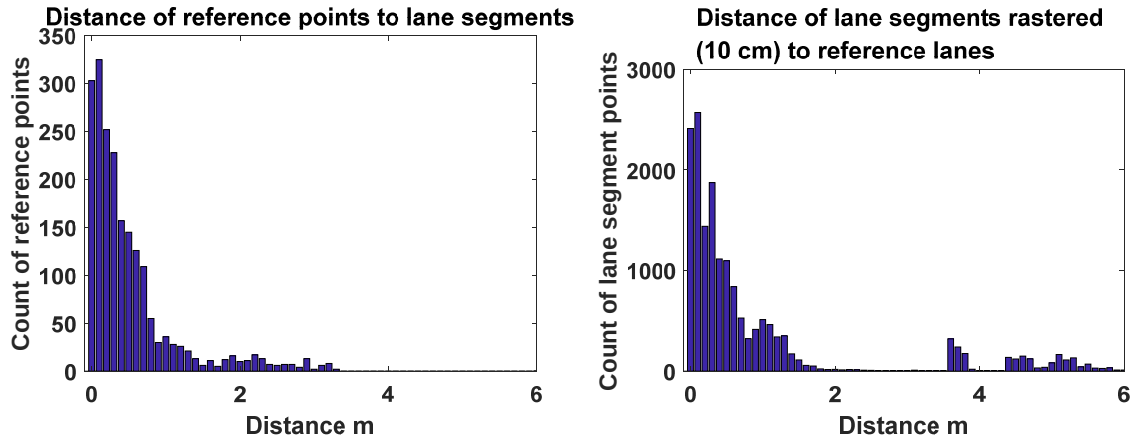


Fig. 8: Histograms of distance between reference lanes and automated lanes distance for each reference point (left) and distances for the automated lane segments (right).

middle axis based on trajectories could not compensate for common driving maneuvers illustrated in the center of the junction; during a turn, vehicles do not drive in the middle of the lane. The histogram (Fig. 8) shows that more than 500 m middle axis were automatically detected in the global UTM coordinate frame with less than 20 cm error.

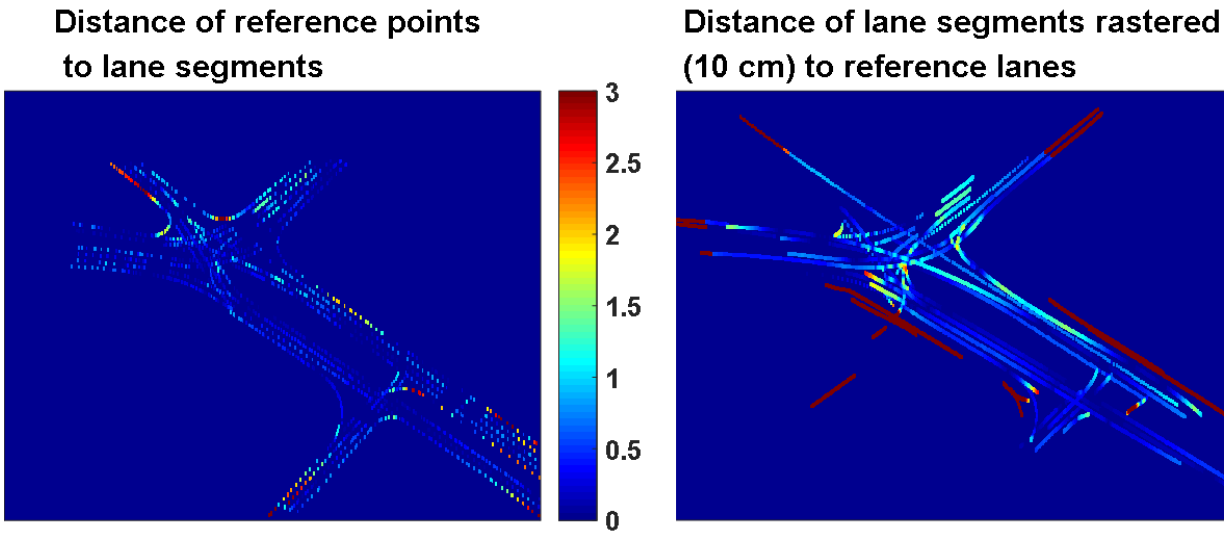


Fig. 9: Heat map of error for reference points (left) and lane segments (right), clamped in the interval 0 to 3 m

6 Discussion and Outlook

We discussed the quality of the result, the problems with mapping the middle axis of lanes by 3D laser scanner measurements and detection problems. We end with an outlook of the problems we addressed and the way we plan to improve our HD-Mapping by analyzing trajectory data.

6.1 Discussion

The problem of missing tracks could be solved by more reliable detections and a more comprehensive measurement via crowd sensing. The mix of vehicles, pedestrians and cyclists participating in the traffic could be distinguished by a classification of detections or other filter solutions. The exact middle positions of detected vehicles could be determined by aligning all point clouds for each detected track and estimating an active shape model to this data (COENEN et al. 2017). Middle axis differences because of driving behaviors could be addressed by generating a more model-based view of a junction. The high accuracy of below 20 cm for 34% percent of our aligned lane segments, ignoring wrong assignments (error > 3 m), shows the potential to generate lane accurate models from trajectories of crowd sensed data like laser scanner measurements.

6.2 Outlook

We plan to complete missing tracks and distinguish traffic participants by improving our detection procedure with pixel segmentation at the depth images by a neuronal network. Thus, we will improve the quantity and accuracy of the trajectories. Furthermore, we will solve the problem of

varying driver behavior and uncertainties by estimating more general junction models using a Markov Chain Monte Carlo optimization. Finally, we will use several vehicles equipped with Velodyne laser scanners to evaluate our approaches in larger inner city areas.

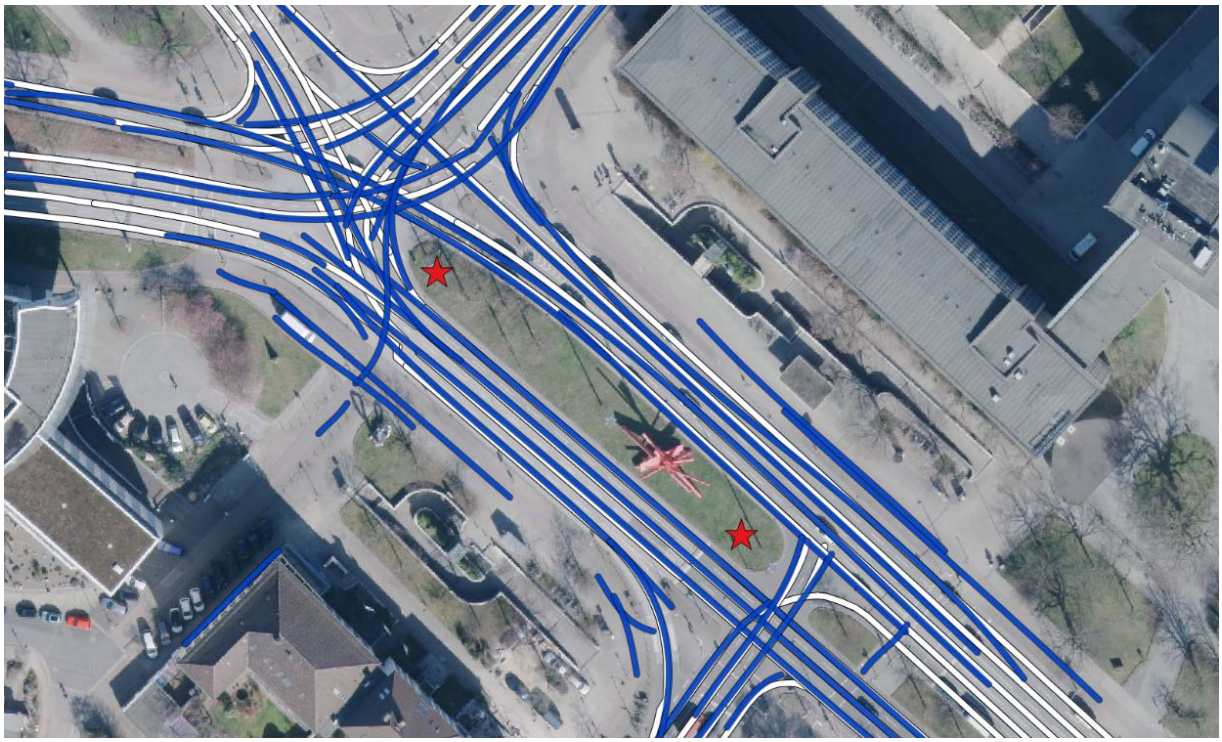


Fig. 10: Comparison of manually labeled (white) and automatic detected (blue) lane axes. Ortho image: (HANNOVER 2017)

7 ACKNOWLEDGEMENTS

Steffen Busch and Jannik Quehl were funded by the German Science Foundation DFG within the priority program SPP1835, “Cooperative Interacting Automobiles”.

8 References

AHMED, M., KARAGIORGOU, S., PFOSER, D. & WENK, C., 2015: A comparison and evaluation of map construction algorithms using vehicle tracking data. GeoInformatica 19, 601-632.

BENDER, P., ZIEGLER, J. & STILLER, C., 2014. Lanelets: Efficient map representation for autonomous driving. in: Intelligent Vehicles Symposium Proceedings, IEEE, 420-425.

BRENNER, C. & HOFMANN, S., 2012: Evaluation of Automatically Extracted Landmarks for Future Driver Assistance Systems. in: Yeh, A.G.O., Shi, W., Leung, Y., Zhou, C. (Eds.), Advances in Spatial Data Handling and GIS. Springer Berlin Heidelberg, 169-181.

CATMULL, E. & ROM, R., 1974: A class of local interpolating splines. in: Computer Aided Geometric Design. Elsevier, 317-326.

COENEN, M., ROTTENSTEINER, F. & HEIPKE, C., 2017. Detection and 3D modelling of vehicles from terrestrial stereo image pairs. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII-1/W1, 505-512. https://doi.org/10.5194/isprs-archives-XLII-1-W1-505-2017

FISCHLER, M.A. & BOLLES, R.C., 1981: Random sample consensus a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM **24**, 381-395.

GOODCHILD, M.F., 2007. Citizens as sensors: the world of volunteered geography. GeoJournal, **69**, 211-221.

GUAN, H., LI, J., CAO, S. & YU, Y., 2016. Use of mobile LiDAR in road information inventory: a review. Int. J. Image Data Fusion, **7**, 219-242. https://doi.org/10.1080/19479832.2016.1188860

HAKLAY, M., 2013. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. in: Crowdsourcing Geographic Knowledge. Springer, Dordrecht, 105-122. https://doi.org/10.1007/978-94-007-4587-2_7

HAKLAY, M. & WEBER, P., 2008. OpenStreetMap: User-Generated Street Maps. IEEE Pervasive Comput., **7**, 12-18. https://doi.org/10.1109/MPRV.2008.80

HANNOVER, 2017: Orthophoto Hannover Koenigsworther Platz [WWW Document]. URL https://www.hannover.de/Leben-in-der-Region-Hannover/Verwaltungen-Kommunen/Die-Verwaltung-der-Landeshauptstadt-Hannover/Dezernate-und-Fachbereiche-der-LHH/Baudezernat/Fachbereich-Planen-und-Stadtentwicklung/Geoinformation/Open-GeoData/Digitale-Orthophotos-DOP20 (accessed 1.26.18).

HOUENOU, A., BONNIFAIT, P., CHERFAOUI, V. & YAO, W., 2013: Vehicle trajectory prediction based on motion model and maneuver recognition. in: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 4363-4369.

HU, M., ZHONG, Z., NIU, Y. & NI, M., 2017: Duration-Variable Participant Recruitment for Urban Crowdsourcing With Indeterministic Trajectories. IEEE Trans. Veh. Technol. 66, 10271-10282. https://doi.org/10.1109/TVT.2017.2718043

JUNG, S., YOUN, J. & SULL, S., 2016: Efficient lane detection based on spatiotemporal images. IEEE Trans. Intell. Transp. Syst., **17**, 289-295.

KNOOP, V.L., DE BAKKER, P.F., TIBERIUS, C.C. & VAN AREM, B., 2017: Lane Determination With GPS Precise Point Positioning. IEEE Trans. Intell. Transp. Syst., **18**, 2503-2513

LANDSIEDEL, C. & WOLLHERR, D., 2017: Road geometry estimation for urban semantic maps using open data. Adv. Robot., **31**, 282-290. https://doi.org/10.1080/01691864.2016.1250675

NEITZEL, F. & KLONOWSKI, J., 2011: Mobile 3D mapping with a low-cost UAV system. Int Arch Photogramm Remote Sens Spat Inf Sci., **38**, 1-6.

PUENTE, I., GONZÁLEZ-JORGE, H., MARTÍNEZ-SÁNCHEZ, J. & ARIAS, P., 2013: Review of mobile mapping and surveying technologies. Measurement, **46**, 2127-2145.

RABE, J., NECKER, M. & STILLER, C., 2016: Ego-lane estimation for lane-level navigation in urban scenarios. in: Intelligent Vehicles Symposium (IV), IEEE, 896-901.

RAZA, A. & ZHONG, M., 2017: Hybrid lane-based short-term urban traffic speed forecasting: A genetic approach. 2017 4th International Conference on Transportation Information and Safety (ICTIS). 271-279. https://doi.org/10.1109/ICTIS.2017.8047776

ROETH, O., ZAUM, D. & BRENNER, C., 2017: Extracting Lane Geometry and Topology Information from Vehicle Fleet Trajectories in Complex Urban Scenarios Using a Reversible Jump Mcmc Method. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. **4**, 51.

RUHHAMMER, C., BAUMANN, M., PROTSCHKY, V., KLOEDEN, H., KLANNER, F. & STILLER, C., 2017: Automated Intersection Mapping From Crowd Trajectory Data. IEEE Trans. Intell. Transp. Syst. **18**, 666-677.

RUSU, R.B. & COUSINS, S., 2011: 3D is here Point Cloud Library (PCL). IEEE International Conference on Robotics and Automation (ICRA), 1-4.

SCHLICHTING, A. & BRENNER, C., 2014: Localization using automotive laser scanners and local pattern matching, in: Intelligent Vehicles Symposium Proceedings, 2014 IEEE. IEEE, 414–419.

SCHUBERT, R., RICHTER, E. & WANIELIK, G., 2008. Comparison and evaluation of advanced motion models for vehicle tracking. 11th IEEE International Conference on Information Fusion, 1-6.

TANG, L., YANG, X., DONG, Z. & LI, Q., 2016: CLRIC: collecting lane-based road information via crowdsourcing. IEEE Trans. Intell. Transp. Syst., **17**, 2552-2562.

TOTH, C. & JÓŹKÓW, G., 2016: Remote sensing platforms and sensors: A survey. ISPRS J. Photogramm. Remote Sens., Theme issue "State-of-the-art in photogrammetry, remote sensing and spatial information science", **115**, 22-36. https://doi.org/10.1016/j.isprsjprs.2015.10.004

TREIBER, M. & KESTING, A., 2013: Trajectory and Floating-Car Data, in: Traffic Flow Dynamics: Data, Models and Simulation. Springer Berlin Heidelberg, Berlin, Heidelberg, 7-12. https://doi.org/10.1007/978-3-642-32460-4_2

VIVACQUA, R.P.D., BERTOZZI, M., CERRI, P., MARTINS, F.N. & VASSALLO, R.F., 2017: Self-Localization Based on Visual Lane Marking Maps: An Accurate Low-Cost Approach for Autonomous Driving. IEEE Trans. Intell. Transp. Syst., 1-16. https://doi.org/10.1109/TITS.2017.2752461

Erfassung von Geometriedaten kleiner Flüsse mit einem unbemannten Wasserfahrzeug als Multisensor-Plattform

HANNES SARDEMANN[1], ANETTE ELTNER[1] & HANS-GERD MAAS[1]

Zusammenfassung: In den letzten Jahren treten an Flüssen mit kleinen und mittleren Einzugsgebieten vermehrt Sturzfluten auf. In dem interdisziplinären Forschungsprojekt EXTRUSO werden hydro-logische Modelle entwickelt, mit denen diese starkregeninduzierten Hochwasserereignisse besser modelliert und vorhergesagt werden können. Dafür werden u.a. hochaufgelöste Geländemodelle der Uferbereiche und des Gewässerbettes benötigt. Des Weiteren kann durch eine Analyse der landschaftlichen Veränderungen der Uferbereiche im Anschluss an ein Hochwasserereignis das Ausmaß des Ereignisses erfasst und somit das Verständnis über das Ereignis verbessert werden, um somit Vorsorgemaßnahmen abzuleiten. Für die Aufnahme von Geländemodellen im Flussbereich wird dabei neben der Befliegung mit einem UAV das Gewässer mit einem unbemannten Wasserfahrzeug (unmanned water vehicle – UWV) befahren. Das UWV dient als Multisensorplattform und digitalisiert sein Umfeld über Wasser mit einer 360°-Kamera und einem mobilen Laserscanner. Die vom Laserscanner aufgenommene Punktwolke der Ufer kann mit den Panoramabildern der 360°-Kamera koloriert werden. Unter Wasser wird das Gerinne mit einem Echolot aufgenommen. Das UWV ist außerdem mit einer Positionierungseinheit aus GNSS-gestützter IMU ausgestattet.
Im Beitrag wird das Konzept des Multisensor-UWV vorgestellt, und es werden Resultate aus ersten Pilotstudien gezeigt.

1 Einleitung

Innerhalb der letzten Jahrzehnte ist eine Zunahme von Starkregenereignissen, welche sich durch eine große Niederschlagsmenge in kurzer Zeit auszeichnen, zu beobachten (MÜLLER & PFISTER 2011). Dies erhöht das Risiko für Sturzfluten, die nach kurzen, kleinräumigen Starkregenereignissen insbesondere in Flüssen mit kleinen Einzugsgebieten auftreten (BRONSTERT et al. 2017). Kleine Flüsse reagieren auf Niederschlag mit schnell ansteigendem Wasserstand und es kommt zu plötzlichen Sturzfluten mit oft schweren Folgen. Gerade Flüsse mit kleinen und mittleren Einzugsgebieten werden aber kaum durch Pegelstationen überwacht. Diese Umstände erfordern neue Ansätze zur Sturzflutvorhersage.
In dem Forschungsprojekt EXTRUSO werden neue Methoden zur Modellierung und Vorhersage von Extremereignissen in kleinen und mittleren Einzugsgebieten entwickelt. Dazu werden zum einen die Messnetze für meteorologische, hydrologische und hydro-morphologische Parameter mit kostengünstigen Sensoren verdichtet. Neben low-cost Regensensoren werden kamerabasierte Pegel entwickelt (ELTNER et al. 2018). Außerdem werden neue hydrologische und hydrodynamische Modelle entwickelt und bestehende angepasst, um den Anforderungen für die Sturzflutvorhersage zu entsprechen. Für die Modellierung werden digitale Geländemodelle (DGM) sowohl

[1] Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstr. 10, D-01069 Dresden, E-Mail: [hannes.sardemann, anette.eltner, hans-gerd.maas]@tu-dresden.de

der Ufer als auch des Flussprofils benötigt. U.a. wird im Rahmen des Projekts der Einfluss der Auflösung der Geländemodelle auf die Modellierungsergebnisse untersucht. Neben den vom Amt verfügbaren DGMs mit 0,5 m Auflösung werden eigene, hochaufgelöste DGMs von Flussabschnitten erzeugt. Dafür werden eine Drohne (UAV, unmanned aerial vehicle) und ein ferngesteuertes Messboot (UWV, unmanned water vehicle) eingesetzt. Vom UAV werden mit einer Kamera Bilder des Geländes im Uferbereich aufgenommen und daraus mittels Structure-from-Motion (SfM) eine Punktwolke berechnet. Die Ufer und das Flussprofil sind für die hydrologische und hydrodynamische Modellierung von besonderer Bedeutung und werden vom Wasser aus mit einem UWV aufgenommen. Anschließend sollen die Punktwolken, die mit UAV und UWV generiert wurden, fusioniert und ein gerastertes DGM inklusive Flussprofil erstellt werden.
In diesem Beitrag wird das Konzept eines Multisensor-UWV mit mobilem Laserscanner und 360° Kamera für die Kartierung über Wasser und einem Echolot für die Kartierung unter Wasser vorgestellt. Ergebnisse aus ersten Pilotstudien zeigen das Potential des Systems.

2 Das unbemannte Wasserfahrzeug

Als Plattform für die Vermessung der Ufer und des Flussprofils vom Wasser aus wird das ferngesteuerte Messboot HyDrone von Seafloor Systems verwendet. Das Katamaran wurde vom Hersteller für bathymetrische Messungen mit einem Echolot entwickelt und von den Autoren zu einem Multisensorsystem umgebaut (Abb. 1).

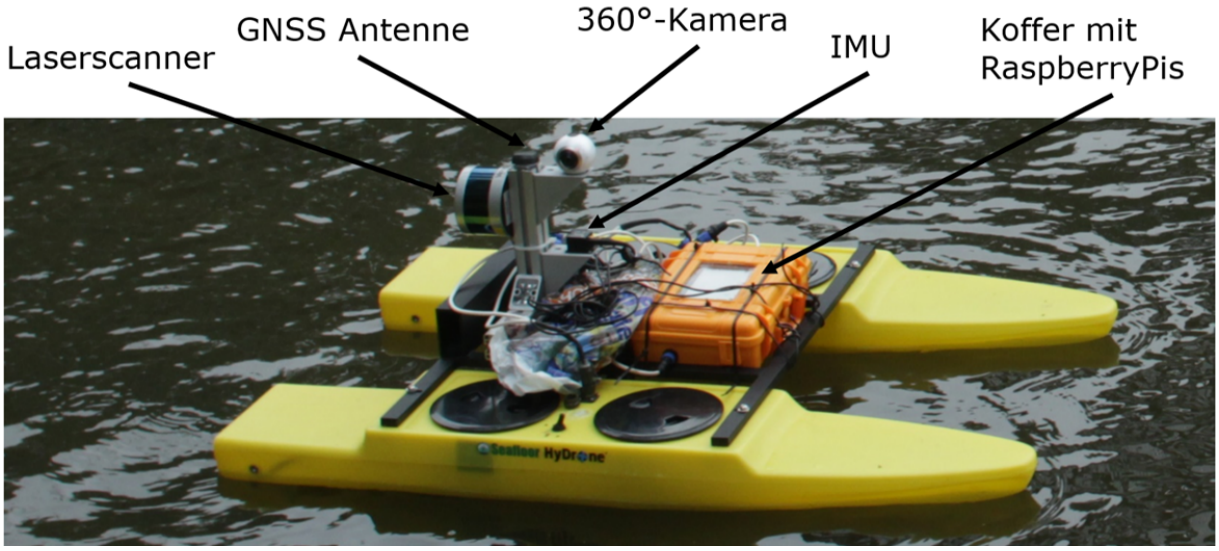


Abb. 1: UWV mit Sensorik. Über Wasser werden Laserscanner und 360°-Kamera eingesetzt. Unter Wasser wird ein Echolot verwendet (in dieser Abbildung verdeckt). Die Positionierung geschieht durch eine IMU mit GPS-Empfänger. Als Recheneinheit sind zwei RaspberryPi Einplatinenrechner an Bord

Über Wasser tastet ein Velodyne Puck Laserscanner die Ufer ab. Der Laserscanner zeichnet parallel 16 Profile mit 0,3 Millionen Punkten pro Sekunde und einer maximalen Reichweite von 100 m auf. Die Messgenauigkeit des Laserscanners beträgt +/- 3 cm. Mit einer Samsung Gear

360 wird ein Panoramavideo der Messfahrt aufgenommen, welches zur Kolorierung der Laserscanner-Punktwolke genutzt werden kann. Die 360°-Kamera besteht aus zwei Fisheye-Objektiven mit Öffnungswinkeln >180°, die in entgegengesetzte Richtungen ausgerichtet sind und ein vollsphärisches Panorama mit 7,3 MP aufnehmen. Für die Positionierung wird die GNSS-gestützte inertiale Messeinheit (IMU, inertial measurement unit) Spatial von Advanced Navigation verwendet. Wassertiefen werden mit dem Einzelpunkt-Echolot ECT400 von Echologger mit einer maximalen Aufnahmerate von 10 Hz gemessen. Durch die modulare Bauweise können weitere Sensoren integriert werden. Die Ansteuerung der Sensoren und die Datenspeicherung werden von zwei RaspberryPi Einplatinenrechnern durchgeführt.

2.1 Zeitliche Synchronisierung der Sensoren

Für eine Fusion der Daten aller Sensoren müssen diese zeitlich synchronisiert bzw. im gleichen Zeitsystem aufgenommen werden. Als Referenz wird die Uhr der IMU genutzt, die mit der GPS-Zeit synchronisiert ist und eine Genauigkeit von 50 ns besitzt. Über eine RS232 Schnittstelle sendet die IMU ihre Zeit und Position per NMEA $GPRMC-Protokoll mit 1 Hz an den Laserscanner. Die interne Uhr des Laserscanners ist bei einer aktiven GPS-Synchronisierung +/- 0,05 ns genau und gibt zu jedem gemessenen 3D-Punkt einen Zeitstempel an. Die Messwerte des Echolots werden, ebenfalls über eine RS232 Schnittstelle, direkt an die IMU gesendet, die diese zusammen mit den Messungen der IMU abspeichert und mit einem Zeitstempel versieht.
Die Samsung Gear 360 wurde als Consumer-Kamera für die Anwendung mit einem Smartphone entwickelt und kann nicht von der IMU oder einem Computer getriggert werden. Um die Bilder trotzdem zu synchronisieren, ist ein Display auf dem Boot angebracht, auf dem der aktuelle Zeitstempel der IMU mit 20 Hz Aktualisierungsrate angezeigt wird. Der Bildschirm ist während der kompletten Messfahrt im Sichtbereich der 360° Kamera. So kann für jede Aufnahme die Zeit (per Texterkennung oder manuell) abgelesen werden.

2.2 Kalibrierung der relativen Orientierungen

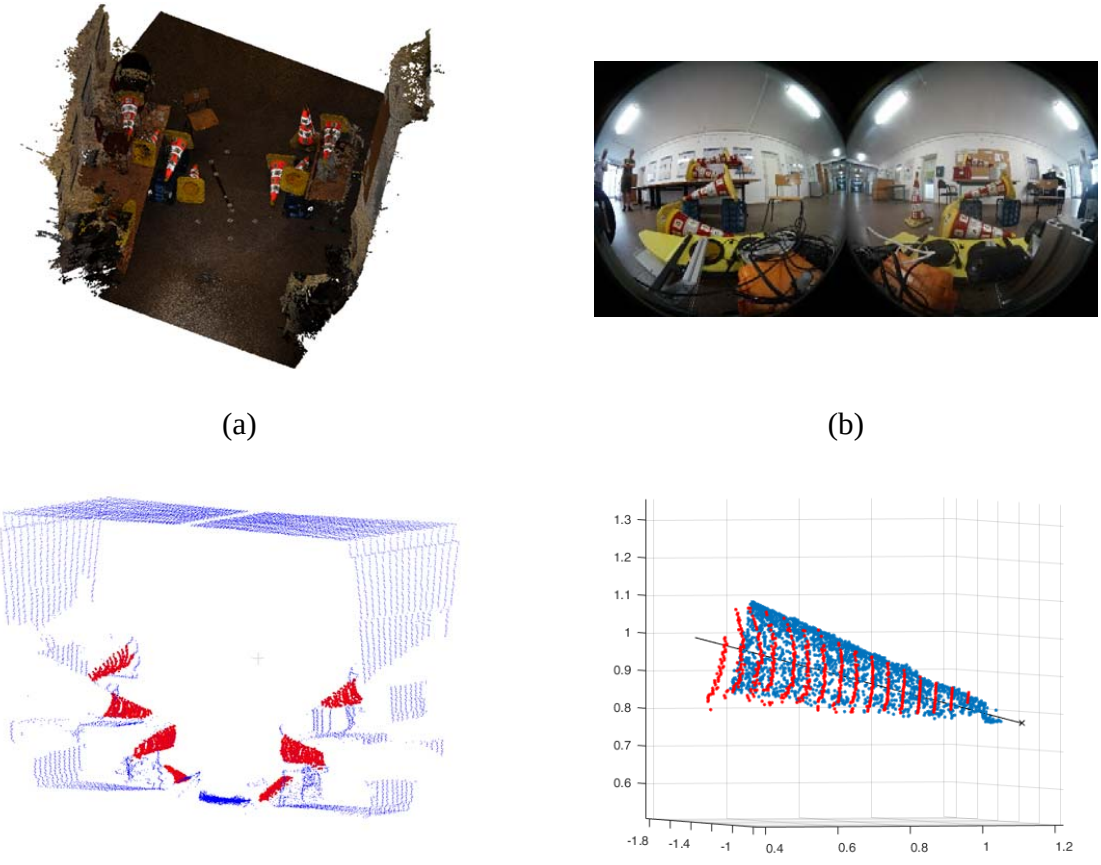
Für die Fusion der Sensordaten im gleichen Koordinatensystem muss die relative Orientierung der einzelnen Sensoren kalibriert werden. Die Kalibrierung wird mit dem in MADER et al. (2014) vorgestellten Verfahren durchgeführt. Das Kalibrierfeld wird aus Kegeln, die mit Messmarken beklebt wurden, aufgebaut und mit photogrammetrischer Mehrbildauswertung eingemessen (Referenzmessung, Abb. 2a). Das UWV wird in der Mitte des Kalibrierfelds platziert, so dass die Kegel sowohl von der 360° Kamera (Abb. 2b) als auch vom Laserscanner (Abb. 2c) aufgenommen werden. Mit einem räumlichen Rückwärtsschnitt mit Fisheye-Kameramodell kann die äußere Orientierung der beiden Bilder der 360° Kamera bestimmt werden. Die Abbildung von Objektpunkten im Referenzsystem ($\boldsymbol{X}_{\text{ref}}$) setzt sich aus einer Transformation M_{ref}^{gear} in das Kamerakoordinatensystem und einer Projektion in den Bildraum zusammen. Die Transformation besteht aus drei Rotationen (ausgedrückt durch die Rotationsmatrix R_{Ref}^{Gear}) und drei Translationen (Translationsvektor T_{Ref}^{Gear}):

$$(1) \quad \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}_{Kam} = \begin{bmatrix} R_{ref}^{gear} & T_{ref}^{gear} \\ [0\ 0\ 0] & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}_{ref} = M_{ref}^{gear} \cdot \boldsymbol{X}_{ref}$$

Das Kamerakoordinatensystem hat seinen Ursprung im Projektionszentrum der Kamera und liegt parallel zum Bildraum. Entsprechend eines äquidistanten Fisheye-Modells werden die Kamerakoordinaten in den Bildraum projiziert:

$$(2) \quad x_{Bild} = x_H + c \cdot \alpha \cdot \frac{x_{Kam}}{r_{obj}} + \Delta x \qquad (3) \quad y_{Bild} = y_H + c \cdot \alpha \cdot \frac{y_{Kam}}{r_{obj}} + \Delta y$$

Dabei ist (x_H, y_H) die Hauptpunktlage, c die Kamerakonstante, α der Einfallswinkel des Bildstrahls und $r_{obj} = \sqrt{X_{Kam}^2 + Y_{Kam}^2}$ der Objektpunktradius im Kamerakoordinatensystem.



(a) (b)

(c) (d)

Abb. 2: Kalibrierung der relativen Orientierung zwischen Laserscanner und 360°-Kamera. Das Kegel-Kalibrierfeld wurde mit einer Spiegelreflexkamera als Referenz eingemessen (a). In den Bildern der 360°-Kamera (b) werden die Marken gemessen und die äußere Orientierung mit RRS bestimmt. In der Punktwolke des Laserscanners (c) werden Kegel eingepasst und mit den Kegeln der Referenzpunktwolke zur Übereinstimmung gebracht (d)

Die Kalibrierung der Lage und Orientierung des Laserscanners im Referenzsystem wird über Kegel-Fits durchgeführt (Abb. 2d). Dafür werden in den Punktwolken der Referenzmessung und des Laserscanners Kegel eingepasst. In einer Ausgleichung wird die Transformation M_{ref}^{Puck} vom Referenzsystem in das Laserscannersystem bestimmt, indem die Parameter entsprechender Kegel

gleichsetzt werden. Die Kegel werden dabei durch die Lage der Kegelspitze, die Orientierung der Kegelachse und den Öffnungswinkel des Kegels parametrisiert.
Für die Kalibrierung des Boresight-Alignments (drei Rotationen) und Lever-Arms (drei Translationen) zwischen dem Laserscanner und der IMU wird das Boot um 180° gedreht und das Kegel-Kalibrierfeld erneut mit dem Laserscanner aufgenommen. Aus der relativen Bewegung zwischen beiden Positionen, die mit der IMU gemessen wurden und der Position und Orientierung des Laserscanners in beiden Epochen lässt sich die Transformation M_{Puck}^{IMU} vom Laserscannersystem in das System der IMU bestimmen. Da das Echolot und die IMU am gleichen senkrechten Alu-Profil angebracht sind, besteht nur ein Offset auf der Z-Achse. Für erste Untersuchungen wurde dieser Offset mit einem Zollstock gemessen. Im weiteren Verlauf des Projekts soll ein kombiniertes Kalibrierverfahren für über und unter Wasser entwickelt werden.
Nach der Kalibrierung können die Aufnahmen des Laserscanners, der 360°-Kamera und des Echolots in das System der IMU und somit in ein übergeordnetes Koordinatensystem (z.B. UTM) transformiert werden.

3 Messfahrten mit dem UWV

Das UWV wurde in ersten Messfahrten auf kleinen und mittleren Flüssen in Sachsen getestet. Im Folgenden werden Aufnahmen von Befahrungen eines ca. 1,5 km langen Abschnitts der Freiberger Mulde bei Leisning zwischen Dresden und Chemnitz vorgestellt.

3.1 Auswertung der Laserscanner-Daten

Die Auswertung der Laserscannerdaten erfolgt im Post-Processing. Die einzelnen Streifen werden zu einer gesamten Punktwolke vereinigt. Dazu werden die mit dem Laserscanner gemessenen 3D-Punkte ($\boldsymbol{X}_{\text{Puck}}$) zuerst entsprechend des Boresight-Alignments und Lever-Arms in das Koordinatensystem der IMU transformiert:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}_{IMU} = M_{Puck}^{IMU} \cdot \boldsymbol{X}_{Puck} \tag{4}$$

Da zu jedem 3D-Punkt ein Zeitstempel in der IMU-Zeit vorliegt, kann jeweils die Position (T_{IMU}^{global}) und Orientierung (R_{IMU}^{global}) der IMU zugeordnet werden. Die geographischen Längen und Breiten werden für die weitere Prozessierung zunächst in UTM-Koordinaten umgerechnet. Anschließend werden die 3D-Laserscanner-Punkte in das globale Koordinatensystem transformiert:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}_{global} = \begin{bmatrix} R_{IMU}^{global} & T_{IMU}^{global} \\ [0\ 0\ 0] & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}_{IMU} = M_{IMU}^{gloabl} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}_{IMU} \tag{5}$$

Abb. 3 zeigt die gesamte Punktwolke, die durch die Transformation der einzelnen Profile in das globale Koordinatensystem generiert wurde.

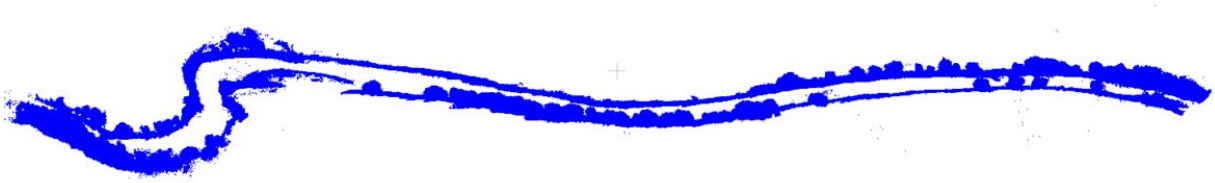


Abb. 3: 3D Punktwolke der Ufer. Die Punktwolke wurde mit dem mobilen Laserscanner generiert.

3.2 Auswertung der Echolotmessungen

Die Echolotmessungen werden ebenfalls im Post-Processing ausgewertet. Um möglichst viel Informationen über die Gewässerbettgeometrie mit dem Einzelpunkt-Echolot zu gewinnen, muss eine Befahrungsstrategie gewählt werden, die sich für eine Interpolation auf das gesamte Gerinne eignet. Nach SANTILLAN et al. (2016) ist eine Befahrung in einzelnen, dicht beieinanderliegenden Querprofilen am besten geeignet. Da aber in einer Messfahrt gleichzeitig die Wassertiefe und die Ufergeometrie mit Laserscanner und Kamera erfasst werden soll, ist eine Befahrung senkrecht zum Ufer ungeeignet. Stattdessen wird eine Zick-Zack-Befahrung durchgeführt und die Interpolation zusätzlich mit 3D-Laserscannermessungen des Übergangs von Wasser zu Ufer unterstützt. Dies kommt der Befahrungsstrategie sehr nah, die von SANTILLAN et al. (2016) am zweitbesten bewertet wurde.

Bei der Interpolation mäandrierender Flüsse kann es im Kurvenbereich passieren, dass Messwerte zwar geographisch nah beieinanderliegen und daher für die Interpolation berücksichtigt werden, obwohl keine topologische Nachbarschaft entlang des Flusses besteht. Um diese Einflüsse zu vermeiden, werden die Tiefenmessungen abgewickelt und in ein flussorientiertes Koordinatensystem mit einer Koordinatenachse entlang der Flussachse überführt. Dieses Koordinatensystem wird als anisotropes Koordinatensystem bezeichnet. Die Interpolation wird mit anisotropem Ordinary Kringing durchgeführt (MERWADE et al. 2006). In Abb. 4 ist das Interpolationsergebnis an dem Abschnitt der Freiberger Mulde zu sehen. Auf der westlichen Seite schließt sich ein Wehr an, so dass die Wassertiefe in diesem Bereich am tiefsten ist. Eine weitere tiefe Stelle ist am schmalsten Stück des Abschnitts zu beobachten.

3.3 Fusion der Punktwolken mit UAV-Daten

Abschließend wird die Laserscanner-Punktwolke der Ufer mit dem interpolierten Flussprofil und einer mit SfM aus UAV-Aufnahmen generierten Punktwolke der angrenzenden Bereiche fusioniert. In Abb. 4 sind die Daten der drei Quellen in UTM-Koordinaten dargestellt. Im weiteren Verlauf des Projekts soll daraus ein gerastertes DGM entstehen.

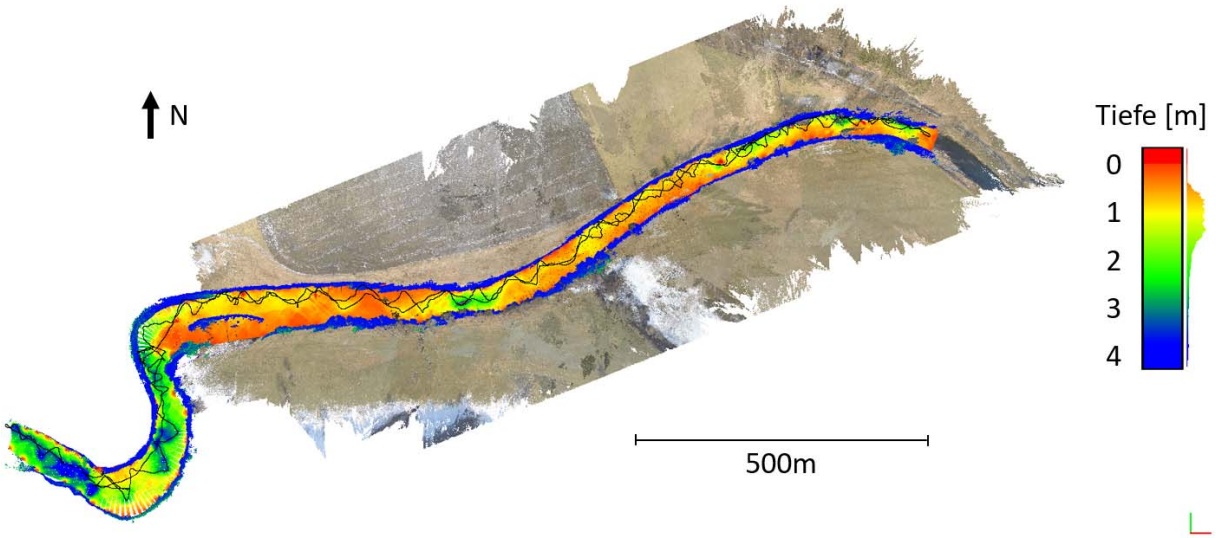


Abb. 4: Ergebnisse einer Testfahrt mit dem UWV und einer UAV-Befliegung. Die Laserscanner Punktwolke deckt ausschließlich die Ufer ab und ist blau eingefärbt. Die Echolotmessungen wurden mit anisotropem Ordinary Kringing interpoliert und sind entsprechend ihrer Tiefe eingefärbt. Die Schwarze Linie zeigt den Track der Befahrung in einem Zick-Zack-Muster auf Hin- und Rückweg. Das Teileinzugsgebiet, das die Ufer umgibt wurde mit einer UAV beflogen eine SfM-Punktwolke erzeugt.

4 Fazit & Ausblick

In diesem Beitrag wurde das Konzept, die Kalibrierung und erste Messungen mit einem UWV als Multisensorplattform vorgestellt. Das UWV ist mit einem Laserscanner und einer 360°-Kamera für die Vermessung der Ufer ausgestattet. Die Sensoren werden zeitlich synchronisiert und relativ zueinander und relativ zur IMU kalibriert. Das Flussprofil wird mit einem Echolot gemessen. Durch geeignete Befahrungsstrategien lässt sich aus den gemessenen Tiefen ein komplettes Gerinneprofil interpolieren. Die Punktwolken Über- und Unterwasser wurden im gleichen Koordinatensystem aufgenommen und lassen sich fusionieren. Es ist ebenfalls möglich, diese mit einer UAV-Punktwolke zu kombinieren.
Die Daten müssen in den folgenden Arbeitsschritten noch für eine hydrologische Modellierung nutzbar gemacht werden. Dafür muss unter anderem die Vegetation aus den Daten gefiltert werden. Es wird analysiert, in wie weit sich die Multi-Pulse Eigenschaft des Laserscanners dafür eignet. Die Zuverlässigkeit der Interpolation von Echolotmessungen ist stark abhängig von der Datenmenge. Zukünftig sollte daher eine dichtere Punktmenge erzeugt werden. Dies gelingt zum Beispiel durch Befahrung in Querprofilen. Alternativ wird ein Triangulations-Sensor aus Linienlaser und Unterwasserkamera entwickelt, der breitere Profile misst.

5 Danksagung

Diese Untersuchungen wurden im Forschungsprojekt “Extremereignisse in kleinen und mittleren Einzugsgebieten (EXTRUSO)“ durchgeführt, welches durch das EU ESF Programm „Nach-

wuchsforschergruppen“ unter der Fördernummer 100270097 unterstützt wird. Die UAV-Punktwolke wurde von der Professur für Geoinformatik der TU Dresden bereitgestellt und ist im Rahmen des Forschungsprojekts „Bootgestütztes Messsystem für longitudinale Gewässerprofile (BOOT-Monitoring)“ entstanden, welches Teil der BMBF-Fördermaßnahme „Regionales Wasserressourcen-Management für den nachhaltigen Gewässerschutz in Deutschland (ReWaM)“ ist.

6 Literaturverzeichnis

BRONSTERT, A., AGARWAL, A., BOESSENKOOL, B., FISCHER, M., HEISTERMANN, M., KÖHN-REICH, L., MORAN, T. & WENDI, D., 2017: Die Sturzflut von Braunsbach am 29. Mai 2016 – Entstehung, Ablauf und Schäden eines „Jahrhundertereignisses“. Teil 1: Meteorologische und hydrologische Analyse. Hydrologie und Wasserbewirtschaftung, **61**, 150-162.

ELTNER, A., SARDEMANN, H., KRÖHNERT, M. & MAAS, H.-G., 2018: Entwicklung eines Low-Cost Kamerapegels zur Erfassung hydrologischer Extremereignisse. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 27, 38. Wissenschaftlich-Technische Jahrestagung der DGPF und PFGK, 07.-09. März 2018 in München.

MADER, D., WESTFELD, P. & MAAS, H.-G., 2014: An Integrated Flexible Self-Calibration Approach for 2D Laser Scanning Range Finders Applied to the Hokuyo UTM-30LX-EW. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **40**(5), 385-393.

MERWADE, V.M., MAIDMENT, D.R. & GOFF, J.A., 2006: Anisotropic considerations while interpolating river channel bathymetry. Journal of Hydrology, **331**(3-4), 731-741.

MUELLER, E. N. & PFISTER, A., 2011: Increasing occurrence of high-intensity rainstorm events relevant for the generation of soil erosion in a temperate lowland region in Central Europe. Journal of Hydrology, **411**(3), 266-278.

SANTILLAN, J.R., SERVIANO, J.L., MAKINANO-SANTILLAN, M. & MARQUESO, J.T., 2016: Influence of River Bed Elevation Survey Configurations and Interpolation Methods on the Accuracy of Lidar DTM-based River Flow Simulations. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **42**(4/W1), 225-235.

Modeling and visualizing the spatial uncertainty of moving transport hubs in urban spaces - a case study in NYC with taxi and boro taxi trip data

ANDREAS KELER[1]

Abstract: Mobility in urban environments is complex due to numerous interacting components, with many of those that are difficult to specify. Examples include the presence of transport hubs, which connect different modes of transport, public and private. The properties of these locations include a temporally changing surface area of operational function that is heavily dependent on the complex and dynamically changing human mobility. Besides public transport services with specified stations, there are taxi services that can be associated with established transport hubs in the way of assigning spatiotemporal service hotspots. This work proposes a technique for relating taxi trip origins and destination hotspots for gaining knowledge on the spatial uncertainties of transport hubs, more precisely their movements within specific times. The case study in NYC focuses on the services of yellow taxi and boro taxi trip data on Saturdays in June 2015. The outcomes of applying the technique are matter of further investigation of spatial uncertainty perception, representation, and visualization. In the stages of the approach, the outcomes of transport hub movements are related with more general functional transport regions resulting from NYC public transport services.

1 Introduction

Urban human mobility relies on the presence of transport hubs, which connect different modes of transport, public and private. This paper proposes a technique for relating taxi trip origins and destination hotspots for gaining knowledge on the spatial uncertainties of transport hubs, more precisely their movements within specific times. Coming from the assumption that different transport modes are usable as a sequence within the day, the taxi services can be seen as one of the key modes, especially in complex urban environments. By assigning functional transport regions, often specific bus routes or metro lines are extracted together with respective stations. This functionality in public transport services is expressible via visualizing accessibilities, as for example by using detailed isochrones. Besides this, the fact of having dynamically changing accessibility surfaces resulting from dynamically changing traffic situations is essential for further analyses.

1.1 Describing the moving transport hub phenomena

Several appearances within the passenger behavior are connectable with the phenomena of the moving transport hub. By definition, movement of urban citizens can be perceived and especially represented differently. One general differentiation consists of grouping movement flows into movements of participants of one mode of transport or into individual movement that is often

[1] Technical University of Munich, Chair of Traffic Engineering and Control, Arcisstraße 21 D-80333 München, E-Mail: andreas.keler@mytum.de

intermodal in urban environments. Keeping this in mind, the movement representation varies with the spatial and temporal scale of analyzing movements in urban environments. On a very general level, one can define inter-city movements, where every city is aggregated by a single symbolization. From this type of representation, it is possible to define functional transport regions that may connect several cities, which comes from early research in transport geography, a discipline originating from human geography. By changing the scale to higher-resolution movement between city districts, it is possible to detect inter-city movements that can also indicate the quality value of selected districts in living and working. At the next higher resolution level, it is feasible to analyze movement between different stations of selected mobility services, especially from the established public transport. Modes using trains such as subways and tramways that usually base on installations of expensive infrastructures can have several mobility stations, such as tramway stations. Some of these stations have connections to nearby situated stations of other transport modes, which can be referred to as transport hubs. These transport hubs are special cases of mobility stations that connect stations of different transport modes within the same space, often with multiple elevation levels. In the possibly most common way in urban environments, the transport hub has numerous transport stations: different spatially fixed stations with respective connections usually situated in walking distance, and, stations that are dynamically changing in space and base on local knowledge or are inferable from vehicle movement observations. The latter can base on the vehicle movements of vehicle services such as taxi services. In this case, it is possible to estimate typical taxi trip origin and destination hotspots, which are connected with the taxi customer behavior of entering and leaving the taxi. These hotspots, especially when generated from taxi movement data acquisitions, are dynamically changing point clusters that occupy selected areas in space. Within these areas, it is possible to have numerous built transport infrastructural elements. Therefore, the appearance of a transport hub is in general more associable with a polygonal then with a pointwise representation in space. These ideas can be explained via an example from the operation of daily bus routes of the public transport in Shanghai.

Fig. 1 shows an example from BURGFELD & LIEBHARDT (2015) of such an appearance, mainly consisting of the comparison between two officially assigned bus routes (a) and not-preprocessed GNSS positioning data from respective buses from February 28, 2007 (b). Both bus routes in Shanghai, 20 and 36, are highly influenced by daily traffic congestion events. This makes it understandable that at certain times of working days, bus movements differ from official routes due to traffic congestion avoidance strategies. More precisely, these traffic congestion avoidance patterns appear often as slight variations of the route, as taking the byroad that is less congested. This is pictured in Fig. 1 by the large variations of spatial distribution of tracked bus movements, especially for bus route 20 (red).

In another way, we can speak of moving transport hubs, when bus stations are shifting in space within specific time windows. The example in Fig. 1b shows that buses of route 20 (in red) are moving around a larger area in the northwestern part of the map view, which is not assigned as an official part of the bus route. Serving specific needs of commuters, together with local knowledge can result in specific movement patterns. These patterns can appear around only one specific station or include two or more stations of different mobility services and modes of transport.

In another way, moving transport hubs are definable as slight variances of the mobility hotspots within the transport hub area.
These connections are in some cases visible, as in the specific example of bicycle parking stations above metro stations underground. The concentrations of bicycles, namely the bicycle densities within the transport hub vary spatially and, in some cases, spatiotemporally, always depending on an enormous selections of different factors influencing the urban mobility in general.

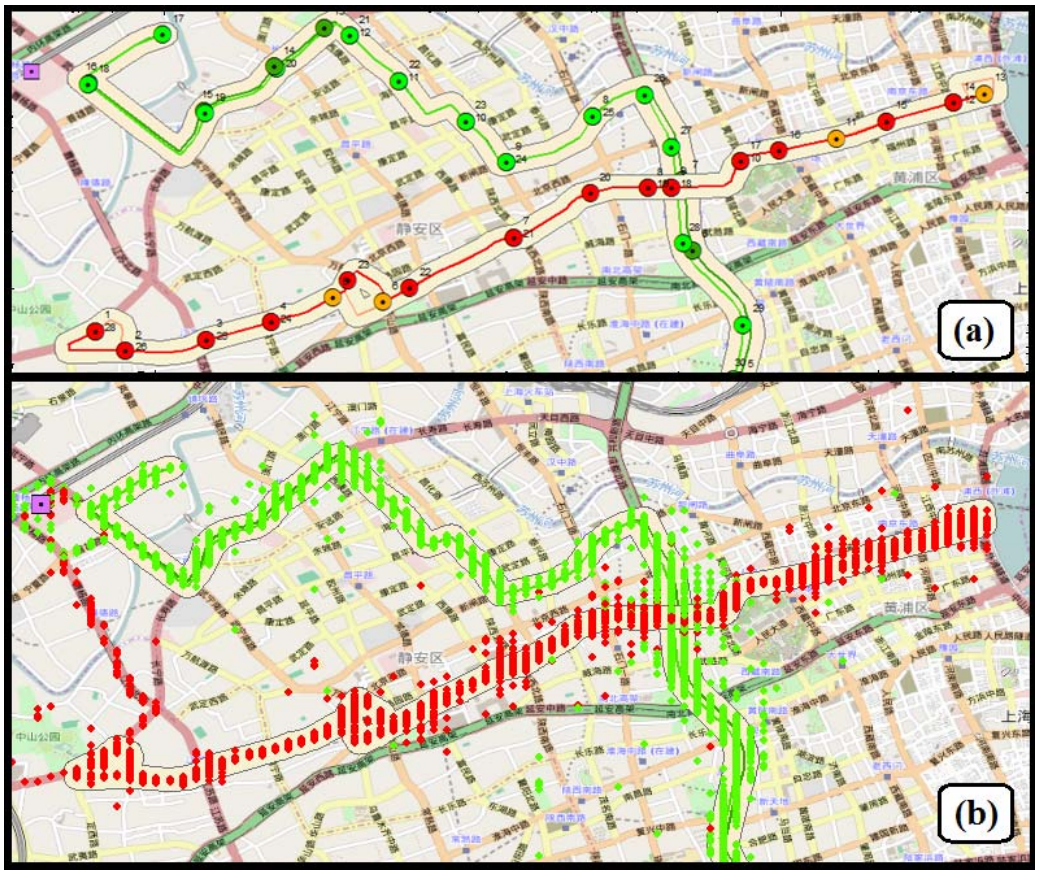


Fig. 1: Example for the presence of moving transport hubs with a comparison of (a) the official bus routes 20 (red) and 36 (green), and, (b) plotted raw positioning data[2] from respective bus movements from February 28, 2007, in Shanghai (BURGFELD & LIEBHARDT 2015)

1.2 Motivation

Urban environments are characterized by complex and dynamically changing human mobility. Vehicle movement trajectories of urban vehicle fleets, as taxi fleets, can help extracting this knowledge. Taxi trip data is usable in many different applications of different research domains.
Due to often-immense sizes of the data extracts, many applications make only use of taxi trip origin and destination points. In an aggregated view, these points may be clustered into origin

[2] 'SUVnet-Trace Data' was obtained from the Wireless and Sensor networks Lab (WnSN) at Shanghai Jiao Tong University, http://wirelesslab.sjtu.edu.cn/taxi_trace_data.html

and destination hotspots, which are connectable, especially, when there are different taxi fleets, with specific operational restrictions. This promotes the automated deduction of local knowledge on smaller scales of investigation. In a simple way, this deduction may consist of spatial queries for the presence of mobility service stations within agglomerations of origins and destinations of different taxis of different taxi fleets.

The question here is if it is possible to define the moving transport hub phenomena as a general collection of different appearances of urban mobility services, as pictured in Fig. 2 by the tree cases. Case 1 of Fig. 2 shows a size variation and movement of the defined taxi service hotspot polygons t1 and t2 around only one railway station. Dependent on the inspected time windows that correspond with t1 and t2, there are different conditions and traffic situations that influence the shapes of the two taxi service hotspot polygons. These polygons are then connected with on specific railway station that is situated within both. For case 1, this is the only condition for defining one transport hub polygon, which is definable via the polygon intersection, where the railway station is situated. Case 2 in Fig. 2 shows the appearance of having two intersecting taxi service hotspot polygons that have an intersection area without the presence of mobility stations. Nevertheless, mobility stations are respectively included in every polygon. One definition of a moving transport hub for case 2 would consist of a spatially larger area of influence (combination of t1 and t2) and one core intersection area. One important aspect of case 2 is the condition of having connections between the mobility stations in walking distance. Additionally, it is possible to reason on the connectivity on the operational level, as for example the specific official public bus route and respective driving directions of operating buses. Case 3 is defined, when these mentioned conditions do not apply. For case 3 in Fig. 2 the mobility stations are within the polygons t1 and t2 but are for example situated on different road segments, which indicates greater walking distances. Case 3 appears often at complex signalized intersections and indicates additional waiting times between specific mobility stations within the moving transport hub.

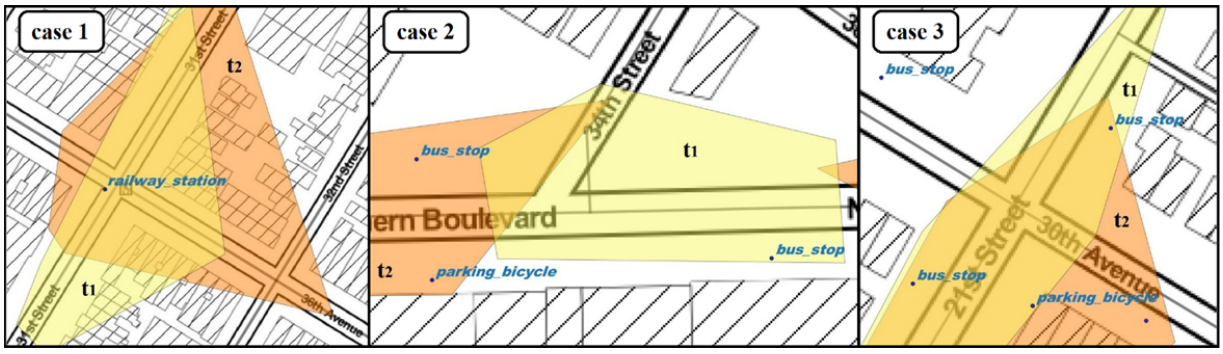


Fig. 2: Proposed definition of the urban transport hub phenomena as a collection of different appearances

All three cases in Fig. 2 represent the general ideas on modelling the moving transport hub phenomena. Understanding time-dependencies and dynamics of moving transport hubs can be beneficial for planning, decision-making, and various ongoing research in connection with smart cities and intelligent transportation systems (ITS). Besides gaining possible benefits for microscop-

ic traffic simulations, there are options for improving activity models of agent-based traffic simulations, which is partially also the aim of this work.

1.3 Clustering vehicle movement trajectories and representation as moving polygons

One simple way to gain more knowledge on massive vehicle movement data set is using point clustering techniques, with and without respecting the temporal component.
Besides selecting suitable clustering techniques, it is important to specify the data representation of the vehicle movement. By selecting whole trajectories of vehicle trips, it should first be clear at which accuracy the trajectories were acquired and with which sensor. Additionally, the state of quality estimation of the vehicle movement data set is important. Raw and not-preprocessed trajectories might evolve as data with a considerable proportion of outliers that can go up to 50% (KELER 2017). Correcting these outliers might be complex, especially when applying map matching techniques. The large number of variations in implementing the latter can cause complex adaptations onto specific data sets. When raw vehicle movement data from urban environments includes systematic signal losses due to the presence of higher buildings, selected parts of the urban road network are fully or partially missing. One possibility for extending these missing partitions is the usage of different interpolation techniques that may help reconstruct missing trajectory elements.
Depending on the application and the way of analysis, extracts of trajectories can serve as data input. Extracting points or positions from a movement trajectory can facilitate selected applications as for example defining spatiotemporal trip destination hotspots (KELER & KRISP 2016). In case of taxis, the usual origin and destination correspond with taxi customer pick-ups and drop-offs.
After defining the trajectory point extracts, it is possible to segment the data sets into data extracts based on selected time window as for example hour-wise for the time of the day or day-wise for day of the week. By comparing the clustering results of these partitions, usually with the same clustering parameters, it is possible to formulate first insights after visualization and visual inspection of the outcomes. In a more general way, these clusters can be aggregated into convex hulls, which improves in many cases the visual inspection, reduces data volumes, and, allows the inspection of matching polygons on a more general level. Matching between different types of areal polygons relies on the Dimensionally Extended nine-Intersection Model (DE-9IM) as introduced by CLEMENTINI et al. (1994).

2 Methodological approach for modelling and visualizing spatially uncertain moving transport hubs

The methods of the proposed technique include density-based point clustering, convex hull estimations, polygon matching, and polygon intersection. The components are part of an extendible workflow, which is pictured in Fig. 3.
Fig. 3a shows the workflow components that make use of spatial positions of taxi trip origins and destinations as data input for computing density-based point clusters for a selected data partitions based on the OPTICS algorithm (ANKERST et al. 1999). The subsequent step bases on JARVIS

(1973) and consists of computing the convex hull for every density-based point cluster. The aim of the mentioned polygonization of point clusters is providing visibility at certain map scales by creating larger polygons for spatially extensive taxi trip origin or destination distributions at for example multiple connected road segments. These taxi trip origins and destinations are respectively clustered based on their spatial distribution within a selected time window as for a whole day.

The subsequent procedure consists of polygon matching in Fig. 3b. Polygon intersection in Fig. 3c is one part of defining moving transport hubs in the sense of analyzing the intersection patterns of consecutive polygons within a time series. Matching between different types of areal polygons relies on the Dimensionally Extended nine-Intersection Model (DE-9IM) as introduced by CLEMENTINI et al. (1994). The definition of moving transport hubs comes along with the step of inferring the indicators for the certain movement from different mobility services. In general, this step is not specified in detail and there are different possible options. One option is to extract information from OpenStreetMap that can indicate selected functionalities related to transport.

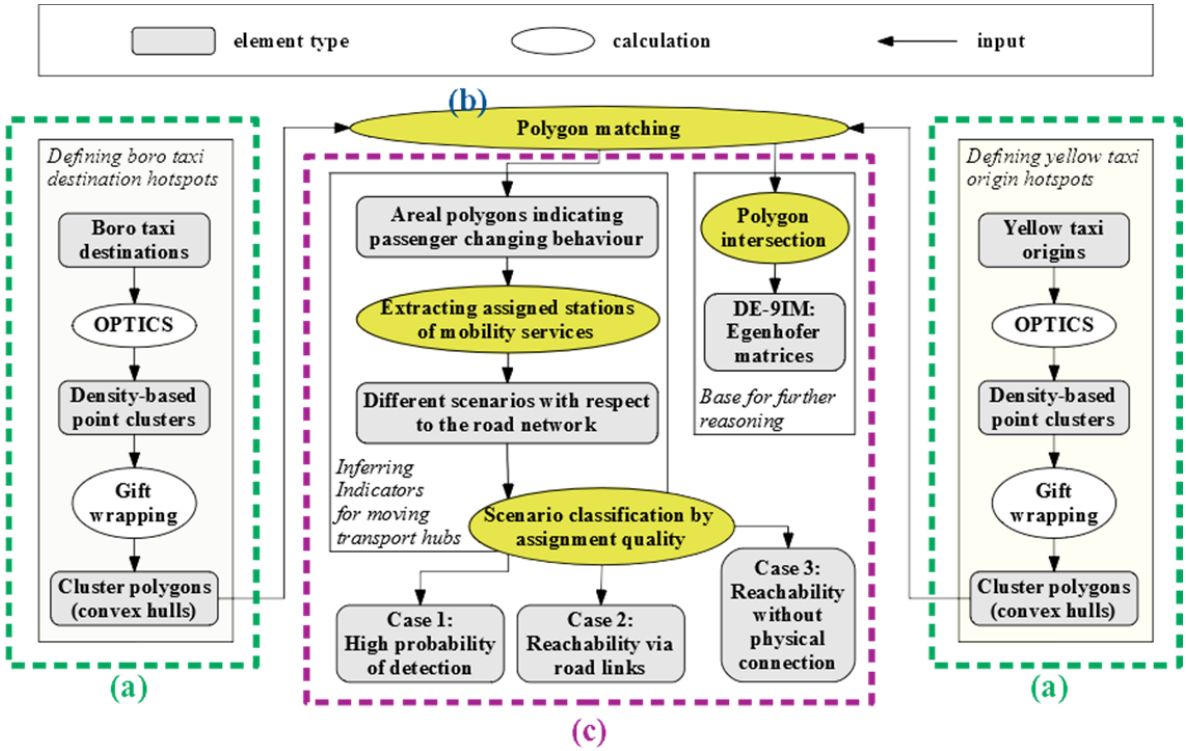


Fig. 3: Workflow of (a) generating origin and destination hotspot polygons, (b) its subsequent matching as consecutive sequences and same time windows, and (c) the definition of the moving transport hubs

By matching temporally subsequent taxi trip origin or destination hotspot polygons, it is possible gain more knowledge on the spatiotemporal distribution of service areas with similar functionalities. Fig. 4 shows four layers of boro taxi trip destination hotspots with introduced opacities. The darker the areas appear, the more common is the assignment as typical boro taxi trip destination hotspots. Besides the resulting boro taxi trip destination hotspots in Fig. 4, there are also ex-

tractable yellow taxi trip origin hotspots and functional transport regions resulting from public transport stations of subway and bus lines. This has a semantic connection to the restricted yellow zone and specific commuting patterns between Manhattan and the outer boroughs of NYC. The further steps include matching these spatiotemporal hotspots with characteristic transport hubs, and testing a specific visualization scheme for representing spatial and spatiotemporal uncertainties of the mentioned transport hubs. The outcomes of applying the techniques are matter of further inspections and visual analysis.

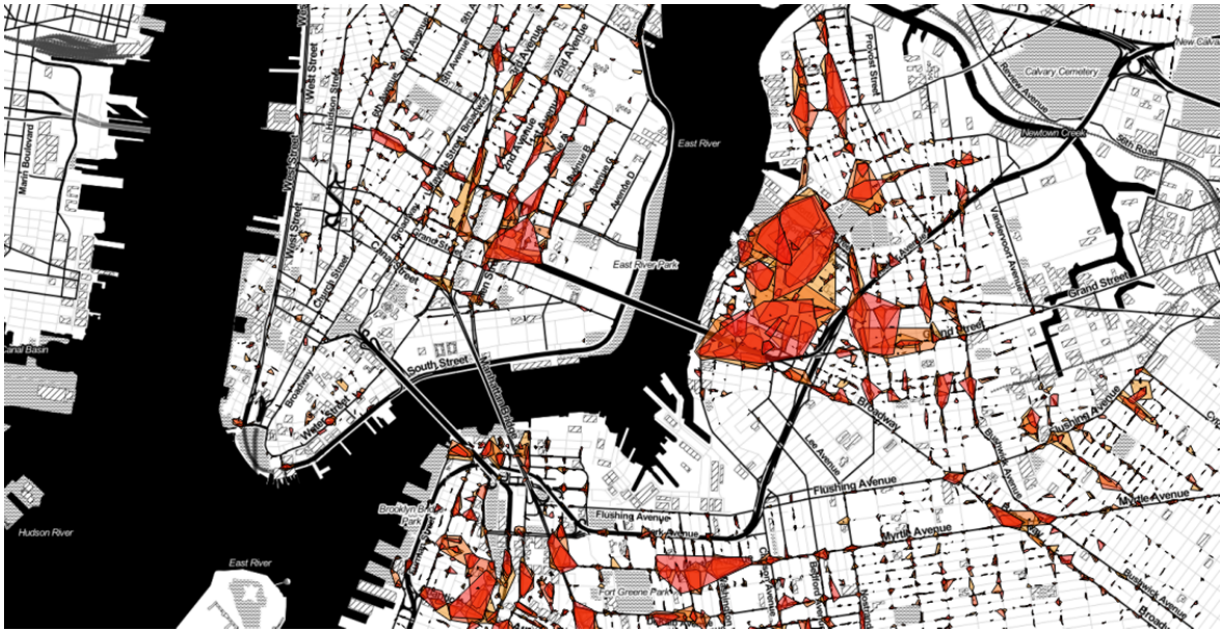


Fig. 4: Visualization of the polygon matching results, showing the spatial uncertainty of boro taxi trip destination hotspots for four successive time windows

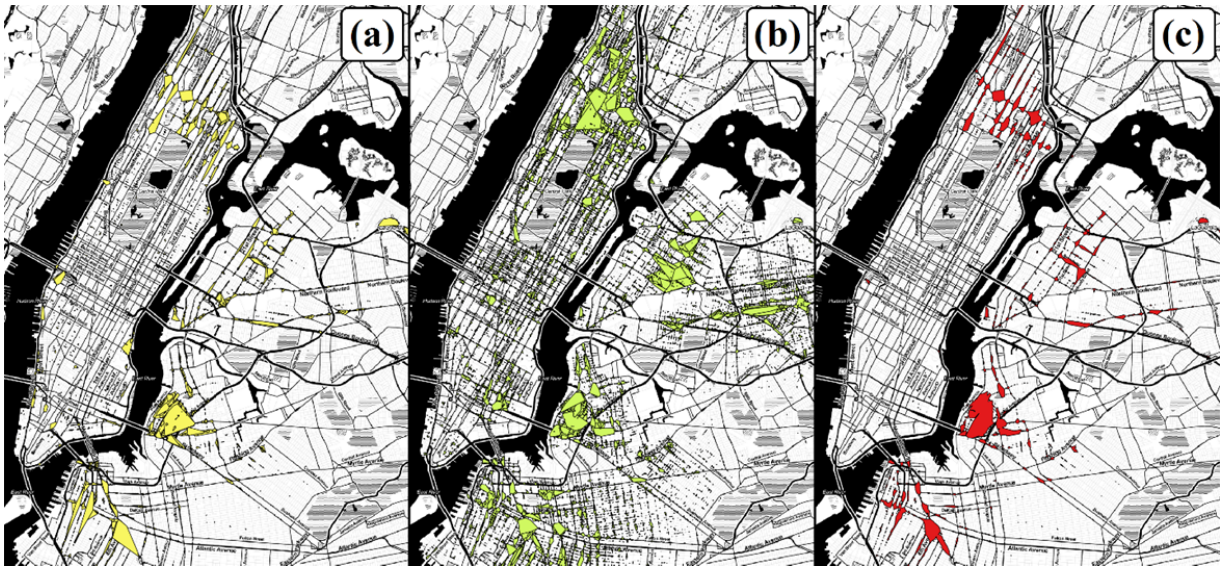


Fig. 5: Taxi hotspots polygons with (a) yellow taxi trip origins, and, (b) boro taxi trip destinations, and, (c) the intersection areas between yellow and boro taxi hotspots

Fig. 5 is interpretable as the merged product of all polygon layers within a time series. In case of Fig. 5, all four successive time windows of June 2015 are matched. Fig. 5a and 5b picture yellow taxi origin hotspots and green boro taxi destination hotspots. Fig. 5c shows the intersection product of both types of hotspots via red polygons. By further visual inspection, it is feasible to detect certain clusters in the spatial distribution of the intersection areas. One typical property of every intersection product between yellow taxi origin and boro taxi destination hotspots is the inclusion of the usually larger polygons around the border of the yellow zone, which has a restriction in customer pick-ups for boro taxi drivers. From this intersection product in Fig. 5c, it is possible to provide a preselection of traffic-related points of interest (POI's) or mobility service stations. Extracts from OSM serve as inputs for assigning those taxi hotspot intersection areas as moving transport hub, simply based on the presence of its existence in space. From the whole area of the State of New York, three selected types of points are extracted from OSM, which is pictured in Tab. 1.

Tab. 1: Selected information extracted from OpenStreetMap connected with functional transport regions and stations of mobility services in NYC

Type of information	POI's	traffic	transport
Feature classes or values	bicycle rental; car rental; car sharing	parking; parking bicycle; parking multistorey; parking underground	bus station; bus stop; ferry terminal; railway halt; railway station; taxi; tram stop
Number of extracted records	551	7 053	4 733

From a total number of 12 337 OSM point extracts 3 511 (= 28.5%) are matching with the intersection product in Fig. 5c. These extracted OSM points serve for the matching with every taxi hotspot intersection layer of every Saturday in June 2015.

3 Results and discussion

The presented methodology is tested with the mentioned selection of boro taxi and yellow taxi data from Saturdays in June 2015. Tab. 2 shows an overview on the used data partitions, the inferred taxi service hotspots, and, the generated moving transport hubs. By observing the low number of defined moving transport hubs in Tab. 2, it is possible to state that many yellow taxi origin and boro taxi destination hotspots are not located nearby stations of other mobility services in NYC.

Tab. 2 shows the unexpected outcomes of applying the workflow in Fig. 3. Despite the fact that only 33 496 yellow taxi trips appeared on June 6, all other yellow taxi trip on Saturdays are around thirteen times higher. Due to the fact that Tab. 2 only shows yellow taxi trip origins, we have higher origin densities at selected taxi stands, especially at the southern part of Manhattan, which is difficult to observe in Fig. 5a due to the smaller spatial extents of the hotspot polygons.

Tab. 2: Overview of inspected results for the Saturdays in June 2015 in NYC

2015	June 6	June 13	June 20	June 27
Boro taxi destination points (= total number of trips)	50 367	47 675	44 859	51 704
Boro taxi destination hotspots	5 006	4 919	4 769	5 233
Yellow taxi origin points (= total number of trips)	33 496	427 994	411 855	445 852
Yellow taxi origin hotspots	1 169	1 248	1 234	1 137
Number of defined (moving) transport hubs	172	178	167	179

The defined transport hubs are represented visually in Fig. 6 with an example from NYC showing an area at the Broadway. The example itself describes the dynamic size variations and spatial configurations of extracted polygons. Every defined moving transport can also be further intersected with each other for gaining typical moving transport hub patterns. In general, the example in Fig. 6 shows clearly the partially separation and aggregation of two nearby situated clusters of mobility stations: in the northwestern part and in the southeastern part of the map views. The transport hub areas on June 6 and June 27, 2015, have more the appearance of two separated polygons, whereas the transport hub area on June 13 and June 20, 2015 have respectively one more aggregated polygon.

Fig. 6: Visualization of the moving transport hub phenomena based on an example in NYC with four successive Saturday in June 2015 as time windows

Since the polygons base on the intersections of boro taxi destinations and yellow taxi origins, it is possible to imagine this example as a typical area for changing the mode of transport.

4 Outlook

First results of the presented approach show promising insights into the operational behavioral aspects of specific taxi drivers in urban environments. Since there are multiple examples that include the presence of mobility service stations and indicators, it is possible to assign moving transport hubs with different manifestations. Future work can focus on including information from traffic light signaling, together with respective average green times for every intersection, which may provide further insights into the analysis of transport hub movement. The following step would be including functional transport regions into the analysis on the level of specific bus or metro lines with the connectivity of specific stations. This would facilitate the classification into the introduced cases 1 to 3 (see Fig. 2) with knowing the accessible road segments. Furthermore, a segmentation into vehicle, bicycle and pedestrian lanes can improve the validation of extracted moving transport hubs.

One further future step consists of inspecting the correlation between the activity of appearing events that have a disproportionally large mentioning in social data extracts (e.g. from Twitter, FourSquare and Wikipedia). Knowing the typical or recurrent movement patterns of moving transport hubs can benefit the understanding of perceived attractiveness for selected demographical groups in urban environments. KELER et al. (2017) have for example defined measures of urban attractiveness for selected time windows in Augsburg, resulting in a spatially-high-resolution surface (1 m). The extension of this would be to include the dynamical aspect by generating a series of rasters as a time series. Building on this, it would be possible to not only estimate urban attractiveness, but as well to provide a connection to the usage of daily transport services in relation to the extracted moving transport hub phenomena. This might be one step in validating the proposed urban appearance, together with real world observations as for example via usage of multiple video devices.

5 References

ANKERST, M., BREUNIG, M.M., KRIEGEL, H.-P. & SANDER, J., 1999: OPTICS: Ordering Points To Identify the Clustering Structure. ACM SIGMOD international conference on Management of data, ACM press, 49-60.

CLEMENTINI, E., SHARMA, J. & EGENHOFER, M. J., 1994: Modelling topological spatial relations: Strategies for query processing. Computers & Graphics, **18**(6), 815-822.

BURGFELD, O. & LIEBHARDT, L., 2015: GI-Anwendungsprojekt - FCD-Analyse. Project report, Seminar GI-Projekt WS 2014/2015, January 27, 2015. Geoinformatics Group Augsburg, University of Augsburg.

JARVIS, R.A., 1973: On the identification of the convex hull of a finite set of points in the plane. Information Processing Letters, **2**, 18-21.

KELER, A., 2017: Ableitung von Verkehrsengpässen anhand Staudetektion mit Fahrzeugtrajektorien. Geoinformationssysteme 2017 – Beiträge zur 4. Münchner GI-Runde, Kolbe, T.H., Bill, R. & Donaubauer, A. (eds), Wichmann, Berlin/Offenbach, 172-184.

KELER, A. & KRISP, J.M., 2016: Is there a relationship between complicated crossings and frequently visited locations? – A case study with boro taxis and OSM in NYC. Proceedings of

the 13th International Symposium on Location-Based Services, Gartner, G. & Huang, H. (eds), 186-190.

KELER, A., LI, X., MAZIMPAKA, J.D., WANG, L. & KRISP, J.M., 2017: Designing measures for urban attractiveness using VGI and Open Data-a case study in Augsburg. Workshop on semantic information in geosocial data for the mobile age (SemGeoSoc'17), September 21-22, 2017, Zürich, Switzerland.

Investigations on the Potential of Binary and Multi-class Classification for Object Extraction from Airborne Laser Scanning Point Clouds

MARTIN WEINMANN[1], ROSMARIE BLOMLEY[1], MICHAEL WEINMANN[2] & BORIS JUTZI[1]

Abstract: In this paper, we address the strategies of a binary classification and a multi-class classification for the pointwise semantic labeling of airborne laser scanning data. For both strategies, we use a collection of spherical and cylindrical neighborhoods as the basis for extracting geometric multi-scale features for each point of a considered point cloud. The extracted features, in turn, are provided as input to a standard Random Forest classifier. The results achieved for multi-class classification indicate a better classification across different classes, which is important for a subsequent spatial regularization. The results achieved for binary classification addressing the detection of cars, buildings and trees, respectively, show the potential for a subsequent extraction of individual objects.

1 Introduction

Nowadays, the semantic interpretation of airborne laser scanning (ALS) point cloud data still commonly relies on a supervised classification. However, the small amount of training data which is typically available for this task not only hinders the use of modern deep learning techniques, but also the use of standard contextual classification techniques such as Conditional Random Fields (CRFs). The latter can cope with the irregular point sampling, while providing a theoretically well-founded approach for point cloud classification. However, CRFs require a sufficient number of training examples to 1) relate the data representations derived for individual 3D points and the defined classes to obtain an initial labeling and then 2) robustly infer relations among neighboring 3D points to improve the initial labeling by imposing spatial regularity. Thereby, the initial labeling is often achieved by using a classifier with probabilistic output, and this output, in turn, is used to define the unary potential also known as the association potential of a CRF. In most cases, the unary potentials are obtained via classic supervised classification techniques like Support Vector Machines or Random Forest classifiers.

When analyzing the performance of classic supervised classification techniques for point cloud classification, different influencing factors have to be taken into account. The most prominent ones are represented by the number and similarity of the defined classes, the set of involved features and the classification strategy itself. The latter becomes particularly important for applications which focus on object extraction in terms of detecting only objects corresponding to a specific class (e.g. “Car”, “Building” or “Tree”), while all remaining classes could also be merged to a common class “Background”.

[1] Karlsruhe Institute of Technology, Institute of Photogrammetry and Remote Sensing, Englerstraße 7, D-76131 Karlsruhe, E-Mail: [Martin.Weinmann, Rosmarie.Blomley, Boris.Jutzi]@kit.edu
[2] University of Bonn, Institute of Computer Science II, Friedrich-Ebert-Allee 144, D-53113 Bonn, E-Mail: mw@cs.uni-bonn.de

In this paper, we investigate the potential of different classification strategies for object extraction from ALS data. As baseline, we take into account the standard strategy of a multi-class classification which delivers a labeling with respect to multiple labels as indicated in Fig. 1. In addition, we focus on the strategy of a binary classification with respect to different objects of interest as e.g. represented by the semantic classes “Car”, “Building” and “Tree”. For each classification strategy, the classification itself is performed using a Random Forest classifier and based on a set of geometric features that has recently been proven to be appropriate for classifying ALS data (BLOMLEY & WEINMANN 2017). To demonstrate the performance of our classification approach, we perform tests on a publicly available ALS benchmark dataset for which a corresponding semantic labeling is provided.
After briefly summarizing related work (Section 2), we explain the proposed methodology for the semantic classification of ALS data in detail (Section 3). Subsequently, we present and discuss the derived experimental results with a specific focus on binary and multi-class classification of ALS data given user-defined class labels such as “Car”, “Building” and “Tree” (Section 4 and Section 5). Finally, we provide concluding remarks as well as suggestions for future work (Section 6).

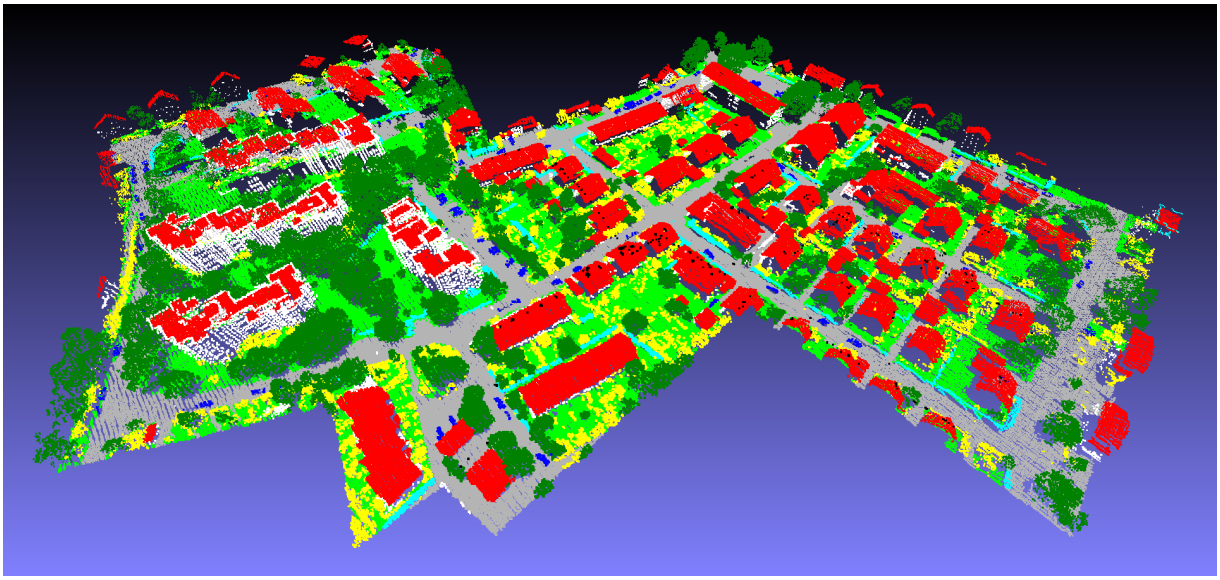


Fig. 1: ALS point cloud colored with respect to the classes “Roof” (red), “Façade” (white), “Impervious Surfaces” (gray), “Car” (blue), “Tree” (dark green), “Low Vegetation” (bright green), “Shrub” (yellow), “Fence/Hedge” (cyan) and “Powerline” (black). The point cloud comprises about 754k 3D points

2 Related Work

In this section, we summarize related work and thereby focus on the definition of appropriate local neighborhoods (Section 2.1) as the basis for extracting geometric features (Section 2.2) which, in turn, are provided as input to a classification framework (Section 2.3).

2.1 Neighborhood Definition

Commonly applied neighborhood definitions are represented by a cylindrical neighborhood (FILIN & PFEIFER 2005) or a spherical neighborhood (LEE & SCHENK 2002; LINSEN & PRAUTZSCH 2001). Such neighborhoods are parameterized by a scale parameter represented by either a radius (FILIN & PFEIFER 2005; LEE & SCHENK 2002) or the number of nearest neighbors (LINSEN & PRAUTZSCH 2001). This scale parameter is typically selected by involving knowledge about the scene and data, and an identical value is chosen for all points of the considered point cloud. Intuitively, however, a suitable value for the scale parameter might also depend on the local 3D structure and, consequently, a locally adaptive neighborhood definition seems to be more appropriate. This can for instance be achieved via dimensionality-based scale selection (DEMANTKÉ et al. 2011) favoring a highly dominant behavior of one of the dimensionality features (i.e. linearity, planarity and sphericity) or via eigenentropy-based scale selection (WEINMANN et al. 2015a) favoring the minimal disorder of 3D points. Indeed, the consideration of locally adaptive neighborhoods has been proven to be advantageous in comparison to the consideration of identically parameterized neighborhoods for all 3D points.

Instead of considering local point cloud characteristics at a single scale, it has also been proposed to consider a collection of such neighborhoods to derive a multi-scale representation for the local 3D structure. This allows a description of geometric properties at different scales and thereby implicitly accounts for the way in which these properties change across scales. Straightforward approaches for defining a multi-scale neighborhood only use differing values of the involved scale parameter. In this regard, a collection of cylindrical neighborhoods with infinite extent in the vertical direction and radii of 1m, 2m, 3m and 5m has been proposed (NIEMEYER et al. 2014) as well as a collection of spherical neighborhoods with different radii (BRODU & LAGUE 2012). Further multi-scale neighborhoods have for instance been proposed with a multi-scale voxel neighborhood (HACKEL et al. 2016) and with a collection of neighborhoods of different scale and type, e.g. in the form of voxels, blocks and pillars (HU et al. 2013), in the form of spatial bins, planar segments and local neighborhoods (GEVAERT et al. 2016) or in the form of cylindrical and spherical neighborhoods (BLOMLEY & WEINMANN 2017).

In this paper, we use a multi-scale, multi-type neighborhood composed of several cylindrical and spherical neighborhoods as this has been proven to be favorable for the classification of airborne laser scanning data compared to the use of a collection of spherical neighborhoods or a collection of cylindrical neighborhoods (BLOMLEY & WEINMANN 2017).

2.2 Feature Extraction

The spatial arrangement of 3D points within a defined local neighborhood can be encoded with a variety of geometric features. In this regard, metrical features describing local point cloud characteristics by evaluating certain geometric measures within the neighborhood are often used such as shape measures which are rather intuitive and represent one single property of the local neighborhood by a single value (WEST et al. 2004; JUTZI & GROSS 2009; MALLET et al. 2011; WEINMANN et al. 2015a). Furthermore, sampled features are often used which focus on a sampling of specific properties within the neighborhood e.g. in the form of histograms (OSADA et al. 2002; BLOMLEY et al. 2014). All extracted features are typically concatenated to a feature vector which is provided as input for classification.

In this paper, we focus on the consideration of both metrical features and sampled features to account for different geometric properties of the 3D points within the local neighborhood. In addition, we take into account that the scene is not necessarily flat and therefore use the spatial coordinates to approximate the local topography of the scene and thus derive the height above ground as additional feature for each 3D point. Instead of an accurate ground filtering of LiDAR data for automatically generating a digital terrain model (MONGUS & ZALIK 2012; SITHOLE & VOSSELMAN 2004; KRAUS & PFEIFER 1998), we assume that a coarse approximation of the local topography is already sufficient to derive normalized heights.

2.3 Classification

To classify the derived feature vectors, the concept of a supervised classification is typically applied. Given representative training data, the internal parameters of the classifier are tuned so that the classifier can afterwards well generalize to new data and assign appropriate class labels. Among the standard approaches for supervised classification, a Support Vector Machine (MALLET et al. 2011) or a Random Forest classifier (HACKEL et al. 2016; BLOMLEY & WEINMANN 2017) are often used. As such classifiers treat each 3D point individually by only considering the respective feature vector, a visualization of the derived labeling might reveal a "noisy" behavior.

To take into account that the labels of neighboring 3D points tend to be correlated and a spatial regularity of the derived labeling should be given, contextual information preserved in the feature vectors and labels of neighboring points can be involved. Such a contextual classification is often realized by using a CRF (NIEMEYER et al. 2014; WEINMANN et al. 2015b; STEINSIEK et al. 2017) or Associative and non-Associative Markov Networks (MUNOZ et al. 2009; SHAPOVALOV et al. 2010), but in general a diversity of structured regularization approaches can be used (LANDRIEU et al. 2017).

In this paper, we investigate the potential of different classification strategies for object extraction from ALS data given a small amount of training data. For this purpose, we follow the standard strategy of a multi-class classification which delivers a labeling with respect to multiple labels, and we additionally involve the strategy of a binary classification with respect to different objects of interest. For both cases, we make use of a Random Forest classifier which is a representative of modern discriminative methods.

3 Methodology

To investigate the potential of binary and multi-class classification for object extraction from ALS point clouds, we propose a framework consisting of three components. Assuming input data in the form of only spatial coordinates of 3D points without additional information, the first component exploits the given spatial information to derive suitable multi-scale, multi-type neighborhoods (Section 3.1). These multi-scale, multi-type neighborhoods represent the basis for describing the local 3D structure at different scales by extracting a set of low-level geometric features (Section 3.2). The extracted geometric features, in turn, are provided as input to a standard supervised classification approach (Section 3.3), whereby the strategies of a binary classification or a multi-class classification may be followed.

3.1 Neighborhood Definition

As the selected neighborhood definition should allow appropriately describing local point cloud characteristics, we focus on the combination of (i) a cylindrical multi-scale neighborhood, (ii) a spherical multi-scale neighborhood, (iii) a locally adaptive spherical neighborhood and (iv) spatial bins as the basis for feature extraction:

- The cylindrical multi-scale neighborhood consists of four cylindrical neighborhoods which are aligned along the vertical direction, have infinite extent in the vertical direction and are parameterized by radii of 1m, 2m, 3m and 5m, respectively (NIEMEYER et al. 2014).
- The spherical multi-scale neighborhood consists of four spherical neighborhoods which are parameterized by radii of 1m, 2m, 3m and 5m, respectively.
- The locally adaptive spherical neighborhood is derived via eigenentropy-based scale selection (WEINMANN et al. 2015a) which delivers an individual spherical neighborhood comprising the optimal number of nearest neighbors with respect to the Euclidean distance in 3D space.
- The spatial bins are derived by partitioning the scene with respect to a horizontally oriented plane into quadratic bins with a side length of 20m, and the bins in turn are only used as the basis for approximating the topography of the considered scene (BLOMLEY & WEINMANN 2017).

Thus, a collection of 10 neighborhoods is used to achieve an advanced multi-scale, multi-type neighborhood providing the basis for feature extraction.

3.2 Feature Extraction

On the basis of the defined local neighborhoods, we derive a set of low-level geometric features comprising (i) covariance features, (ii) geometric 3D properties, (iii) shape distributions and (iv) the normalized height feature:

- The covariance features are extracted from the 3D structure tensor, a 3D covariance matrix derived from the spatial coordinates of all points within the considered local neighborhood. The eigenvalues of the 3D structure tensor are normalized by their sum and then used to define the features of linearity, planarity, sphericity, omnivariance, anisotropy, eigenentropy, sum of eigenvalues and change of curvature (WEST et al. 2004; PAULY et al. 2003).
- The geometric 3D properties are derived by evaluating intuitive geometric point cloud statistics within the considered local neighborhood. In this regard, we use the features represented by the local point density, the verticality, and the maximum difference as well as the standard deviation of the height values corresponding to those points within the local neighborhood (WEINMANN et al. 2015a). For the locally adaptive spherical neighborhood derived via eigenentropy-based scale selection, we additionally consider the radius of the local neighborhood as feature.

- The shape distributions (OSADA et al. 2002) have originally been introduced to describe the shape of complete objects. More specifically, shape distributions are histograms of shape values typically derived from random point samples by applying distance metrics or angular metrics. In this regard, the angle between any three random points (A3), the distance of one random point from the centroid of all points within the neighborhood (D1), the distance between two random points (D2), the square root of the area spanned by a triangle between three random points (D3) and the cubic root of the volume spanned by a tetrahedron between four random points (D4) represent commonly used metrics. In our work, we use the adaptation of shape distributions to describe local point cloud characteristics within a selected neighborhood (BLOMLEY et al. 2014). For each of the metrics A3, D1, D2, D3 and D4, we randomly select 255 minimal point samples from the considered neighborhood, evaluate the metric for each point sample, put the resulting values in a histogram and finally consider the distribution of histogram counts. Thereby, we use histograms with 10 bins and we estimate the binning thresholds in an adaptive procedure based on 500 exemplary local neighborhoods (BLOMLEY et al. 2014).
- The normalized height feature is derived from an approximation of the scene topography as shown in Fig. 2. First, absolute height minima are determined on a large grid for which we define a sampling distance of 20m. Subsequently, a linear interpolation is performed among those coarsely gridded minimum values and evaluated on a fine grid of 0.5m sampling distance. Finally, the difference between the height value of a 3D point and the topographic height value of the corresponding grid cell is considered to derive a normalized height value for that 3D point (BLOMLEY & WEINMANN 2017).

Thus, we derive 62 features per neighborhood parameterized by a fixed radius, 63 features for a neighborhood determined via eigenentropy-based scale selection, and the normalized height feature that is additionally used.

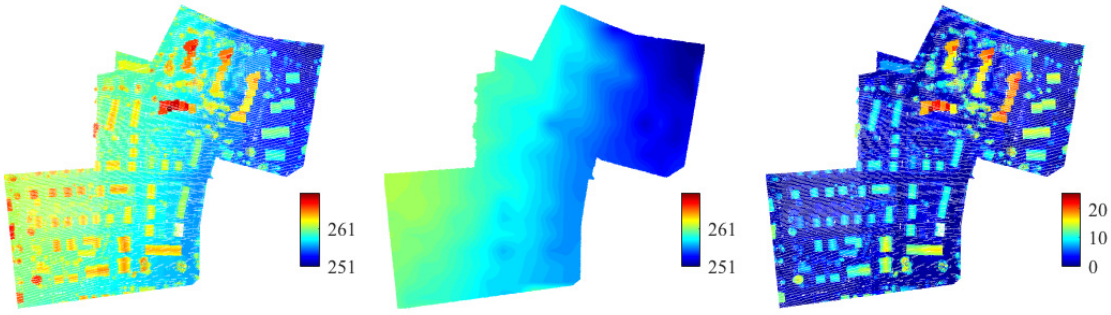


Fig. 2: Height minima on a 0.5m grid (left), approximation of the scene topography (center) and the normalized heights (right). The considered scene is the same as in Fig. 1

As the considered features address different quantities and may therefore be associated with different units as well as a different range of values which, in turn, might have a negative impact on the classification results, we introduce a normalization of the derived feature vectors. For the covariance features, the geometric 3D properties and the normalized height feature, we use a linear mapping to the interval [0,1]. Thereby, we reduce the effect of outliers by determining the

new range of the data via the 1st percentile and the 99th percentile of the training data (BLOMLEY & WEINMANN 2017). For the shape distributions, the normalization is achieved by dividing each histogram count by the total number of pulls from the local neighborhood (i.e. by 255).

3.3 Classification

For classification, we focus on a supervised classification and use a Random Forest classifier (BREIMAN 2001) which relies on the principle of ensemble learning, where the main idea consists in strategically generating a set of weak learners and combining them to achieve a strong learner. In the training stage, the weak learners are generated via bootstrap aggregating also known as "bagging" (BREIMAN 1996). More specifically, different subsets of the training data are randomly drawn with replacement so that an individual decision tree can be trained for each subset. Thereby, each decision tree is trained via the successive splitting of the considered training data to smaller subsets whose instances belong to the same class, respectively. In the prediction stage, the generated ensemble of randomly trained decision trees is used to assign a label to new feature vectors. Each of the decision trees casts a vote for one of the defined lasses, and the majority vote across all decision trees is used to estimate the respective class label.

4 Experimental Results

To demonstrate the performance of our framework, we perform tests on a benchmark dataset (Section 4.1). Involving a variety of commonly used evaluation metrics (Section 4.2), we provide a detailed quantitative assessment of the quality of the derived classification results (Section 4.3).

4.1 Dataset

For our experiments, we use the Vaihingen Dataset (CRAMER 2010; ROTTENSTEINER et al. 2012) which is provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) and available upon request. This dataset has been acquired with a Leica ALS50 system over a small German village with many detached buildings surrounded by trees and small multi-story buildings. In total, the dataset consists of about 1.166M 3D points and a split into a training scene with about 754k 3D points and a test scene with about 412k 3D points is already provided. While a reference labeling with respect to the classes "Powerline", "Low Vegetation", "Impervious Surfaces", "Car", "Fence/Hedge", "Roof", "Façade", "Shrub" and "Tree" is provided for the training scene (see Fig. 1), it is missing for the test scene so that derived classification results have to be submitted to the organizers of the ISPRS Benchmark on 3D Semantic Labeling who perform the evaluation externally.

4.2 Evaluation Metrics

For performance evaluation, we consider commonly used evaluation metrics that allow quantifying the quality of derived classification results on a per-point basis. We consider global evaluation metrics represented by the overall accuracy (OA), the kappa-index (κ) and the mean F_1-score across all classes (mF_1). Furthermore, we take into account that an imbalanced distribution of the occurrence of single classes might introduce a bias in the global evaluation

metrics and we therefore additionally consider the classwise evaluation metrics represented by recall (R), precision (P) and F_1-score (F_1), where the latter is a compound metric combining recall and precision with equal weights.

4.3 Results

First, we focus on the strategy of a multi-class classification. We train the involved Random Forest classifier and thereby take into account that an imbalanced distribution of training examples across all classes might have a detrimental effect on the training process (CRIMINISI & SHOTTON 2013). Accordingly, we introduce a class re-balancing by randomly sampling an identical number of training examples per class for the training phase. For those classes for which less training examples are available, this results in a duplication of training examples. Using 10,000 training examples per class to train the involved classifier, the classification results achieved for the test scene correspond to an overall accuracy of 68.1%, a kappa index of 60.5% and a mean F_1-score of 52.6%. To better account for the variability given in some of the classes, we also use 100,000 training examples per class to train the classifier. The respective classification results achieved for the test scene correspond to an overall accuracy of 71.5%, a kappa index of 64.3% and a mean F_1-score of 58.3%, i.e. a significant improvement of the classification results is given ($\Delta OA = 3.4\%$, $\Delta\kappa = 3.8\%$, $\Delta mF_1 = 5.7\%$). For this case, the classification results are visualized in Fig. 3 and the corresponding confusion matrix as well as more details about classwise evaluation metrics are provided in Tab. 1.

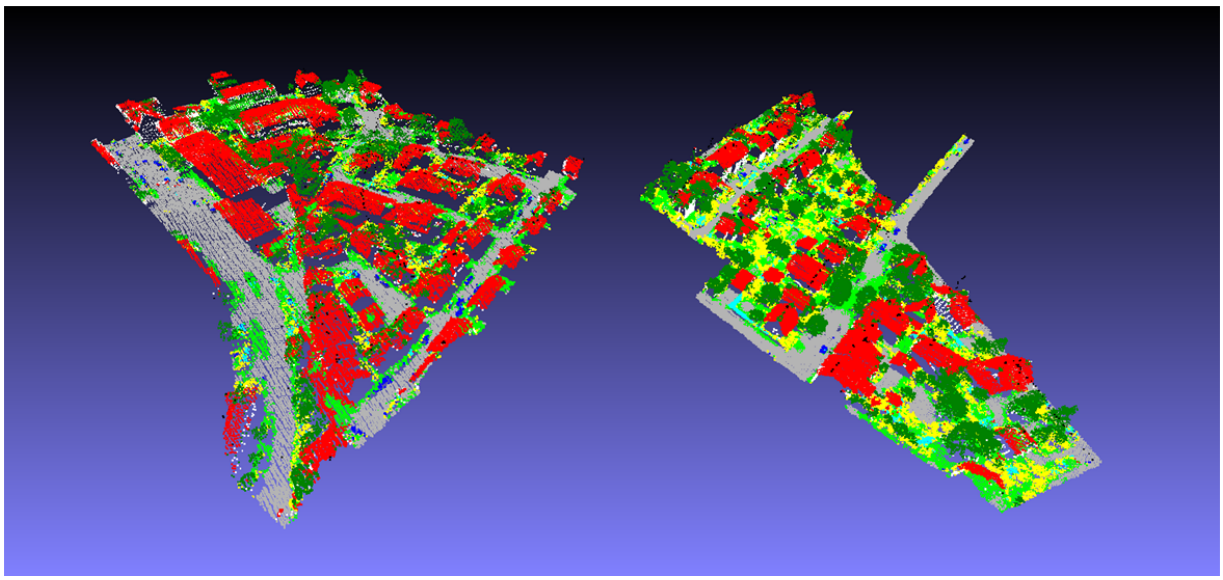


Fig. 3: Classified point cloud colored with respect to the classes “Roof” (red), “Façade” (white), “Impervious Surfaces” (gray), “Car” (blue), “Tree” (dark green), “Low Vegetation” (bright green), “Shrub” (yellow), “Fence/Hedge” (cyan) and “Powerline” (black). The point cloud comprises about 412k 3D points

Tab. 1: Confusion matrix and the classwise evaluation metrics of recall, precision and F_1-score (in %) corresponding to the result of multi-class classification (OA=71.5%, κ=64.3%, mF_1=58.3%)

Class		Estimated Label								
		Power-line	Low Veg.	Imp. Surf.	Car	Fence/ Hedge	Roof	Façade	Shrub	Tree
Reference Label	Powerline	432	1	0	0	0	44	20	2	101
	Low Vegetation	0	56068	15233	125	1692	8859	595	14128	1990
	Impervious Surfaces	1	19068	80162	146	118	1502	38	850	101
	Car	0	644	112	1082	428	121	7	1280	34
	Fence/ Hedge	0	934	48	53	1212	256	87	4059	773
	Roof	181	1534	183	7	503	90920	2164	2494	11062
	Façade	26	365	13	33	67	808	5668	1468	2776
	Shrub	7	3177	239	152	1182	549	546	15351	3615
	Tree	28	546	25	7	277	1977	1266	6575	43525
Recall		72.0	56.8	78.6	29.2	16.3	83.4	50.5	61.9	80.3
Precision		64.0	68.1	83.5	67.4	22.1	86.6	54.5	33.2	68.0
F_1-Score		67.8	61.9	81.0	40.7	18.8	84.9	52.4	43.2	73.6

Besides the strategy of a multi-class classification, we also focus on the strategy of a binary classification with respect to the classes “Car”, “Building” and “Tree”, respectively. While the classes “Car” and “Tree” are coincident with their counterparts in the provided reference labeling, the class “Building” is composed of the classes “Roof” and “Façade”. Again, we randomly sample an identical number of 100,000 training examples per class to train the classifier. For the binary classification addressing “Car” and “Background”, we achieve an overall accuracy of 99.0%, a kappa index of 43.7% and a mean F_1-score of 71.9%. For the binary classification addressing “Building” and “Background”, we achieve an overall accuracy of 91.4%, a kappa index of 79.5% and a mean F_1-score of 89.8%. For the binary classification addressing “Tree” and “Background”, we achieve an overall accuracy of 89.0%, a kappa index of 62.6% and a mean F_1-score of 81.1%. The classwise evaluation metrics for all three cases are provided in Tab. 2. A comparison to Tab. 1 reveals that the binary classification leads to a better detection of cars and buildings, while the detection of trees is worse in comparison to the results achieved with the strategy of a multi-class classification.

Tab. 2: Classwise evaluation metrics of recall, precision and F_1-score (in %) corresponding to the results of a binary classification with respect to “Car” (left), “Building” (center) and “Tree” (right).

	Car	Background
Recall	44.1	99.5
Precision	44.3	99.5
F_1-Score	44.2	99.5

	Building	Background
Recall	87.7	92.9
Precision	83.7	94.8
F_1-Score	85.7	93.9

	Tree	Background
Recall	91.6	88.7
Precision	55.1	98.6
F_1-Score	68.8	93.9

The derived results for multi-class classification and binary classification are visualized in Fig. 4 for the test scene with about 412k 3D points.

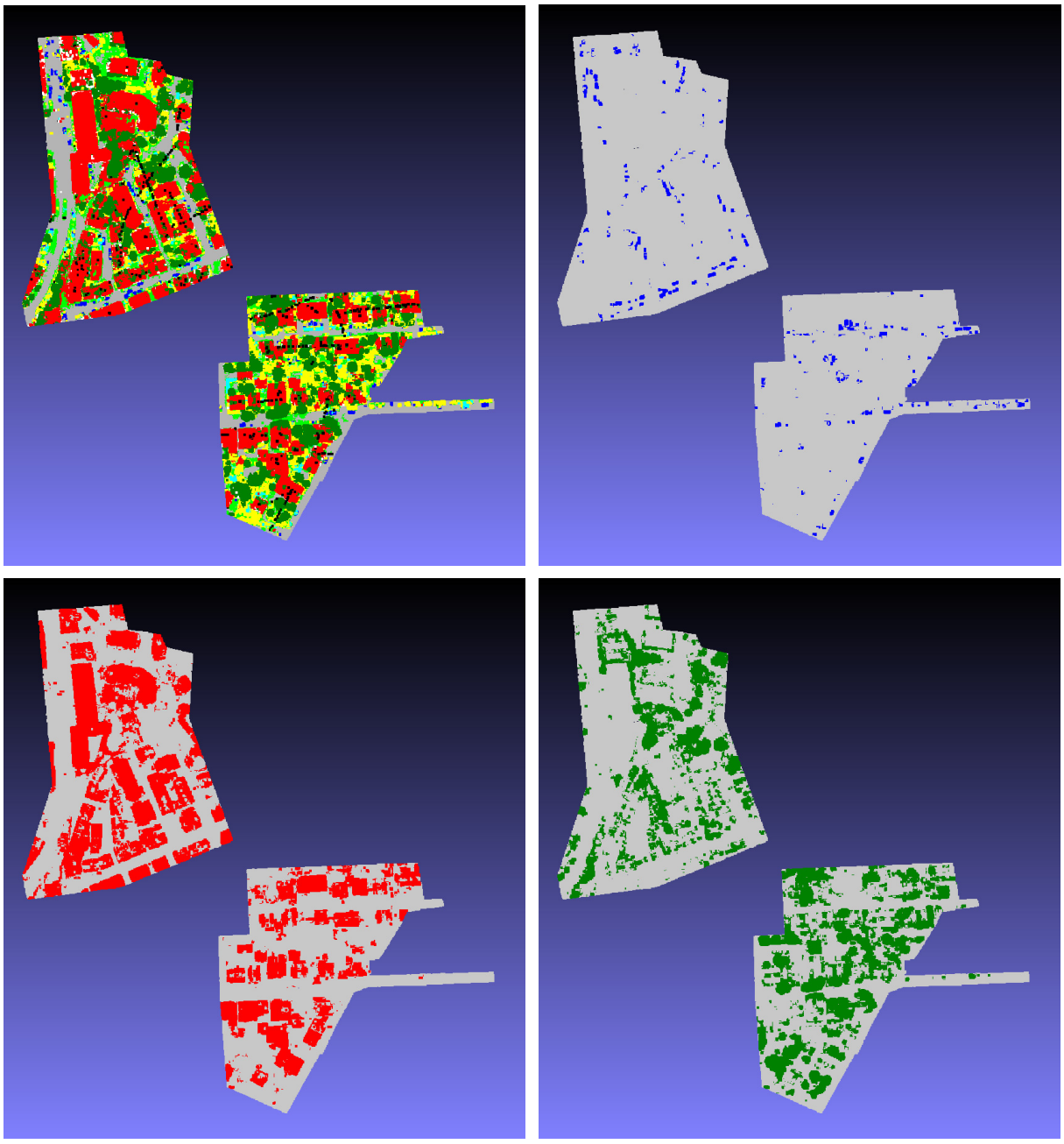


Fig. 4: Classified point clouds: the top left part shows the result of a multi-class classification with respect to “Roof” (red), “Façade” (white), “Impervious Surfaces” (gray), “Car” (blue), “Tree” (dark green), “Low Vegetation” (bright green), “Shrub” (yellow), “Fence/Hedge” (cyan) and “Powerline” (black); the top right part shows the result of a binary classification with respect to “Car” (blue) and “Background” (gray); the bottom left part shows the result of a binary classification with respect to “Building” (red) and “Background” (gray); the bottom right part shows the result of a binary classification with respect to “Tree” (green) and “Background” (gray).

5 Discussion

The derived results for multi-class classification indicate that the Vaihingen Dataset with the nine defined classes is rather challenging for the proposed methodology, as the achieved overall accuracy is only about 71.5%. The reason for this is that only geometric features are provided as input to the classifier and that several of the defined classes are characterized by a high geometric similarity. Particularly the classes "Low Vegetation" and "Impervious Surfaces" as well as the classes "Shrub" and "Tree" exhibit a similar geometric behavior and therefore misclassifications among these classes occur quite often (Tab. 1). This is in accordance with other recent investigations involving the same dataset (STEINSIEK et al. 2017; BLOMLEY & WEINMANN 2017). A more detailed comparison to the results achieved with other classification approaches is provided in Tab. 3 and reveals that the proposed methodology involving a Random Forest for classification outperforms other approaches using the same classifier. The significant improvement across several classes thereby comes at the cost of a loss with respect to the predominant classes represented by "Low Vegetation" and "Impervious Surfaces". The mean F_1-score is however 8.3% higher than in recent investigations (STEINSIEK et al. 2017), which indicates a better classification of the different classes. Consequently, interpreting the votes of the involved Random Forest classifier in a probabilistic way would provide a better association potential for a CRF imposing spatial regularity on the derived classification results which, in turn, is likely to allow the CRF to further increase the quality of the classification results.

Tab. 3: Classwise F_1-scores as well as the overall accuracy (OA) and the mean F1-score across all classes (mF_1) (in %) for different approaches (RF^1: classification based on a Random Forest (STEINSIEK et al. 2017); CRF^1: classification based on a CRF approach (STEINSIEK et al. 2017); RF^2: classification based on a Random Forest (BLOMLEY & WEINMANN 2017); RF^3: classification based on the proposed approach).

	F_1									OA	mF_1
	Power-line	Low Veg.	Imp. Surf.	Car	Fence/ Hedge	Roof	Façade	Shrub	Tree		
RF^1	14.3	65.8	86.1	24.9	19.8	84.8	43.9	40.8	69.5	71.0	50.0
CRF^1	69.8	73.8	91.5	58.2	29.9	91.6	54.7	47.8	80.2	80.5	66.4
RF^2	32.1	57.9	80.0	44.1	17.5	81.8	47.5	41.5	70.9	68.1	52.6
RF^3	67.8	61.9	81.0	40.7	18.8	84.9	52.4	43.2	73.6	71.5	58.3

Furthermore, the derived results reveal an improvement for the classes "Car" and "Building" if a binary classification is considered instead of a multi-class classification (Tab. 1 and Tab. 2). However, the binary classification delivers worse results for the class "Tree" which might be due to shrubs revealing geometrically similar characteristics. Indeed, there are many misclassifications among the classes "Shrub" and "Tree" in the case of a multi-class classification (Tab. 1). As the elements of the class "Shrub" are added to the class "Background" for binary classification, the relevant classes "Tree" and "Background" contain more similar instances which, in turn, results in lower detection rates. A consideration of the results for binary classification also reveals that the class "Building" can easily be identified, while the class "Car" seems to be more challenging (Tab. 2). The latter is due to the similar geometric behavior of the

classes “Car”, “Fence/Hedge” and “Shrub” for the given point density, which results in many misclassifications between these classes in the case of a multi-class classification (Tab. 1). Note that, for a binary classification, the overall accuracy does not necessarily provide an appropriate conclusion about the objects of interest. For the classification with respect to the classes “Car” and “Background”, the overall accuracy is 99.0%, although the detection rates are rather low for the class “Car”. This is due to the small amount of 3D points that should be labeled as “Car” in comparison to a huge amount of 3D points that should be labeled as “Background”. Hence, even classifying all 3D points as “Background” would still result in a high overall accuracy, while the kappa-index allows reasoning about the class separability and hence would indicate a poor classification result. For this example, the mean F_1-score across all classes would also indicate a poor classification result as it accounts for the appropriate detection of both defined classes.

6 Conclusions

In this paper, we have investigated the potential of binary and multi-class classification as the basis for object extraction from ALS data. We have focused on the use of a multi-scale, multi-type neighborhood composed of several cylindrical and spherical neighborhoods as the basis for extracting a variety of geometric features. The extracted features, in turn, have been provided as input for classification, whereby the focus has been put on the strategies of binary and multi-class classification based on a Random Forest classifier. The classification results derived for a benchmark dataset have clearly revealed the potential of the proposed methodology for pointwise classification. Regarding the strategy of a multi-class classification, the achieved results indicate a significantly better classification across all classes in comparison to other similar approaches which tend to better classify the dominant classes. Regarding the strategy of a binary classification, we have addressed the detection of cars, buildings and trees, respectively. The achieved classification results provide a good initial labeling for a subsequent extraction and counting of single objects in the scene, yet additional effort is required to account for misclassifications and only retain segments which correspond to real objects in the scene.
In future work, we intend to improve the classification results by first using the proposed methodology to achieve an initial labeling and then imposing spatial regularity by using spatial regularization techniques (NIEMEYER et al. 2014; STEINSIEK et al. 2017; LANDRIEU et al. 2017). Thereby, the proposed methodology allows for a better classification of different classes and thus a better initial labeling which serves as input to the CRF via the association potentials and this, in turn, is likely to allow the CRF to further increase the quality of the classification results compared to other approaches. Furthermore, we plan to address the step towards the extraction of individual objects in the scene.

7 Acknowledgements

The Vaihingen Dataset was kindly provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (CRAMER 2010): http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html .

8 References

BLOMLEY, R., WEINMANN, M., LEITLOFF, J. & JUTZI, B., 2014: Shape distribution features for point cloud analysis – A geometric histogram approach on multiple scales. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **II**(3), 9-16.

BLOMLEY, R. & WEINMANN, M., 2017: Using multi-scale features for the 3D semantic labeling of airborne laser scanning data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **IV**(2/W4), 43-50.

BREIMAN, L., 1996: Bagging predictors. Machine Learning, **24**(2), 123-140.

BREIMAN, L., 2001: Random forests. Machine Learning, **45**(1), 5-32.

BRODU, N. & LAGUE, D., 2012: 3D terrestrial LiDAR data classification of complex natural scenes using a multi-scale dimensionality criterion: applications in geomorphology. ISPRS Journal of Photogrammetry and Remote Sensing, **68**, 121-134.

CRAMER, M., 2010: The DGPF-test on digital airborne camera evaluation – Overview and test design. PFG Photogrammetrie – Fernerkundung – Geoinformation, **2/2010**, 73-82.

CRIMINISI, A. & SHOTTON, J., 2013: Decision forests for computer vision and medical image analysis. Springer, London, UK.

DEMANTKÉ, J., MALLET, C., DAVID, N. & VALLET, B., 2011: Dimensionality based scale selection in 3D LiDAR point clouds. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(5/W12), 97-102.

FILIN, S. & PFEIFER, N., 2005: Neighborhood systems for airborne laser data. Photogrammetric Engineering & Remote Sensing, **71**(6), 743-755.

GEVAERT, C.M., PERSELLO, C. & VOSSELMAN, G., 2016: Optimizing multiple kernel learning for the classification of UAV data. Remote Sensing, **8**(12), 1-22.

HACKEL, T., WEGNER, J.D. & SCHINDLER, K., 2016: Fast semantic segmentation of 3D point clouds with strongly varying density. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **III**(3), 177-184.

HU, H., MUNOZ, D., BAGNELL, J. A. & HEBERT, M., 2013: Efficient 3-D scene analysis from streaming data. Proceedings of the IEEE International Conference on Robotics and Automation, 2297-2304.

JUTZI, B. & GROSS, H., 2009: Nearest neighbour classification on laser point clouds to gain object structures from buildings. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(1-4-7/W5), 1-6.

KRAUS, K. & PFEIFER, N., 1998: Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing, **53**(4), 193-203.

LANDRIEU, L., RAGUET, H., VALLET, B., MALLET, C. & WEINMANN, M., 2017: A structured regularization framework for spatially smoothing semantic labelings of 3D point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, **132**, 102-118.

LEE, I. & SCHENK, T., 2002: Perceptual organization of 3D surface points. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **34**(3A), 193-198.

LINSEN, L. & PRAUTZSCH, H., 2005: Local versus global triangulations. Proceedings of Eurographics, 257-263.

MALLET, C., BRETAR, F., ROUX, M., SOERGEL, U. & HEIPKE, C., 2011: Relevance assessment of full-waveform LiDAR data for urban area classification. ISPRS Journal of Photogrammetry and Remote Sensing, **66**(6), S71-S84.

MONGUS, D. & ZALIK, B., 2012: Parameter-free ground filtering of LiDAR data for automatic DTM generation. ISPRS Journal of Photogrammetry and Remote Sensing, **67**, 1-12.

MUNOZ, D., BAGNELL, J. A., VANDAPEL, N. & HEBERT, M., 2009: Contextual classification with functional max-margin Markov networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 975-982.

NIEMEYER, J., ROTTENSTEINER, F. & SOERGEL, U., 2014: Contextual classification of LiDAR data and building object detection in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing, **87**, 152-165.

OSADA, R., FUNKHOUSER, T., CHAZELLE, B. & DOBKIN, D., 2002: Shape distributions. ACM Transactions on Graphics, **21**(4), 807-832.

PAULY, M., KEISER, R. & GROSS, M., 2003: Multi-scale feature extraction on point-sampled surfaces. Computer Graphics Forum, **22**(3), 81-89.

ROTTENSTEINER, F., SOHN, G., JUNG, J., GERKE, M., BAILLARD, C., BENITEZ, S. & BREITKOPF, U., 2012: The ISPRS benchmark on urban object classification and 3D building reconstruction. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **I**(3), 293-298.

SHAPOVALOV, R., VELIZHEV, A. & BARINOVA, O., 2010: Non-associative Markov networks for 3D point cloud classification. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(3A), 103-108.

SITHOLE, G. & VOSSELMAN, G., 2004: Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, **59**(1-2), 85-101.

STEINSIEK, M., POLEWSKI, P., YAO, W. & KRZYSTEK, P., 2017: Semantische Analyse von ALS- und MLS-Daten in urbanen Gebieten mittels Conditional Random Fields. Tagungsband der 37. Wissenschaftlich-Technischen Jahrestagung der DGPF, 521-531.

WEINMANN, M., JUTZI, B., HINZ, S. & MALLET, C., 2015a: Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. ISPRS Journal of Photogrammetry and Remote Sensing, **105**, 286-304.

WEINMANN, M., SCHMIDT, A., MALLET, C., HINZ, S., ROTTENSTEINER, F. & JUTZI, B., 2015b: Contextual classification of point cloud data by exploiting individual 3D neighborhoods. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **II** (3/W4), 271-278.

WEST, K.F., WEBB, B.N., LERSCH, J.R., POTHIER, S., TRISCARI, J.M. & IVERSON, A.E., 2004: Context-driven automated target detection in 3-D data. Proceedings of SPIE, **5426**, 133-143.

Dense Matching mit WorldView-4 und Kompsat-3 Bildern

KARSTEN JACOBSEN [1] & UMUT SEFERCIK[2]

Zusammenfassung: Mit WorldView-4 steht ein zweiter zivil verfügbarer optischer Satellit mit einer Bodenauflösung von 31cm zur Verfügung, der eine Konkurrenz für Luftbilder darstellt. Im Bereich von Istanbul wurde neben einem WorldView-4 Stereomodell auch ein Kompsat-3 Stereomodell, mit 70cm Objektpixelgröße, mit flächenbasierter Bildzuordnung nach der Methode der kleinsten Quadrate und Semi Global Matching (SGM) ausgewertet. Die Auswertung der Kompsat-3 Bilder ist von der Methode und der Art der vorhandenen Bilder abhängig.

Die WorldView-4 Aufnahmen sind, entsprechend der von GeoEye bevorzugten Bildgeometrie, auf eine Objektebene konstanter Höhe projiziert und damit geometrisch unproblematisch auszuwerten. Die Bildzuordnung nach der Methode der Kleinsten Quadrate führt zu den bekannten Verwischungen der Höhenmodelle. SGM hat zwar die gleichen Probleme mit Verdeckungen in den engen Straßenschluchten, gibt die Dachformen aber entsprechend der Objektpixelgröße gut wieder, soweit die Dächer nicht durch Bäume überragt werden. Von Kompsat-3 wurden L1R und L1G-Aufnahmen ausgewertet – L1G ist auf das SRTM-DSM bezogen, dessen Berücksichtigung bei der Auswertung nicht vollständig rückgängig gemacht werden konnte.

1 Einleitung

Höhenmodelle, erstellt mit Satellitenbildern, stellen eine ökonomische Lösung dar, die entsprechend der geforderten Genauigkeiten, mit der basierend auf Luftbildern oder Laserscanning konkurrieren kann. Mit einer Objektauflösung von 31cm liegen die WorldView-4-Aufnahmen im Bereich der Luftbilder und können auch aufgrund der Bildqualität damit verglichen werden. Kompsat-3 Aufnahmen haben dagegen eine Objektauflösung von 71cm und damit kann nicht die gleiche Genauigkeit erreicht werden, die linear von der Objektpixelgröße abhängig ist. Mit beiden Weltraumsystemen wurde das Hauptsiedlungsgebiet der Stadt Istanbul aufgenommen für das ein Laserscanning-Höhenmodell (ALS) mit 4m Punktabstand als Referenz vorliegt.

2 Problemstellung

Höhenmodelle sind ein grundlegender Bestandteil jedes Geoinformationssystems. Für einige Anwendungen, wie z.B. Unterstützungen im Krisenfall, sind aktuelle Höhenmodelle erforderlich, für die sich Satellitenstereopaare anbieten, da sie eine geringere Vorlaufzeit als Luftbildflüge benötigen. Die Erstellung von Höhenmodellen aus optischen Satellitenbildern ist kein neues Problem und auch die dichte Punktbestimmung mittels Semi Globaler Bildzuordnung (SGM) ist nicht neu (u.a. HIRSCHMÜLLER 2005, HIRSCHMÜLLER 2008, HEINRICHS et al. 2007, ALOBEID et al.

[1] Leibniz Universität Hannover, Nienburger Str. 1, D-30167 Hannover,
E-Mail: jacobsen@ipi.uni-hannover.de
[2] Bülent Ecevit University, Zonguldak, Türkei; E-Mail: UGSefercik@hotmail.com

2010, HAALA 2011). Jedes Projekt hat aber spezielle Herausforderungen, wie die dichte Bebauung und die engen Straßenschluchten von Istanbul, die zu Verdeckungen und Rauschen führen, und die verfügbaren Bildprodukte L1R und L1G der Kompsat-3 Aufnahmen vom März 2017. Die Auswertung von L1G-Aufnahmen wurde bislang nicht publiziert. Die koreanische Weltraumbehörde KARI vertreibt die Satellitenbilder mit beiden Bildarten. L1R-Bilder entsprechen den geometrisch korrigierten und zusammengesetzten originalen Satellitenbildern (Basic Imagery), während L1G-Bilder auf das SRTM-Höhenmodell mit 3 Bogensekunden Punktabstand projiziert und damit genäherte Orthobilder sind. Die für SGM erforderlichen Epipolarbilder können durch ebene Rotation der L1G-Produkte auf die Basisrichtung des Stereopaares erstellt werden, während L1R-Bilder vorher auf eine Ebene konstanter Höhe, sinnvollerweise im Landeskoordinatensystem, zu projizieren sind. Die mittels L1G erstellten Höhenmodelle führen zu Höhenunterschiede gegenüber dem SRTM-Referenzhöhenmodell. Bei entsprechenden QuickBird-Standard-Aufnahmen, die ebenfalls auf das SRTM Höhenmodell projiziert sind, traten keinerlei Probleme mit diesem Bildtyp auf und es wurden ähnliche Genauigkeiten wie mit QuickBird-Basic-Imagery, die L1R entsprechen, erzielt (JACOBSEN 2006). KARI teilte jedoch mit, dass bei der Definition der SRTM-Höhenmodelle Probleme vorhanden waren. Verschiebungen des SRTM- gegenüber dem ALS Referenzhöhenmodell wurden durch Ausgleichung bestimmt und berücksichtigt.

3 Bildorientierung

3.1 Kompsat-3

Wie oben erwähnt, liegt das Kompsat-3 Stereomodell als L1G (genäherte Orthobilder bezogen auf das SRTM 3-Bogensekunden-Oberflächenmodell) und als L1R (zusammengesetzte und um Kalibrierung verbesserte Originalbilder) vor. Epipolarbilder, die für SGM erforderlich sind, können aus L1G-Bildern durch horizontale Drehung auf die Bildbasis erzeugt werden, während L1R-Bilder vor der Drehung auf eine Ebene konstanter Höhe zu projizieren sind, was Orthobildern mit einer konstanter Höhe entspricht. Die Orientierung der L1G- und der L1R-Bilder erfolgt üblicherweise mittels bias-korrigierter rationaler Polynomkoeffizienten. In den Kompsat-3 Bildern wurden 71 Passpunkte bestimmt.

Die Orientierungsgenauigkeiten in den Tabellen 1 und 2 zeigen, dass die Orientierung der L1G-Bilder (projiziert auf das SRTM 3-Bogensekunden-DSM) nicht so gute Ergebnisse liefern, wie die Orientierung der Basic Imagery L1R. Die schlechtere Genauigkeit in Blickrichtung deutet auf Probleme in der Referenzhöhe hin, die sich quer zur Blickrichtung nicht auswirken. Die Auswirkung der Höhenkorrektur durch das SRTM-DSM konnte somit bei der Orientierung nicht vollständig berücksichtigt werden. Das SRTM 3-Bogensekunden-Höhenmodell wurde vor der Orientierung durch Ausgleichung auf das Laserscanning-Referenz-DSM verschoben, ohne diese Verschiebung waren die Ergebnisse der L1R-Orientierungen schlechter. Die Ergebnisse der L1R-Bilder (Basic Imagery) sind deutlich besser und stellen in Relation zu 0.71m Objektpixelgröße (GSD) ein gutes Ergebnis dar. Für das Basis zu Höhenverhältnis von 1:1,23 ist auch die Höhengenauigkeit der 3D-Orientierung mit SZ=1,09m zufriedenstellend, während dieses bei der L1G-Lösung mit SZ=2.07m nicht der Fall ist.

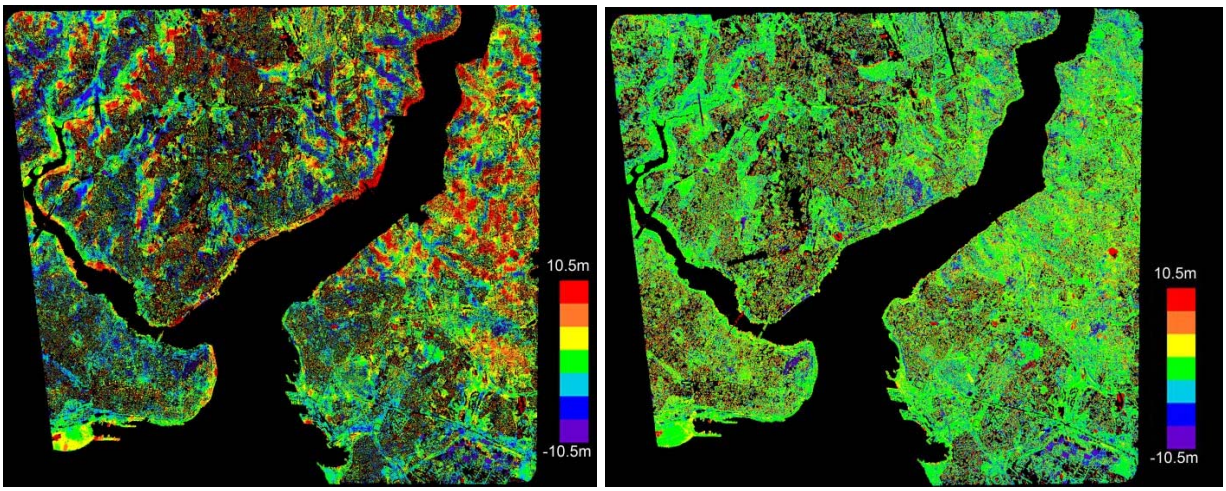
Die Orientierungsprobleme der L1G-Bilder wurden deswegen durch eine flächenbasierte Bildzuordnung nach der Methode der kleinsten Quadrate, untersucht. Als Referenz diente das Laserscanning-Oberflächenmodell mit 4m Punktabstand. Dieser Punktabstand führt bei der dichten Bebauung in Istanbul zu nicht unerheblichen Höhendifferenzen an den Gebäudeumringen und beeinflusst die errechneten Höhengenauigkeiten nicht unerheblich.

Tab. 1: Genauigkeit der Kompsat-3 L1G Bildorientierung – basierend auf 71 Passpunkten

Einzelbildorientierung L1G				
	SX	SY	In Blickrichtung	Quer dazu
Bild 5935	0.95 m	0.97 m	1.03 m	0.88m
Bild 5820	0.84 m	1.24 m	1.29 m	0.78 m
Dreidimensionale Orientierung				
	SX	SY	SZ	
Stereomodell	0.82 m	0.90 m	2.07 m	

Tab. 2: Genauigkeit der Kompsat-3 L1G Bildorientierung – basierend auf 71 Passpunkten

Einzelbildorientierung L1R (Basic Imagery)				
	SX	SY	In Blickrichtung	Quer dazu
Bild 5935	0.68 m	0.67 m	0.68 m	0.67 m
Bild 5820	0.45 m	0.40 m	0.44 m	0.41 m
Dreidimensionale Orientierung				
	SX	SY	SZ	
Stereomodell	0.53 m	0.55 m	1.09 m	



Höhendifferenzen L1G Höhendifferenzen L1R

Abb. 1: Farbkodierte Höhendifferenzen der Kompsat-3 Höhenmodelle gegenüber Laserscanning

Die in Abbildung 1 dargestellten Höhenunterschiede zwischen den Kompsat-3 Höhenmodellen und der Referenz verdeutlichen die Probleme der erforderlichen Höhenkorrektur des L1G-Höhenmodells durch das SRTM 3-Bogensekunden Oberflächenmodells. Die in Abbildung 1,

links, deutlichen lokalen Höhendifferenzen in den steileren Lagen bestätigen die von KARI mitgeteilten Probleme mit dem SRTM-Referenzhöhenmodell. Die etwas dunkleren Bereiche in Abbildung 1, rechts, sind Punktausfälle durch Verdeckungen in den dicht bebauten Gebieten. Die lokalen systematischen Höhendifferenzen sind nicht durch Verschiebungen der Höhenmodelle gegeneinander zu erklären – diese wurden untersucht und berücksichtigt. Damit sind die Kompsat-3 L1G-Bilder nicht für dreidimensionale Auswertungen zu empfehlen, auch ihre Lagegenauigkeit ist problematisch. Dieses gilt ausdrücklich nicht für die L1R-Bilder, die unproblematisch auszuwerten sind.

3.2 WorldView-4

WorldView-4 Stereobilder werden mit der Geometrie der Standard Imagery vertrieben, die einer Projektion der zusammengesetzten und um die Kalibrierung verbesserten Originalbilder auf eine Ebene konstanter Höhe im Landeskoordinatensystem entsprechen. Üblicherweise wird die Orientierung mittels Rationaler Polynomkoeffizienten (RPC) mit Biaskorrektur durchgeführt. Die verwendeten Bilder vom Juni 2017 haben jedoch eine Größe von 25600×68616 Pixeln, was an die Genauigkeitsgrenze der RPC geht. Für die volle Ausnutzung der Bildgenauigkeit sind Passpunkte erforderlich. Es sind 7 klar identifizierbare Passpunkte mit ausreichender Genauigkeit vorhanden. Eine Biaskorrektur nur mit einer ebenen Affintransformation führte nicht zu Subpixelgenauigkeit, es waren zusätzlich zwei Parameter erforderlich, die die Blickrichtung verbessern. Von den sechs Affinparametern der Biaskorrektur waren allerdings nur zwei beziehungsweise drei Parameter erforderlich, womit letztendlich insgesamt 4 beziehungsweise 5 Bias-Parameter für die Bildorientierung benutzt wurden. Damit wurde neben der Lagegenauigkeit von 0,4 GSD auch eine Höhengenauigkeit an den Passpunkten im Subpixelbereich von 0.9 GSD erreicht.

Tab. 3: Genauigkeit der WorldView-4 Bildorientierung

Einzelbildorientierung			
	Unbekannte	SX	SY
Bild 5615	5	0.17 m	0.09 m
Bild 5723	4	0.22 m	0.25 m
Dreidimensionale Orientierung			
	SX	SY	SZ
Stereomodell	0.15 m	0.07 m	0.27 m

4 Untersuchung der erzeugten Höhenmodelle

4.1 Kompsat-3

Es wurde sowohl ein Höhenmodell mittels flächenbasierter Bildzuordnung nach der Methode der kleinsten Quadrate, als auch mittels SGM erstellt. Die in Abbildung 1 dargestellten Höhendifferenzen basieren auf der flächenbasierten Zuordnung, die mit Ausnahme der Details an Höhensprüngen, wie Gebäuden, Brücken und Straßendämmen, ein zufriedenstellendes Ergebnis liefern. Die in Abbildung 1, rechts, dargestellten farbkodierten Höhendifferenzen gegenüber dem La-

serscanning Oberflächenmodell aus dem Jahr 2015 zeigen einzelne rote Flecken, die auf Höhendifferenzen von etwa 10m hinweisen. Diese großen Differenzen sind auf Änderungen zwischen dem Zeitpunkt Laserscanaufnahme im Jahre 2015 und der Satellitenbilder im Jahr 2017, wie durch den Bau einer Brücke über das Goldene Horn, eine neue Stadionüberdachung und neue Gebäude zurückzuführen sind. Eine Auswirkung haben auch die Differenzen der Oberflächendefinition in Parkanlagen durch Laserscanner und durch Bildzuordnung von optischen Bildern, die in Abbildung 1 blau erscheinen.

Tab. 4: Vergleich des Kompsat-3 Höhenmodells basierend auf flächenhafter Bildzuordnung mit der Laserscanner-Referenz

L1R	SZ	NMAD	Relatives SZ
Alle Punkte	2.90 m	2.34 m	1.99 m
Für Neigungen < 10%	1.94 m	1.20 m	
Funktion der Neigung	SZ=2.35m + 1.07m * tan a	NMAD=1.82m + 1.55m * tan a	
L1G	SZ	NMAD	Relatives SZ
Alle Punkte	5.40 m	4.80 m	3.60 m
Für Neigungen < 10%	4.45 m	3.60 m	
Funktion der Neigung	SZ=4.82m + 0.99m * tan a	NMAD=4.19m + 1.39m * tan a	

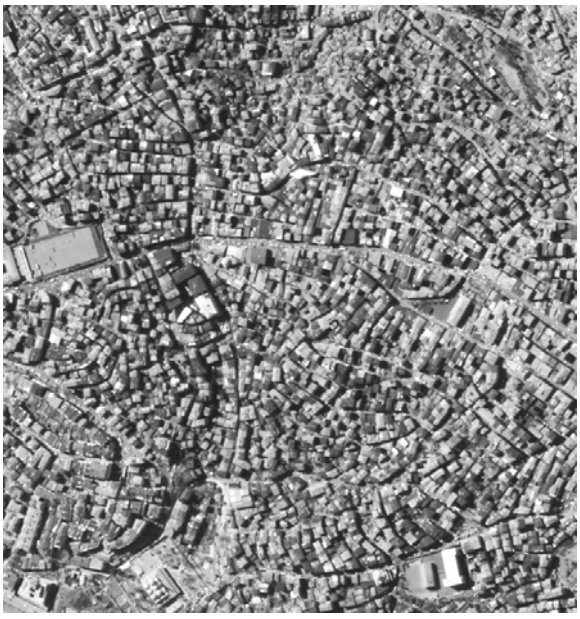


Abb.2: Kompsat-3-Bildausschnitt

Abb. 3: Kompsat-3 Parallaxenbild

Auch der Vergleich der durch flächenhafte Bildzuordnung erstellten Kompsat-3 Höhenmodelle mit Laserscandaten (Tabelle 4) zeigt die Probleme der L1G-Daten, während das Höhenmodell basierend auf L1R-Bildern innerhalb der Erwartung ist. Ein NMAD (Normalized Median Absolute Deviation (HÖHLE & HÖHLE 2009)) von 1,21m für Geländeneigungen < 10% stellt für diese Methode in dem dicht bebauten Gebiet ein zufriedenstellendes Ergebnis dar. Dass die Stan-

dardabweichung mit 1.94m deutlich höher liegt, ist auf den Einfluss von Höhensprüngen an den Gebäudekanten zurückzuführen, die durch die flächenbasierte Bildzuordnung verwischt werden.

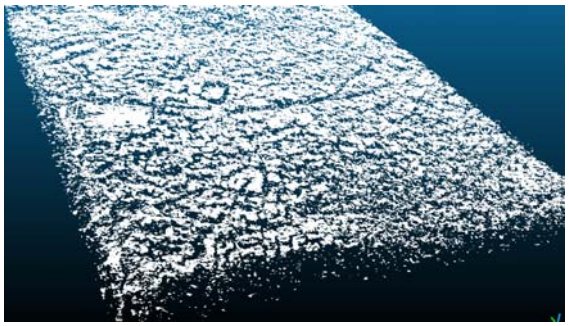


Abb. 4: Originale Punktwolke

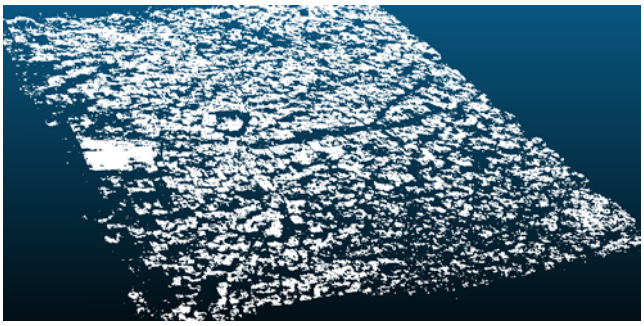


Abb. 5: gefilterte Punktwolke

Die Bestimmung der dreidimensionalen Punktwolke durch SGM (Abb. 3-Abb. 5) ist durch die Verdeckungen in den engen Straßenschluchten verrauscht (Abb. 4) und wurde durch einen dreidimensionalen Nachbarschaftsfilter verbessert (Abb. 5). Entsprechend der Objektpixelgröße von 0.71m sind die Dachformen zufriedenstellend erfasst.

Tab. 5: Vergleich des Kompsat-3 Höhenmodells basierend auf SGM mit der Laserscanner-Referenz

L1R	SZ	NMAD	Relatives SZ
Alle Punkte	3.03 m	2.16 m	2.16 m
Für Neigungen < 10%	1.97 m	1.21 m	
Funktion der Neigung	SZ=2.44m + 0.82m * tan a	NMAD=1.74m + 1.31m * tan a	

Die Genauigkeitsmaße der Kompsat-3 L1R Höhenmodelle basierend auf flächenhafter Bildzuordnung und auf SGM (Tabellen 4 und 5), verglichen mit dem Laserscanning-Oberflächenmodell mit 4m Punktabstand, unterscheiden sich nicht wesentlich. Durch das Laserscanning Oberflächenmodell mit 4m Punktabstand kann die dicht bebaute städtische Landschaft (Abb. 2 und 9) nicht ausreichend beschrieben werden. Lediglich das NMAD in Bereichen mit Neigungen unter 10% gibt eine einigermaßen realistische Genauigkeitsangabe, wobei allerdings zu beachten ist, dass der seitliche Gebäudeabstand in dem Untersuchungsgebiet deutlich unter 4m liegt (Abb. 9). Dieses zeigt sich auch bei der Standardabweichung, die deutlich über dem NMAD liegt – eine größere Anzahl von größeren Höhendifferenzen als der Normalverteilung entsprechend, erhöht die Standardabweichung deutlich, während das bei dem auf dem Median basierenden NMAD nicht der Fall ist.
Die ähnlichen Genauigkeiten der flächenhaften Bildzuordnung durch Kleinste Quadrate und von SGM wurden bereits in ALOBEID et al. (2010) festgestellt. Der wesentliche Vorteil von SGM liegt in der korrekten Erfassung von Höhensprüngen, was sich bei der guten Darstellung der Dachflächen und -formen (Abb. 5) zeigt.

4.2 WorldView-4

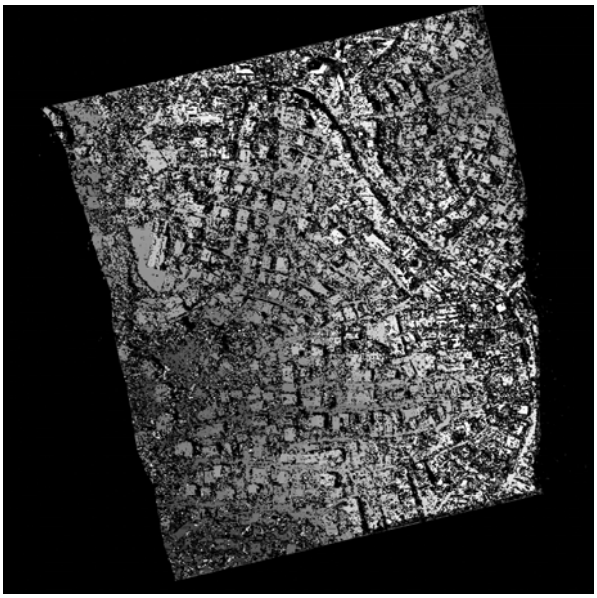


Abb. 6: Ausschnitt aus dem Parallaxenbild SGM Bildzuordnung

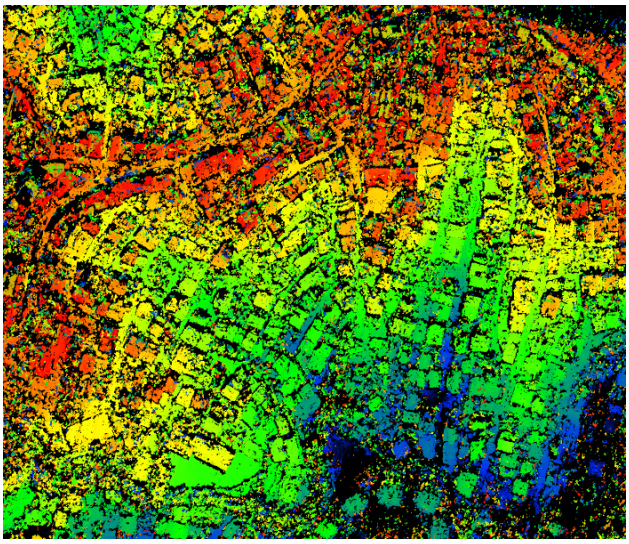


Abb. 7: Farbcodiertes SGM-Oberflächenmodell

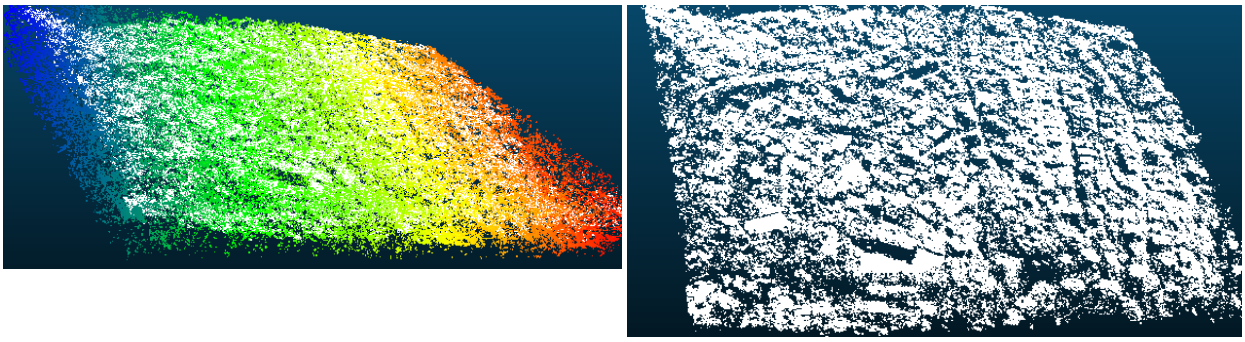


Abb. 8: Punktwolke des Oberflächenmodells erzeugt mit Semi Global Matching. Ausschnitt aus originalem SGM-DHM mit Rauschen (links) und nach Filterung mit Nachbarschaftsfilter (rechts)

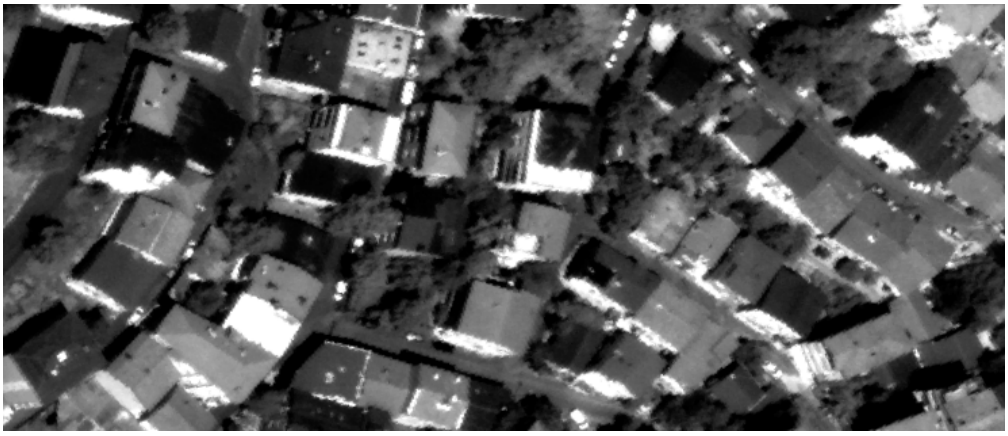


Abb. 9: Detail des WorldView-4-Bildes

Abb. 10: Laserscanning-Referenz-Punktwolke

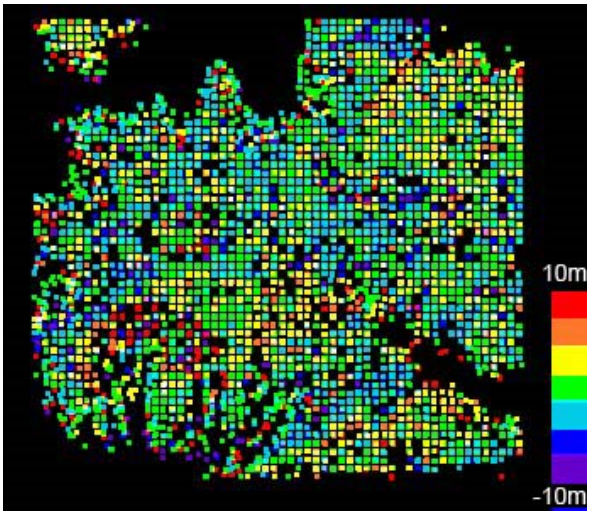


Abb. 11: Farbkodierte Höhendifferenzen

WorldVierw-4 hat mit 31cm GSD eine Objektauflösung, die teilweise auch bei Luftbildern verwendet wird (Abb. 9). Dementsprechend detailliert sind die durch SGM erzeugten Höhenmodelle (Abb. 7). Wie auch bei den Kompsat-3-Daten zeigt sich allerdings auch, dass ein Referenzhöhenmodell mit 4m Punktabstand in dem extrem dicht bebauten Gebiet nicht ausreicht. Die Gebäudeumrisse können damit nicht genau genug beschrieben werden. Hinzu kommen die Verdeckungen in Orbitrichtung durch das Basis-Höhenverhältnis von 1:1,2 etwa entsprechend einer Verdeckung von 40% der Gebäudehöhe. Quer zum Orbit gibt es nur eine Neigung der Blickrichtung von 3°, wodurch sich die Querverdeckungen in Grenzen halten. Das in Tabelle 6 gezeigte NMAD von 90cm für Geländeneigungen unter 10% entspricht nicht der Genauigkeit der Einzelpunkte, die liegt bei gutem Kontrast in der Größenordnung der Höhengenauigkeit der Passpunkte von 27cm (Tabelle 3).

Tab. 6: Vergleich des WorldView-4 Höhenmodells basierend auf SGM mit der Laserscanner-Referenz

SGM	SZ	NMAD	Relatives SZ
Alle Punkte	2.10 m	2.02 m	1.02 m
Für Neigungen < 10%	1.18 m	0.90 m	
Funktion der Neigung	SZ=1.73m + 0.41m * tan a	NMAD=1.48m + 0.72m * tan a	

5 Schlussfolgerungen

Wie leider häufiger, ist die Bildorientierung der Satellitenbilder nicht ohne Probleme, wenn das Genauigkeitspotential der Auswertungen voll ausgeschöpft werden soll. Die Kompsat-3 L1G-Bilder sind für dreidimensionale Auswertungen nicht zu empfehlen, aber auch die Lagegenauigkeit ist eingeschränkt. Mit den Kompsat-3 L1R-Bildern gibt es dagegen keine Probleme und die erzielten Genauigkeiten entsprechen dem GSD. Die WorldView-4-Aufnahmen sind mit Ausnahme der erforderlichen zusätzlichen Bias-Parameter geometrisch unproblematisch.

Wie bekannt, können durch eine flächenbasierte Bildzuordnung nach der Methode der kleinsten Quadrate nicht die Objekte mit senkrechten Begrenzungen, wie die Gebäude und Brücken, nicht ausreichend scharf bestimmt werden, hier gibt keine Probleme mit der Semi Globalen Bildzuordnung, durch die die Gebäudedächer entsprechend der Objektauflösung gut erfasst werden. Durch das Basis-Höhenverhältnis beider Satellitenbildtypen von etwa 1:1,2 gibt es in dem sehr dicht bebauten Aufnahmegebiet Verdeckungen, so dass die Straßenhöhen nicht überall gut erfasst werden. Hier wäre ein kleinerer Konvergenzwinkel der Aufnahmen von Vorteil. Luftbilder mit höherer Längs- und Querüberdeckung haben den Vorteil, dass die Verdeckungen kleiner sind, die Genauigkeit ist allerdings bei gleicher Objektpixelgröße in der gleichen Größenordnung.

6 Literaturverzeichnis

ALOBEID, A., JACOBSEN, K. & HEIPKE, C., 2010**:** Comparison of Matching Algorithms for DSM Generation in Urban Areas from IKONOS Imagery. Photogrammetric Engineering & Remote Sensing, **76**(9) 1041-1050.

HAALA, N., 2011: Multiray Photogrammetry and Dense Image Matching, Photogrammetric Week 2011, Wichmann Verlag, 183-195.

HEINRICHS, M., HELLWICH, O. & RODEHORST, V., 2007: ARTIST: Architectural Model Refinement using Terrestrial Image Sequences from a Trifocal Sensor, Photogrammetrie Fernerkundung Geoinformation, **2007**(6), 405-414.

HIRSCHMÜLLER, H., 2005: Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. IEEE CVPR, San Diego, USA, Juni 2005.

HIRSCHMÜLLER, H., 2008: Stereo Processing by Semi-Global Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, **30**(2), 328-341

HÖHLE, J. & HÖHLE, M., 2009. Accuracy Assessment of Digital Elevation Models by means of Robust Statistical Methods, ISPRS Journal of Photogrammetry and Remote Sensing, **64**, 398-406

JACOBSEN, K., 2006: Comparison of Image Orientation by IKONOS, QuickBird and OrbView-3. EARSeL 2006, Warsaw, in “New Developments and Challenges in Remote Sensing”, edited by Z. Bochenek, Millpress, Rotterdam 2007 ISBN 978-90-5966-053-3, 667-676

Ein systematischer Vergleich verschiedener Multi-View Stereo-Lösungen für die luftbildgestützte dreidimensionale Infrastrukturkartierung

PASCAL SCHÄR[1], STEFAN CAVEGN[1,2], DAVID NOVAK[3], BENJAMIN LOESCH[3], HANNES EUGSTER[1,3] & STEPHAN NEBIKER[1]

Zusammenfassung: Aus der Vogelperspektive abgebildete Infrastrukturanlagen über einen Webdienst am Arbeitsplatz zugänglich zu machen, verspricht grosses Potential. Der Ablauf, um von beliebigen Luftbildaufnahmen zu den in den Webdienst integrierbaren 3D-Bildern zu kommen, beinhaltet die zwei Hauptteile Bildorientierung (Structure-from-Motion) und Multi-View Stereo (MVS). Um die beste Lösung für den zweiten Teil zu bestimmen, werden die drei MVS-Tools SURE, COLMAP und CMVS/PMVS evaluiert. Zur effizienten und pragmatischen Untersuchung wurde ein Workflow implementiert, welcher die automatische Berechnung der Vergleichsresultate ermöglicht. Dieser wird sowohl für den Vergleich, wie auch für die Untersuchung ausgewählter Steuerparameter dieser drei MVS-Algorithmen verwendet. Alle MVS-Untersuchungen werden anhand von vier realen Szenen von Infrastrukturanlagen vorgenommen. Diese zeichnen sich durch verschiedenartige Infrastrukturtypen, Befliegungsmuster und Aufnahmesensoren aus. COLMAP generiert die grösste Punktmenge und SURE rekonstruiert die homogenste Punktwolke. Durch gezielte Parametersteuerung können die guten Resultate der Standardparameter weiter verbessert werden.

1 Motivation und Ausgangslage

Seit einigen Jahren werden Infrastruktur-Inspektionen und 3D-Kartierungen zunehmend nicht mehr vor Ort sondern anhand von Webdiensten effizient am Arbeitsplatz vorgenommen. Die vorausgehende Datenerfassung erfolgt in der Regel entlang von Strassen- und Schienenkorridoren aus der Fahrzeugperspektive (EUGSTER et al. 2013; PUENTE et al. 2013; STANEK & EUGSTER 2017). Dazu werden für die präzise Erfassung fixe Stereokonfigurationen von Kameras in Kombination mit GNSS und Inertialnavigationssystemen (INS) eingesetzt (BURKHARD et al. 2012; CAVEGN & HAALA 2016).

Immer öfter werden Infrastrukturanlagen und Objekte auch mit Luftbildern inspiziert und vermessen (NEITZEL & KLONOWSKI 2011; REMONDINO et al. 2011; NEX & REMONDINO 2014). Dazu werden unterschiedliche Trägerplattformen wie Flugzeuge, Helikopter oder UAVs verwendet. Die aus Luftbildaufnahmen ermittelten Produkte wie Punktwolken, Orthophotos oder Vermaschungen (DOM, DGM) erfüllen die Ansprüche der Kunden, in einfacher Form Informationen abzuleiten, oft nicht zur Genüge. Diese Endprodukte lassen sich zudem für spezifische Fragestel-

[1] Fachhochschule Nordwestschweiz, Institut Geomatik, Gründenstrasse 40, CH-4132 Muttenz, Schweiz
E-Mail: [pascal.schaer, stefan.cavegn, stephan.nebiker]@fhnw.ch
[2] Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart
[3] iNovitas AG, Oberrohrdorferstrasse 1c, CH-5405 Baden-Dättwil, Schweiz
E-Mail: [david.novak, benjamin.loesch, hannes.eugster]@inovitas.ch

lungen nur umständlich bis gar nicht einsetzen oder die Weiterverarbeitung ist für die Kunden komplex und problematisch.

Das Forschungsprojekt *INFRA BirdView* des Instituts Geomatik (IGEO) der Fachhochschule Nordwestschweiz (FHNW) in Zusammenarbeit mit der iNovitas AG hat zum Ziel, beliebige Luftbildaufnahmen einer Infrastrukturanlage in einfacher und effizienter Form mittels 3D-Bildern dem Kunden über einen Webdienst zur Verfügung zu stellen. Die Bilder aus der Luft, je nach Objekt ergänzt durch terrestrische Aufnahmen, können einen entscheidenden Mehrwert zur Beurteilung und Vermessung von Infrastrukturanlagen bieten. Der Web-Viewer ermöglicht eine virtuelle Befliegung der Infrastrukturanlage inklusive Mess- und Erfassungsmöglichkeiten. Die Generierung der dazu nötigen 3D-Bilder aus den Luftbildern und die entsprechende Einbindung in einen Webdienst bergen im Vergleich zur Stereokonfiguration einige neue Herausforderungen. Aufgrund der deutlich vergrösserten Aufnahmedistanz zu den abzubildenden Infrastrukturobjekten sind physische Stereomesssysteme unvorteilhaft, was die Bestimmung von virtuellen Stereobasen mittels bildbasierter Georeferenzierung erfordert. Zudem ist die effiziente Handhabung deutlich grösserer Bilddateien mit der Möglichkeit zur massiven Skalierung bzw. Parallelisierung sämtlicher Prozessschritte nicht trivial.

3D-Bilder sind eine Kombination der verzeichnungsfreien Originalbilder und den dazu berechneten Tiefenkarten, welche Tiefeninformationen ausgehend vom Bildstandpunkt beinhalten (NEBIKER et al. 2015). Orthophotos weisen, aufgrund der vereinfachten 2.5D Geometrie, im Gegensatz zu 3D-Bildern einen Informationsverlust bei Untersichten oder vertikalen Flächen auf. Eine kombinierte Nutzung von fahrzeugbasierten Stereobilddaten und Luftbilddaten benötigt eine Co-Registrierung beider Aufnahmen (NEBIKER et al. 2012; NEBIKER et al. 2013; KAMER et al. 2013; GERKE et al. 2016).

Der entwickelte Workflow (vgl. Abb. 1), um Tiefenkarten für die Ausgangsbilder zu ermitteln, unterstützt beliebige Luftbildaufnahmen, folglich auch Befliegungen von Objekten wie Gebäude oder Brücken. Dieser Workflow kann in die drei Hauptteile *Bildorientierung* (Structure-from-Motion (SfM)), *Multi-View Stereo* (MVS) und *Aufbereitung der Tiefenkarten* unterteilt werden. Für den MVS-Teil wurden drei Lösungen untersucht und evaluiert, um diese optimal im Workflow einzusetzen. Zusätzlich konnten für unterschiedliche Szenarien optimierte Steuerungsparameter abgeleitet werden. Auf diese Untersuchungen und Resultate wird in diesem Beitrag vertieft eingegangen.

HAALA (2014) vergleicht und beurteilt die Resultate von elf Dense Image Matching-Tools, die im Rahmen eines Benchmarks mit Nadir-Luftbildaufnahmen generiert wurden. Eine Erweiterung dieses Benchmarks um Oblique-Luftbildaufnahmen einer Leica RCD30 wird in CAVEGN et al. (2014a) sowie HAALA & CAVEGN (2015) vorgestellt. Auch wurden bereits erste Leistungsuntersuchungen von *SURE* und *Agisoft PhotoScan* anhand von terrestrischen Laserscans als Referenzdaten vorgenommen und in HAALA & CAVEGN (2016) weitergeführt. NEX et al. (2015) stellen einen weiteren Benchmark-Datensatz inkl. Referenzdaten vor, welcher Oblique-Aufnahmen von einem Flugzeug, UAV-Aufnahmen, aber auch terrestrische Bildaufnahmen umfasst. Die Leistungsfähigkeit der vier Dense Image Matching-Algorithmen *SURE*, *MicMac*, *PMVS* und *Agisoft PhotoScan* wird von REMONDINO et al. (2014) anhand von acht Datensätzen evaluiert. SCHÖPS et al. (2017) analysierten in ihrem Multi-View Stereo Benchmark vier unterschiedliche MVS-Tools – zwei davon wurden für diesen Beitrag ebenfalls evaluiert – mit hochauflösenden

Spiegelreflexkamerabildern sowie synchronisierten niedrig aufgelösten Stereo-Videos. Dafür nahmen SCHÖPS et al. (2017) sowohl Aussen- wie auch Innenraumszenen auf und bestimmten mit einem Laserscanner Referenzdaten, welche zur Ableitung für Qualitäts- und Genauigkeitsaussagen dienten. SEITZ et al. (2006) untersuchten verschiedene MVS-Algorithmen anhand von 38 Innenraumszenen mit niedrig aufgelösten Bildern (VGA-Auflösung) und Referenzdaten von einem Streifenprojektionsscanner. SCHARSTEIN et al. (2014) lieferten mit 6 Megapixel den höher aufgelösten *Middlebury* Datensatz und zeigten damit Schwierigkeiten für MVS-Algorithmen im Bereich von Auflösung und Szenenkomplexität auf. GEIGER et al. (2012) verwendeten für ihren MVS-Benchmark mit dem *KITTI* Datensatz niedrig aufgelöste Bilder (0.5 Megapixel), welche von einer mobilen autonom fahrenden Plattform erfasst wurden. REMONDINO et al. (2017) beurteilen den automatisierten photogrammetrischen Rekonstruktionsprozess kommerzieller Softwarelösungen anhand terrestrischer aber auch Luftbildaufnahmen. Um die Leistungsfähigkeit unterschiedlicher Dense Image Matching-Lösungen bezüglich Oblique-Luftbildaufnahmen zu vergleichen, entwickelten CAVEGN et al. (2014b) eine Untersuchungsmethodik mit Analysen im Bild- und Objektraum, welche weitgehend system- und formatunabhängig einsetzbar ist. Die Vergleiche im vorliegenden Beitrag basieren auf einer pragmatischen und effizienten Herangehensweise. Dafür werden hochaufgelöste Luftbilder mit bis zu 30 Megapixel verwendet. Zusätzlich kommen durchaus grössere Blickwinkelunterschiede in den für die Untersuchungen verwendeten Luftbilddatensätzen vor.

2 Grundlagen für MVS-Untersuchungen

In diesem Kapitel wird der Gesamtworkflow vom erfassten Luftbild bis zur Tiefenkarte, sowie die untersuchten MVS-Matcher und die Infrastruktur-Anlageszenen beschrieben. Allen Untersuchungen und Daten ist dabei gemein, dass keine Referenzdaten vorliegen oder erfasst werden konnten. Die Resultate der untersuchten MVS-Matcher können untereinander, jedoch nicht mit Soll-Daten verglichen werden. Die Qualitätsbeurteilung der MVS-Algorithmen ist daher nicht auf Grundlage einer Referenz möglich. Da auch HAALA (2014) über keine Referenzdaten verfügte, verwendete er den Median aller Matching-Resultate als Vergleichswert.

Der Gesamtworkflow ist in Abb. 1 visualisiert und zeigt auf, mit welchen Prozessschritten aus Rohdaten in Form von Luftbildern die entsprechenden Tiefenkarten berechnet werden, um diese im Webdienst einzusetzen. Die Bilder können von unterschiedlichen Plattformen und Sensoren erfasst und gemeinsam ausgewertet werden. Für die Bildorientierung (SfM) wird eine adaptierte und erweiterte Version der Open Source Software COLMAP eingesetzt, mit welcher die inneren (IOP) sowie äusseren Orientierungen (EOP) zur Effizienz- und Zuverlässigkeitssteigerung eingebunden werden können. Über Passpunkte (GCP) lassen sich die Luftbilder (präziser) in ein gefordertes Bezugssystem einbinden. Die Angabe von IOP, EOP und GCP sind optional. Bei keiner Eingabe dieser Parameter erfolgt die Berechnung in einem lokalen nicht-metrischen Bezugssystem. Aus dem Bildorientierungsprozess (SfM) resultieren die inneren und äusseren Orientierungsparameter aller Luftbilder, welche die Ausgangslage für die anschliessende Multi-View Stereo (MVS) Berechnung bilden. Das Ziel der MVS-Berechnung ist die komplette Rekonstruktion eines 3D-Modells basierend auf den erfassten Luftbildern mit bekannter innerer und äusserer Orientierung. Dafür existieren die zwei bekannten und in Abb. 1 dargestellten Ansätze.

Der *Stereo Matching* Ansatz berechnet aus jeweils zwei Bildern mit überlappenden Bildbereichen die entsprechenden Normalbilder (Stereonormalfall – Eliminierung der y-Parallaxe), um den Suchraum der dichten Bildkorrelation (Dense Image Matching) einzuschränken. Die daraus resultierenden rohen Tiefenkarten aus jeweils zwei Bildern werden im darauffolgenden Schritt zu einer Tiefenkarte pro Bild fusioniert. Der *Multi-View Matching* Ansatz generiert direkt über die epipolare Beziehung zweier oder mehrerer Bilder 3D-Punkte, welche nach diversen Filteroperationen eine 3D-Punktwolke ergeben. Aus dieser Punktwolke wird anschliessend für jedes Luftbild eine Tiefenkarte zurück gerechnet. Die resultierende Tiefenkarte (TK) pro Luftbild wird danach zusammen mit dem verzeichnungsbereinigten Luftbild im Viewer des Webdienstes eingebunden.

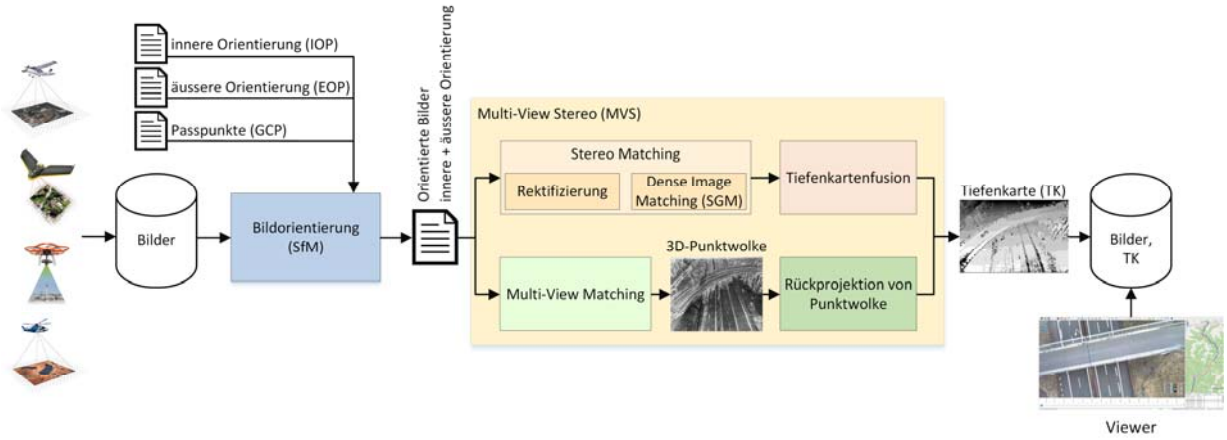


Abb. 1: Ablauf von den erfassten Luftbildern über die Bildorientierung und Multi-View Stereo zu den resultierenden Tiefenkarten

2.1 Untersuchte MVS-Lösungen

Um eine möglichst optimale MVS-Lösung im Gesamtworkflow zu integrieren, wurden die drei MVS-Tools SURE (ROTHERMEL et al. 2012; WENZEL et al. 2013), COLMAP (SCHÖNBERGER et al. 2016) sowie CMVS (FURUKAWA et al. 2010) & PMVS (FURUKAWA & PONCE 2010) evaluiert und untersucht. Die Auswahl erfolgte primär aufgrund der einfachen Verfügbarkeit der Programme, sowie der Möglichkeit zur Einbindung in den Gesamtworkflow. Tab. 1 liefert eine Übersicht über Ansatz, Input und Output der drei MVS-Tools, welche nachfolgend detaillierter beschrieben sind.

2.1.1 SURE

SURE (ROTHERMEL et al. 2012; WENZEL et al. 2013) ist eine kommerziell vertriebene MVS-Software von *nFrames* und wurde ursprünglich am Institut für Photogrammetrie der Universität Stuttgart entwickelt. SURE verwendet einen hierarchischen Ansatz, welcher niedrig aufgelöste Matchingresultate nutzt, um die dichte Bildkorrelationssuche höher aufgelöster Bilder einzuschränken. Dadurch werden Rechenzeit und Rechenressourcen (RAM Speicherbedarf) bei gleichbleibender Qualität der Resultate entscheidend reduziert. Die dichte Bildkorrelation (Dense Image Matching) basiert auf der Semi-Global Matching (SGM) Methode von HIRSCHMÜLLER

(2008). Anschliessend folgt eine rechenaufwändige Fusion der redundanten Tiefeninformationen zu einer vollständigen Tiefenkarte pro Bild. Diese Fusion relativiert die ursprünglich hohe Geschwindigkeit des hierarchischen SGM. SURE bietet unter anderem Schnittstellen für den Import von Bildorientierungen aus *VisualSFM*, *Bundler* oder *Pix4D*. Die Prozessierung kann dabei wahlweise auf der CPU oder einer *CUDA* fähigen *NVIDIA* GPU ausgeführt werden.

Tab. 1: Eingabedaten und Ausgabeprodukte der untersuchten MVS-Algorithmen. Produkte in [ ] können optional ausgegeben werden.

	SURE	**COLMAP**	**CMVS & PMVS**
Ansatz	Stereo Matching	Stereo Matching	Multi-View Matching
Input	Bilder Projektionsmatrix **P** Steuerparameter	Bilder Projektionsmatrix **P** Steuerparameter	Bilder Projektionsmatrix **P** Steuerparameter
Output	Punktwolke [Tiefenkarte] Oberflächenmodell (DSM) Orthophoto	Punktwolke Tiefenkarte Normalenkarte [Konsistenzgraph] Mesh	Punktwolke

2.1.2 COLMAP

Die Open Source Software COLMAP deckt sowohl den Bildorientierungsteil (SfM) (SCHÖNBERGER & FRAHM 2016) wie auch den MVS-Teil (SCHÖNBERGER et al. 2016) ab. Der SfM-Teil wurde für den Einsatz zu Vermessungszwecken adaptiert und erweitert. Der MVS-Teil besteht aus den vier Modulen Bereinigung der Bildverzeichnung, dichte Bildkorrelation, Tiefenkartenfusion und der Vermaschung, welche sequentiell zu durchlaufen sind. Die Berechnung kann wahlweise auf der CPU oder einer *CUDA* fähigen *NVIDIA* GPU ausgeführt werden. COLMAP MVS basiert auf dem Framework von ZHENG et al. (2014). Dieses wurde erweitert und unter anderem durch die pixelweise Schätzung der Normalen oder die pixelweise Filterung unter Verwendung von Triangulationswinkel, Einfallswinkel und Bildauflösung optimiert.

2.1.3 CMVS & PMVS

CMVS und PMVS sind zwei unterschiedliche und grundsätzlich voneinander unabhängige MVS-Tools. CMVS (FURUKAWA et al. 2010) selektiert und clustert Eingabebilder möglichst optimal für MVS-Algorithmen in handhabbare Teilmengen, so dass ein minimaler Verlust von Inhalt und Detail der MVS-Resultate auftritt. Die Verknüpfungspunkte der Bildorientierung (sparse point cloud) werden von CMVS zum Clustering verwendet. Die Bildcluster werden einzeln prozessiert und die Resultate anschliessend zusammengeführt. Bilder, die voraussichtlich einen kleinen Beitrag zur kompletten Rekonstruktion liefern, werden vom Algorithmus herausgefiltert. CMVS unterstützt daher die Berechnung grosser MVS-Projekte und trägt zur Skalierung bei. Zudem bietet CMVS eine nahtlose Schnittstelle zu PMVS (FURUKAWA & PONCE 2010), welches mit einer musterbasierten Multi-View Matching (patch match) Methode dichte Punktwolken erzeugt. Dafür werden Flächenmuster über Verknüpfungspunkte (aus Harris Cor-

ner Detector und Difference of Gaussian) initialisiert. Anschliessend werden ausgehend von diesen Flächenmustern weitere übereinstimmende angrenzende Bildbereiche gesucht, um ein 3D-Modell zu rekonstruieren. PMVS verarbeitet – begründet durch den Multi-View Matching Ansatz – mehrere Bilder simultan, was einen grossen Bedarf an Arbeitsspeicher (RAM) auslöst. In Kombination mit CMVS wird PMVS zu einem einsetzbaren MVS-Tool. Für die Untersuchungen wurde die frei zugängliche und leicht erweiterte Version CMVS-PMVS2 (MOULON et al. 2011) verwendet.

2.2 Szenen von Infrastrukturanlagen

Die MVS-Algorithmen wurden anhand von vier realen Szenen von Infrastrukturanlagen untersucht. Diese Anwendungsfälle zeichnen sich durch verschiedenartige Infrastrukturtypen, Befliegungsmuster und Aufnahmesensoren aus. Zudem stellen heterogene Beleuchtung, homogene Flächen und spiegelnde Flächen grosse Herausforderungen für MVS dar. Die vier Infrastrukturanlageszenen und die entsprechenden Befliegungsdetails werden nachfolgend vorgestellt (vgl. Tab. 2). Abb. 2 zeigt zudem je ein charakteristisches Luftbild pro Szene.

Tab. 2: Übersicht der vier Infrastruktur-Anlageszenen und Befliegungsinformationen

Bezeichnung	**Espark**	**Brenner**	**FHNW-Campus**	**Strommast**
Infrastrukturobjekt	Industriegebäude mit Solardach	zweispurige Autobahn	Grossbaustelle	Strommasten mit Transportleitungen
Befliegungsmuster	Zirkular	Linear	Flächig	Dreidimensional
Sensorplattform	Multikopter UAV	Helikopter	Flächenflügel UAV	Multikopter UAV
Ausdehnung	70x50x10 m	5 km	300x300x70 m	30x30x100 m
Ø Entfernung	20 m	80 m	140 m	15 m
Befliegungsdauer	15 min	1 – 1.5 h	25 min	1 h
Befliegungsdatum	10.08.2016	23.02.2016	18.10.2016	20.09.2016
Anzahl Bilder	365	626’000	170	635

Die Szene des Industriegebäudes *Espark* zeichnet sich durch einige homogene Flächen mit wiederholenden Strukturen aus. Insbesondere der Vorplatz und die Solarpanels auf dem Dach stellen für die MVS-Algorithmen eine Herausforderung dar. Die Luftbilder wurden mit einer schräg ausgerichteten *dji Phantom 3* Kamera (9 MP, 4096×2160 Pixel) erfasst. Die Kamera wurde nicht kalibriert und die verwendete Drohne konnte den Bildern keine Geotags hinzufügen. Die starke Verschwenkung der Bilder erschwert zudem die automatische Bildorientierung.
Bei der Befliegung der *Brenner* Autobahn wurden zwei Kameras mit unterschiedlicher Auflösung mitgeführt. Die *Allied Vision AVT – GE4000 prosilica* Industriekamera war Nadir ausgerichtet und erfasst Bilder mit einer Auflösung von 11 MP (4008×2672 Pixel). Die zweite *Allied Vision* Kamera vom Typ *piA 1900-32GC* war Oblique ausgerichtet und liefert mit 2 MP (1920×1080 Pixel) niedriger aufgelöste Bilder. Zur Kalibrierung der Kameras wurde ein mit ca. 30 Punkten erstelltes 50×50 m grosses Kalibrierfeld überflogen. Mittels direkter Sensororientierung wurden Näherungswerte für die äussere Orientierung auf Grund von GNSS und INS-Messungen berechnet. Der lineare Perimeter erstreckt sich über eine Länge von 5 km und weist

neben der Autobahn ebenfalls Brücken, Ein- und Ausfahrten, Kandelaber und Verkehrsschilder auf. Die Homogenität der Fahrbahn ist herausfordernd für MVS-Algorithmen.
Der *FHNW-Campus* wurde in einer Kreuzbefliegung (Überlappung L=80%, Q=80%, GSD 4 cm) mit einer *Sony DSC-WX220* (CMOS Sensor) mit 18 MP (4896×3672 Pixel) aufgenommen. Die Kamera wurde Nadir ausgerichtet und nicht kalibriert. Neben dem markanten Rohbau prägen Industriegebäude und das angrenzende Einfamilienhausquartier sowie die Bahnlinien die Szene.
Der *Strommasten* mit Transportleitungen ragen ca. 80 m in die Höhe, wobei der Hintergrund sich mit zunehmender Flughöhe weiter weg befindet. Dies ist eine ungünstige Situation für die Bildorientierung und Rekonstruktion des Mastes. Zudem weist der Strommasten eine homogene grüne Oberfläche mit runden Verstrebungen auf. Für die Befliegung wurde bewusst eine für photogrammetrische Zwecke optimale Flugkonfiguration gewählt. Aufgrund der unterschiedlichen Entfernungen zum Objekt bei der Aufnahme, wurden scharfe Bilder höher gewichtet als die Kamerakalibrierung. Die hochauflösenden Vollformat-Bilder mit 36 MP (7360×4912 Pixel) wurden mit einer *SONY ILCE-7R* Kamera mit einer Pixelgrösse von 4.9 µm (CMOS Sensor) aufgenommen.

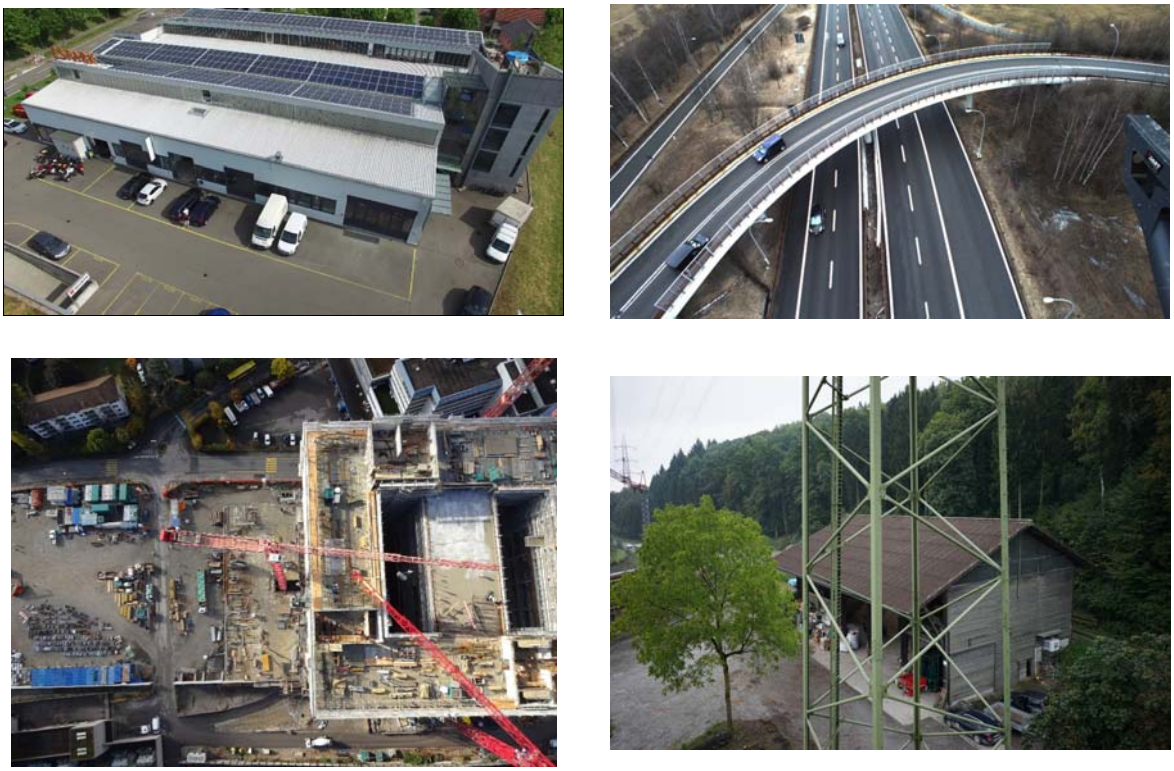


Abb. 2: Charakteristische Luftbilder der vier untersuchten Szenen: Espark (oben links), Brenner-Autobahn (oben rechts), Grossbaustelle FHNW-Campus (unten links) und Strommasten (unten rechts)

3 Untersuchungen zu Multi-View Stereo

Für die effiziente Untersuchung und Evaluierung der in Kap. 2.1 vorgestellten MVS-Algorithmen wurde ein automatischer Workflow umgesetzt. Dieser deckt einen Teil des Gesamtworkflows ab (vgl. Kap. 2) und wird nachfolgend erläutert. Zuvor wird aber auf die Parameteruntersuchungen der drei MVS-Algorithmen eingegangen.

3.1 Parameteruntersuchung

Die Parameteruntersuchungen basieren hauptsächlich auf Vergleiche der MVS-Resultate. Aus Effizienzgründen, den vielfältigen Möglichkeiten für Vergleiche und da alle MVS-Algorithmen eine Punktwolke als Output liefern, wurden diese als Vergleichsbasis gewählt (vgl. Abb. 3). Da das Ableiten von aussagekräftigen Werten, welche die Qualität von Punktwolken beschreiben, aufwändig ist, wird eine pragmatische Herangehensweise gewählt. Die Punktwolken und somit die MVS-Resultate werden anhand von Punktwolkenansichten (Renderings) und Querschnitten (Profilen) von identischen Positionen aus für jede Punktwolke visuell beurteilt (vgl. Abb. 3). Dies hat den Vorteil einer grossen Effizienz, jedoch den Nachteil einer gewissen Aussagenunschärfe mangels abgeleiteter Qualitätswerte. Die Parameteruntersuchung hat zum Ziel, optimale Parameter der drei MVS-Algorithmen für die unterschiedlichen Infrastrukturanlageszenen abzuleiten. Dafür wurden 4-5 Steuerungsparameter pro MVS-Algorithmus ausgewählt, die gemäss Beschreibungen und Erfahrung den grössten Einfluss haben. Zusätzlich wurde versucht, gleiche oder ähnliche Steuerungsparameter festzulegen, so dass eine gewisse Vergleichbarkeit der Resultate möglich bleibt. Die gewählten Steuerungsparameter sind in den nachfolgenden Tab. 3 – 5 aufgelistet. Es wurde für jeden Steuerungsparameter ein Tief- (TW) und ein Hochwert (HW) (faktorieller 2^k-Versuchsplan) bestimmt und prozessiert. Daraus resultieren gesamthaft 96 zu beurteilende Punktwolken (*(1+9+8+6 Parameter)* □ *4 Szenen = 96*).
Die Wahl der Steuerungsparameter fiel hauptsächlich auf die Pyramidenstufe (Auflösungsstufe) der prozessierten Bilder, sowie der Beschränkung der Triangulationswinkel (Minimum und Maximum). Ergänzt wurde die Wahl durch Parameter, welche die minimale Anzahl Modelle, Bilder oder fusionierter Pixel definieren, dass ein 3D-Punkt rekonstruiert und nicht gefiltert wird.

Tab. 3: Ausgewählte Steuerungsparameter von SURE, sowie Standardwert (SW), Tiefwert (TW) und Hochwert (HW) der Parameteruntersuchung

		SURE			
Nr.	**Parameter**	**SW**	**TW**	**HW**	**Kurzbeschreibung**
1	pyr	0	–	1	Pyramidenstufe
2	fold	2	1	3	minimale Anzahl Modelle pro Pixel
3	minangle	4°	0.5°	5°	minimaler Triangulationswinkel
4	maxangle	–	140°	–	maximaler Triangulationswinkel

Tab. 4: Ausgewählte Steuerungsparameter von COLMAP, sowie Standardwert (SW), Tiefwert (TW) und Hochwert (HW) der Parameteruntersuchung

COLMAP					
Nr.	**Parameter**	**SW**	**TW**	**HW**	**Beschreibung**
1	windows_radius	5	3	7	Grösse Kreuzkorrelationsfenster
2	sigma_spatial sigma_color ncc_sigma	5 0.2 0.6	3 0.1 0.4	7 0.3 0.8	(Gewichtungs-) Parameter der Kreuzkorrelationsberechnung
3	min_triangulation_angle	1	0.5°	5°	minimaler Triangulationswinkel
4	filter_min_num_consistent	2	–	3	minimale Anzahl Bilder pro Pixel
5	min_num_pixels	5	2	3	minimale Anzahl fusionierter Pixel

Tab. 5: Ausgewählte Steuerungsparameter von PMVS, sowie Standardwert (SW), Tiefwert (TW) und Hochwert (HW) der Parameteruntersuchung. *SW = 1 wurde für Untersuchungen auf 0 angepasst.

PMVS					
Nr.	**Parameter**	**SW**	**TW**	**HW**	**Beschreibung**
1	level	0*	–	1	Pyramidenstufe
2	threshold	0.7	0.4	0.85	Kreuzkorrelationsgrenzwert
3	maxAngle	10°	0.5°	5°	minimaler Triangulationswinkel
4	minImageNum	3	2	–	minimale Anzahl Bilder pro Pixel
5	wsize	2	5	15	Grösse Kreuzkorrelationsfenster

3.2 Workflow für MVS-Vergleich

Der in Abb. 3 dargestellte Workflow für die Multi-View Stereo (MVS) Untersuchungen ermöglicht die vollautomatische Prozessierung beliebiger Flugprojekte mit den drei in Kap. 2.1 vorgestellten MVS-Algorithmen. Für einen korrekten Vergleich der MVS-Resultate muss die Ausgangslage, d.h. die Bildorientierung zwingend identisch sein. Dafür wird die COLMAP-Bildorientierung (SfM) (SCHÖNBERGER & FRAHM 2016) des Gesamtworkflows (vgl. Abb. 1) eingesetzt. Diese liefert die äussere Orientierung in Form von Rotationsmatrix und Translationsvektor. Die Schnittstelle zum jeweiligen MVS-Tool transformiert diese vorliegenden Bildorientierungsdaten ins jeweilig benötigte Input-Format (vgl. Tab. 1). Für die Berechnung der verzeichnungsbereinigten Luftbilder wird auf das Modul zur Bildverzeichnisbereinigung von *COLMAP* (SCHÖNBERGER et al. 2016) zurückgegriffen (vgl. Kap. 2.1.2). Die definierten Steuerungsparameter (vgl. Tab. 3 – 5) werden den MVS-Tools übergeben, welche jeweils eine Punktwolke als Einzelresultat pro Parametersatz liefern. Für jede Punktwolke werden anschliessend automatisiert und daher basierend auf der äusseren Orientierung der Luftbilder identische Ansichten gerendert und kongruente Vertikalschnitte erzeugt, welche zur visuellen Leistungs- und Qualitätsbeurteilung herangezogen werden. Für das Rendering wird die frei erhältliche Software *VTK 8.0* (KITWARE INC. 2018) eingesetzt.

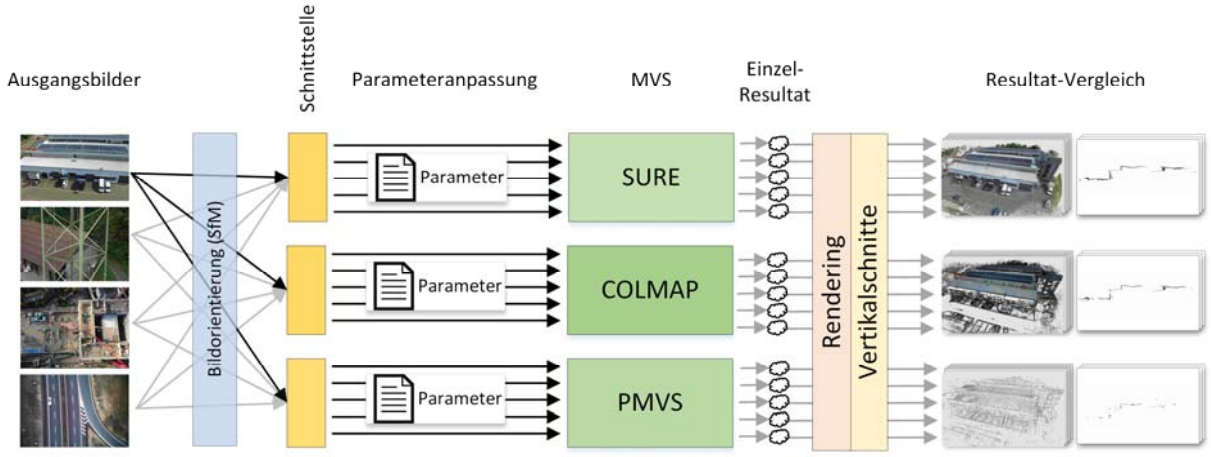


Abb. 3: Implementierter Workflow für effiziente Multi-View Stereo (MVS) Untersuchungen

4 Ergebnisse

Die gerechneten Punktwolkenansichten und Vertikalschnitte lassen eine visuelle Beurteilung der Punktwolkendichte, Punktverteilungen, Streuverhalten, Kantenschärfe sowie Problemzonen und Limitierungen der MVS-Algorithmen zu. Dies sind Qualitäts- und Leistungsmerkmale der untersuchten MVS-Algorithmen. Zusätzlich werden die Gesamtpunktmengen und Prozessierungszeiten des implementierten Workflows dokumentiert. Die Resultate wurden mit einem *DELL Precision Tower 7910* Rechner mit nachfolgend aufgelisteten Hardwarekomponenten generiert. Für die MVS-Prozessierung der 96 unterschiedlichen Konfigurationen benötigte dieser Computer eine Rechenzeit von 233.8 h (9 T 18 h).

- CPU: 3.0 GHz *Intel Xeon* Octo Core
- RAM: 32 GB, + 20 GB ausgelagert auf einer SSD (swap)
- GPU: *CUDA 8.0* fähige *NVIDIA GeForce GTX 1080* mit 8 GB OnBoard Memory

Die Parameteruntersuchung zeigt, dass die Standardparameter des jeweiligen MVS-Tools bereits gute Resultate liefern. Bei guten Oberflächenstrukturen generiert COLMAP die dichteste Punktwolke. Über den ganzen Projektperimeter gesehen erzeugt aber SURE die homogenste Punktwolke der drei MVS-Algorithmen. PMVS liefert lückenhafte Resultate, extrahiert jedoch insbesondere Begrenzungslinien gut. In den nachfolgenden zwei Kapiteln wird detaillierter auf die Resultate der Parameteruntersuchungen und des MVS-Vergleichs eingegangen.

4.1 Parameteruntersuchung

Für die Parameteruntersuchung werden die Resultate aus der Berechnung mit den Standardparametern, mit den jeweiligen des Hoch- und Tiefwerts eines angepassten Steuerungsparameters verglichen. Grundsätzlich lässt sich festhalten, dass mit den Standardparametern gute Resultate generiert werden, welche für gewisse Szenen durch gezielte Anpassung noch optimiert werden können. Der grösste Einfluss bei SURE hat der Steuerparameter *fold* (minimale Anzahl Modelle um 3D-Punkt zu rekonstruieren). Für den Tiefwert (TW) liefert SURE eine beinahe komplette

Punktwolke, welche jedoch viele grobe Fehler aufweist. Der Hochwert (HW) von *fold* bewirkt insbesondere in Randbereichen eine lückenhafte Punktwolke, dafür ist die Streuung geringer und mit COLMAP vergleichbar. Es gilt herauszustreichen, dass der Steuerparameter *pyr = 1* (Bildpyramidenstufe 1 – reduzierte Bildauflösung) erstaunlich gute Resultate liefert und dies bei stark reduzierter Berechnungsdauer. In gewissen Bereichen ist das Resultat sogar besser als bei voller Auflösungsstufe (Bildpyramidenstufe 0). Dies ist vermutlich darauf zurückzuführen, dass die Sensoren aufgrund des Bayermusters die Szene tatsächlich mit einer geringeren Auflösung erfassen als ausgegeben. Den grössten Einfluss der untersuchten Steuerparameter bei COLMAP haben die Parameter *Nr. 2* (vgl. Tab. 7), welche die Kreuzkorrelation steuern. Die tiefen Werte (TW) liefern eine lückenhafte Punktwolke. Die hohen Werte (HW) ergeben eine dichtere Punktwolke bei relativ kleiner Zunahme der Streuung. Die Wahl des COLMAP-Parameters *filter_min_num_consistent* beeinflusst die Zuverlässigkeit der Rekonstruktionen und die Reduktion von Ausreissern bei gleichzeitiger Ausdünnung der Randbereiche und Eliminierung teils korrekter Resultate. Tab. 6-8 fassen die Resultate der Parameteruntersuchungen zusammen, wobei *Zeit* und *Punkte* sich nur auf die Szene des Industriegebäudes *Espark* beziehen. Diese sind für die anderen Szenen vergleichbar.

Tab. 6: Zusammenstellung der Resultate der Parameteruntersuchung von SURE.
(Für Erläuterungen siehe Tabellenlegende weiter unten)

SURE							
Nr	**Parameter**	**SW**	**TW**	**HW**	**Zeit TW, HW**	**Punkte TW,HW**	**Erkenntnis**
0	Standard				1:24	7.2	
1	pyr	0	–	1	0:24	7.0	HW: gute Resultate möglich
2	fold	2	1	3	1:27, 1:20	10.2, 4.3	HW bei hoher Bilddichte
3	minangle	4°	0.5°	5°	1:25, 1:24	9.1, 6.6	Allg. Empfehlung 3°
4	maxangle	–	140°	–	1:29	7.2	bei Objektbefliegung sinnvoll

Tab. 7: Zusammenstellung der Resultate der Parameteruntersuchung von COLMAP.
(Für Erläuterungen siehe Tabellenlegende weiter unten)

COLMAP							
Nr	**Parameter**	**SW**	**TW**	**HW**	**Zeit TW, HW**	**Punkte TW, HW**	**Erkenntnis**
0	Standard				3:45	12.9	
1	windows_radius	5	3	7	3:46, 3:46	12.9, 12.9	TW: Reduktion Ausreisser
2	sigma_spatial sigma_color ncc_sigma	5 0.2 0.6	3 0.1 0.4	7 0.3 0.8	3:28, 3:51	3.7, 13.4	TW: lückenhaftes Punktw. HW: dichter, grössere Streuung
3	min_triangulation_angle	1	0.5°	5°	4:00, 3:47	12.9, 12.9	Allg. Empfehlung 2°-3°
4	filter_min_num_consistent	2	–	3	3:51	11.5	Randbereiche stark gefiltert
5	min_num_pixels	5	2	3	4:00, 3:47	12.9, 12.9	Allg. Empfehlung 4

Tab. 8: Zusammenstellung der Resultate der Parameteruntersuchung von PMVS.
(Für Erläuterungen siehe Tabellenlegende weiter unten)

PMVS							
Nr	**Parameter**	**SW**	**TW**	**HW**	**Zeit TW, HW**	**Punkte TW, HW**	**Erkenntnis**
0	Standard				0:29	0.7	
1	level	0*	–	1	0:07	0.3	HW: weniger Ausreisser
2	threshold	0.7	0.4	0.85	1:20, 0:14	0.8, 0.6	mehr Punkte & Ausreisser
3	maxAngle	10°	0.5°	5°	0:27, 0:26	0.7, 0.7	Allg. Empfehlung 2°-3°
4	minImageNum	3	2	–	1:14	0.5	
5	wsize	2	5	15	0:26, 1:11	0.6, 1.08	Empfehlung 8-10

Tabellenlegende

grün Parameter mit den besten Resultaten
orange Parameter mit schlechten Resultaten
blau optimale Parameter liegen zwischen diesen Werten
gelb stark szenen- und objektabhängig inkl. Aufnahmegrund
Zeile ohne Einfärbung – Steuerparameter hat keinen sichtbaren Einfluss
Zeit Berechnungszeit der Infrastrukturszene *Espark*
Punkte Punktmenge der Infrastrukturszene *Espark* in der Grössenordnung Million
Die Berechnungszeiten und Punktmengen der Infrastrukturszene *Espark* stehen stellvertretend für alle Szenen, welche ähnliche Resultatverhältnisse aufweisen.

4.2 Vergleich der MVS-Tools

Der Vergleich der MVS-Tools erfolgt anhand von Resultaten, die mit Standardparametern prozessiert wurden. SURE weist über das gesamte Bild bzw. den gesamten Projektperimeter die homogenste Punktwolke auf. Dies vor allem da SURE auch bei homogenen Oberflächenstrukturen wie Fenster, Belag, Schattenwurfbereiche und vertikale homogene Fassadenelemente 3D-Punkte rekonstruieren kann. SURE weist zudem eine geringe Anzahl an Ausreisser auf. Dafür ist die Streuung deutlich höher als bei COLMAP, insbesondere bei homogenen Oberflächenstrukturen. Die Kantenschärfe bei SURE ist ebenfalls niedriger als bei COLMAP. Dies macht sich besonders bei abstehenden Objekten wie Kandelaber, Fahrleitungsmaste, Strassenschilder oder kahle Baumstämme bemerkbar, welche SURE mässig bis gar nicht und COLMAP zumindest erkennbar rekonstruiert. SURE hat zudem Mühe, 3D-Punkte abzuleiten, wenn ein vorgelagertes Objekt dahinterliegende Bereiche teilweise verdeckt.
COLMAP rekonstruiert mit 1.1 bis 1.8 Mal mehr als SURE und bis zu 25 Mal mehr als PMVS die grösste Punktmenge der drei Tools. Im Besonderen bei markanten Oberflächenstrukturen liefert COLMAP die höchste Punktdichte, weist jedoch bei Fenstern, Belag, Schattenwurfbereichen und vertikalen homogenen Fassadenelementen wenige bis keine Punkte auf. Ein Vorteil von COLMAP ist die geringe Streuung und hohe Kantenschärfe. Vereinzelt kommen jedoch grobe Ausreisser vor. Diese Ausreisser sind Teilflächen, welche grob falsch sind, jedoch mit zunehmender Bildanzahl bzw. Bildüberlappung abnehmen.

PMVS liefert lückenhafte Resultate mit der kleinsten Punktmenge, extrahiert dennoch Begrenzungslinien und somit Kanten mit einer geringen Streuung gut.
COLMAP benötigt deutlich am längsten für die Prozessierungen, dies trotz des GPU-Einsatzes. SURE ist mindestens um Faktor zwei schneller als COLMAP und dies obwohl die Prozessierung auf der CPU erfolgte. Es wird erwartet, dass der Einsatz einer GPU die Prozessierungsdauer von SURE senkt. PMVS benötigt mit Abstand am wenigsten Zeit für die Berechnung. Tab. 9 ergänzt diese Erkenntnisse zum MVS-Vergleich in übersichtlicher Weise. Der darauffolgende Abschnitt geht auf Detailerkenntnisse des MVS-Vergleichs der einzelnen Szenen ein.

Tab. 9: Zusammenfassung der wichtigsten Resultate und Erkenntnisse des Vergleichs der MVS-Tools. Zur Erläuterung des Tabelleninhalts wird auf die untenstehende Tabellenlegende verwiesen.

Kriterien	**SURE**	**COLMAP**	**PMVS**
Punktmenge	E: 7'212'690 B: 5'530'526 C: 8'891'132 M: 25'236'429	E: 12'921'428 B: 8'544'771 C: 14'887'627 M: 27'858'406	E: 775'443 B: 1'174'864 C: 762'936 M: 1'098'145
Zeit [h:m]	E: 1:24 B: 0:27 C: 1:30 M: 1:52	E: 3:45 B: 3:13 C: 7:29 M: 3:48	E: 0:29 B: 0:17 C: 0:30 M: 1:07
Punktdichte	homogene Dichte über ganzen Bereich	hohe Dichte bei Strukturen geringe-mittlere Dichte bei homogenen Oberflächen	stark lückenhaft
Problemzonen	freistehende Objekte: (Kandelaber, Fahrleitungsmast, Baumstamm)	homogene und/oder repetitive Strukturen (Fenster, Belag, Fassade, Fahrbahn, Schattenwurf)	homogene und/oder repetitive Oberflächenstrukturen
keine oder geringe Problemzonen	- Filterung bewegter Objekte - homogene oder repetitive Strukturen	- Filterung bewegter Objekte - freistehende Objekte	- Filterung bewegter Objekte
Streuung (Genauigkeit)	mittel, wenig Ausreisser	gering, z.T. grobe Ausreisser	{mittel}
Kantenschärfe	mittel	hoch	{mittel}
CPU / GPU	**CPU** & GPU	CPU & **GPU**	**CPU**
Lizenz	kommerzieller Vertrieb	GNU General Public License 3	GNU General Public License 3
Kosten	- Lizenzkosten - Prozessierungskosten (GPU-Prozessierung löst höhere Kosten aus) - Speicherbedarf	- Prozessierungskosten (GPU-Prozessierung löst höhere Kosten aus) - Speicherbedarf	- Prozessierungskosten (CPU-Prozessierung geringere Kosten) - Speicherbedarf

Tabellenlegende

[ ]	Einheitsangabe
{ }	aufgrund der Resultate schwierig zu beurteilen
gelb	gemessene Grössen
orange	visuelle Beurteilung basierend auf Punktwolkenrenderings
violett	visuelle Beurteilung basierend auf Profilen (Vertikalschnitt) durch die Punktwolke
E B C M	vier untersuchte Szenen von Infrastrukturanlagen (E: **E**spark, B: **B**renner-Autobahn, C: FHNW-**C**ampus, M: Strom**m**ast)
fett	verwendete Hardwarekomponente, um Untersuchungsresultate zu prozessieren
Punktmenge	basierend auf Standardparameter
Zeit	basierend auf Standardparameter
GNU 3	www.gnu.org/copyleft/gpl.html, www.gnu.org/licenses/quick-guide-gplv3

Insbesondere bei der Szene des Industriegebäudes *Espark* (vgl. Abb. 4 & Abb. 5) ist die homogene Dichte von SURE, im Gegensatz zu COLMAP und PMVS, ersichtlich. Vor allem die Solarpanels auf dem Dach, die vertikale Fassade und die Glaskonstruktionen werden von SURE erstaunlich gut rekonstruiert. COLMAP hat Mühe mit der homogenen vertikalen Frontfassade, dem Belag und den Fenstern. PMVS liefert bei homogenen Oberflächenstrukturen und Glaselementen keine rekonstruierten Punkte. In den Profilen lässt sich die grosse Streuung bei homogenen Oberflächen mit SURE, sowie die geringe Streuung von COLMAP und die damit verbundene Kantenschärfe erkennen. Bei der Szene der *Brenner* Autobahn ist auffallend, dass COLMAP im Vergleich zu SURE eine deutlich dichtere Punktwolke von der Wiese mit starken Strukturen liefert. SURE hat des Weiteren im Gegensatz zu COLMAP Mühe, eine Autobrücke, Kandelaber und kahle Baumstämme zu rekonstruieren. Für den Bereich am Perimeterrand beim *Brenner* liefert SURE keine und COLMAP durchaus plausible Resultate. Dies hängt wohl mit den unterschiedlich aufgelösten Bildern zusammen, da der Randbereich nur von den niedrig aufgelösten Bildern abgedeckt wird. Alle drei Tools filtern bewegte Objekte wie Autos heraus. Anders als bei der *Brenner* Szene weist COLMAP in den Randbereichen des *FHNW-Campus* lückenhafte Resultate auf. Die hauptsächlich auf den klassischen Luftbildfall ausgerichtete Software SURE generiert für diese Szene klar die besten Resultate. Der Unterschied ist besonders in Schattenwurfbereichen, Strassenbereichen (Belag) und am Perimeterrand erkennbar. Mit der Glasfassade des Campus und dem in der Höhe stark versetzten sowie mit vielen Objekten versehenen Baustellendach haben wie erwartet alle Algorithmen Mühe. Beim *Strommast* liefern SURE und COLMAP sehr ähnliche Resultate. Einzig die teils vom Masten verdeckten Bereiche kann SURE im Gegensatz zu COLMAP nicht rekonstruieren. PMVS extrahiert vor allem den Mast, die Hintergrundszene schlecht bis gar nicht. Es lässt sich zusammenfassend sagen: COLMAP – geringe Streuung mit hoher Kantenschärfe, bei homogenen Flächen wenig Rekonstruktionen, SURE – geringe Kantenschärfe bei mittlerer Streuung, extrahiert jedoch homogene Flächen am besten, PMVS – lückenhafte Resultate.

Abb. 4: Lage des extrahierten Profils in Abb. 5 rechts

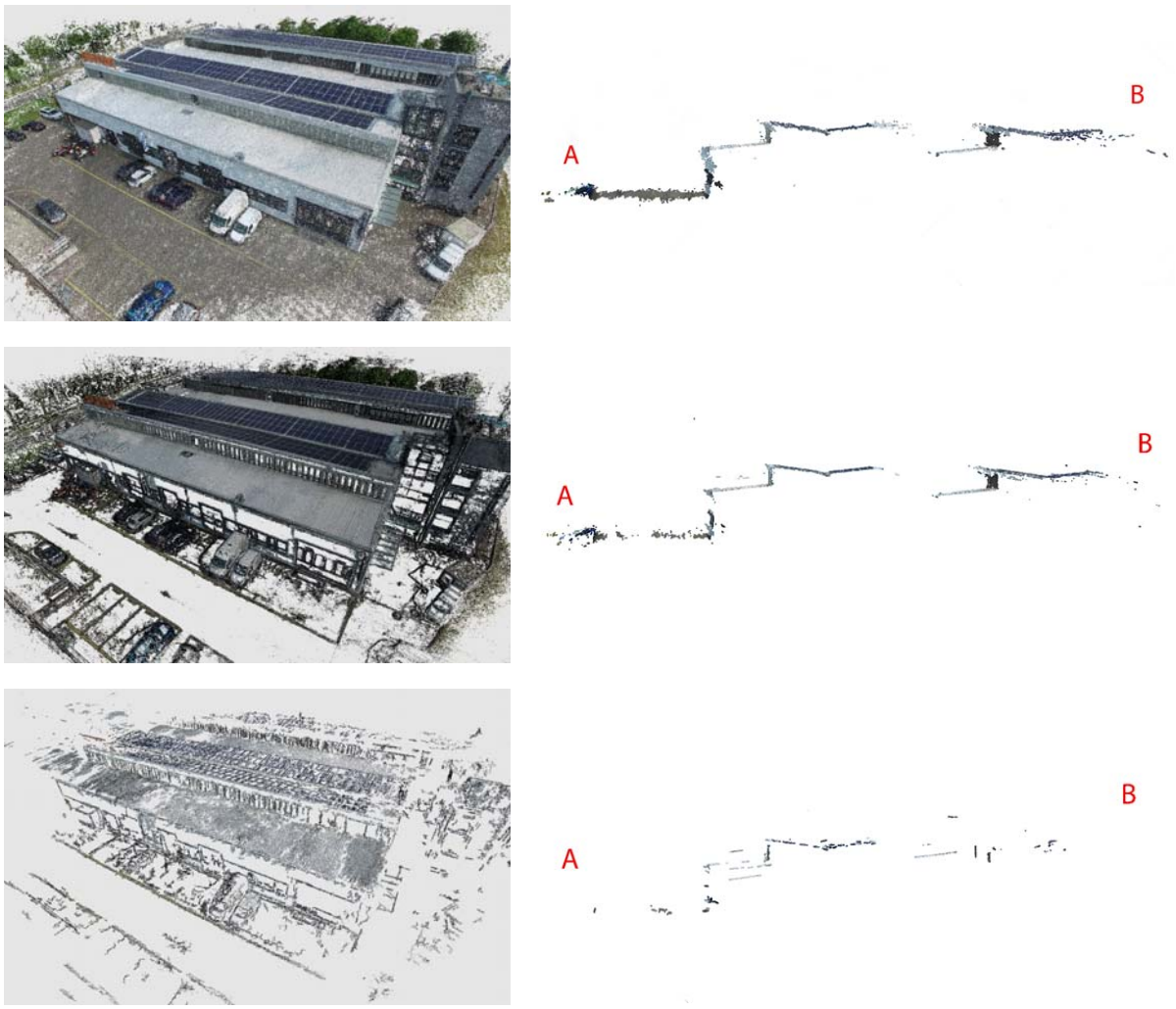


Abb. 5: Beispielhafte Beurteilungsgrundlagen für die Szene *Espark*. Punktwolkenrendering (links) und Profile (rechts) von SURE (oben), COLMAP (Mitte) sowie PMVS (unten).

5 Fazit und Ausblick

In diesem Beitrag wurden die drei MVS-Algorithmen SURE, COLMAP und PMVS zur Integration in einem Gesamtworkflow evaluiert. Dafür wurden die drei Tools mit dem implementierten Workflow, welcher die vollautomatische Prozessierung von Punktwolken, Renderings und Profilen ermöglicht, verglichen. Dieser Workflow lässt ebenfalls die Untersuchung der Steuerparameter zu, welche für jeweils vier oder fünf Parameter pro MVS-Algorithmus umgesetzt wurde. Die hauptsächlich pragmatische visuelle Beurteilung der Resultate ist sehr effizient – insbesondere bei den 96 zu untersuchenden Punktwolken von vier unterschiedlichen Infrastrukturszenen – verunmöglicht jedoch präzise quantitative Aussagen wie Differenzen oder Standardabweichungen.
Der Vergleich der MVS-Tools zeigt, dass COLMAP eine geringe Streuung und hohe Kantenschärfe aufweist, jedoch homogene Flächen nur spärlich rekonstruiert. SURE weist dagegen eine geringe Kantenschärfe bei mittlerer Streuung auf, rekonstruiert aber homogene Oberflächen ohne grobe Ausreisser am besten. PMVS weist lückenhafte Punktwolkenresultate auf. Die Parameteruntersuchung zeigt, dass die Standardparameter bereits gute Resultate liefern. Diese lassen sich jedoch mit minimalen Anpassungen je nach Szene verbessern. Minimale Unterschiede der Steuerparameterauswirkungen liessen sich besser anhand von Differenzbildern beurteilen.
Die effiziente Handhabung deutlich grösserer Bildformate mit der Möglichkeit zur massiven Skalierung bzw. Parallelisierung sämtlicher Prozessschritte ist nicht trivial, wie bereits von FURUKAWA et al. (2010) festgehalten. Eine Abhilfe dafür könnte ein Clustering des Projekts schaffen. Lineare Objekte eignen sich beispielsweise dazu. Die Skalierung erfordert neben grosser RAM-Kapazität ebenfalls grosse Speicherkapazität, da die MVS-Tools grosse Datenmengen erzeugen. Gerade COLMAP benötigt für die Prozessierung der gleichen Projekte mehr RAM. Die gemessenen Zeiten sind stark abhängig von dem zur Verfügung stehenden RAM bzw. der Leistung des Computers. Aufgrund der Erkenntnisse durch die Untersuchungen wird für klassische Luftbildfälle oder bei Verwendung von Tiefenkarten der Einsatz von SURE empfohlen, hauptsächlich begründet durch die homogenen Resultate und kurze Berechnungsdauer. Liegt der Fokus einer Aufnahme auf sich abgrenzende und strukturierte Objekte wie Strommasten, Kandelaber oder ähnliches, wird der Einsatz von COLMAP empfohlen.
Bei den untersuchten Datensätzen lagen keine Kamerakalibrierungen (innere Orientierungen) vor. Es wird erwartet, dass die Qualität der Resultate bei kalibrierten Kameras steigt. Allerdings ist eine konstant zu haltende Kammerakonstante für Aufnahmen mit stark unterschiedlichen Objektdistanzen durchaus eine Herausforderung, um immer noch scharfe Bilder aufzunehmen.

6 Dank

Das Forschungsprojekt *INFRA BirdView* wird vom Forschungsfonds Aargau (FFAG) finanziell unterstützt.

7 Literaturverzeichnis

BURKHARD, J., CAVEGN, S., BARMETTLER, A. & NEBIKER, S., 2012: Stereovision Mobile Mapping: System Design and Performance Evaluation. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **39**(B5), 453-458.

CAVEGN, S., HAALA, N., NEBIKER, S., ROTHERMEL, M. & TUTZAUER, P., 2014a: Benchmarking High Density Image Matching for Oblique Airborne Imagery. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **40**(3), 45-52.

CAVEGN, S., NEBIKER, S. & DEUBER, M., 2014b: Dense Image Matching mit Oblique Luftbildaufnahmen – Ein systematischer Vergleich verschiedener Lösungen mit Aufnahmen der Leica RCD30 Oblique Penta. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 23, Seyfert, E., Gülch, E., Heipke, C., Schiewe, J., Sester, M. (Hrsg.).

CAVEGN, S. & HAALA, N., 2016: Image-Based Mobile Mapping for 3D Urban Data Capture. Photogrammetric Engineering & Remote Sensing, **82**(12), 925-933.

EUGSTER, H., GOTTSMANN, F., KÄSER, C., NEBIKER, S. & KOCH, R., 2013: 3DRoadTV-Service – Georeferenzierter 3D-Bilddatendienst für die Infrastrukturpflege und -verwaltung. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 22, E. Seyfert (Hrsg.), 33. Wissenschaftlich-Technische Jahrestagung der DGPF, Dreiländertagung D - A - CH, 27. Februar – 1. März 2013 in Freiburg i. B., Tagungsband auf CD-ROM, 344-353.

FURUKAWA, Y., CURLESS, B., SEITZ, S. M. & SZELISKI, R., 2010: Towards Internet-Scale Multi-View Stereo. IEEE Conference on Computer Vision and Pattern Recognition, 1434-1441.

FURUKAWA, Y. & PONCE, J., 2010: Accurate, Dense, and Robust Multiview Stereopsis. IEEE Transactions on Pattern Analysis and Machine Intelligence, **32**(8), 1362-1376.

GEIGER, A., LENZ, P. & URTASUN, R., 2012: Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite. IEEE Conference on Computer Vision and Pattern Recognition, 3354-3361.

GERKE, M., NEX, F. & JENDE, P., 2016: Co-Registration of Terrestrial and UAV-Based Images - Experimental Results. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **40**(3/W4), 11-18.

HAALA, N., 2014: Dense Image Matching Final Report. EuroSDR Publication Series, Official Publication No., **64**, 115-145.

HAALA, N. & CAVEGN, S., 2015: Benchmark zur Evaluation dichter Bildzuordnungsverfahren in Luftbildern. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 24, T. Kersten (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 16.-18. März 2015 an der Universität zu Köln, Tagungsband auf CD-ROM, 244-253.

HAALA, N. & CAVEGN, S., 2016: High Density Aerial Image Matching: State-of-the-Art and Future Prospects. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **41**(B4), 625-630.

HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, **30**(2), 328-341.

KAMER, K., EUGSTER, H. & NEBIKER, S., 2013: Integrierte Georeferenzierung von luft- und bodengestützten Stereobilddaten im Strassenbereich: Strategien und Ergebnisse. Geomatik Schweiz, **111**(2013/9), 496-501.

KITWARE INC., 2018: VTK - The Visualization Toolkit. https://www.vtk.org (31.01.2018).

MOULON, P., HIESTAND, R. & BOUDRAND, É., 2011: CMVS-PMVS2. GitHub. https://github.com/pmoulon/CMVS-PMVS (31.01.2018).

NEBIKER, S., CAVEGN, S., EUGSTER, H., LAEMMER, K., MARKRAM, J. & WAGNER, R., 2012: Fusion of Airborne and Terrestrial Image-Based 3D Modelling for Road Infrastructure Management - Vision and First Experiments. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **39**(B4), 79-84.

NEBIKER, S., KAMER, K. & EUGSTER, H., 2013: Integrierte Georeferenzierung von luft- und bodengestützten Stereobilddaten im Straßenbereich – Strategien und Ergebnisse. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 22, E. Seyfert (Hrsg.), 33. Wissenschaftlich-Technische Jahrestagung der DGPF, Dreiländertagung D - A - CH, 27. Februar – 1. März 2013 in Freiburg i. B., Tagungsband auf CD-ROM, 499-508.

NEBIKER, S., CAVEGN, S. & LOESCH, B., 2015: Cloud-Based Geospatial 3D Image Spaces—A Powerful Urban Model for the Smart City. ISPRS International Journal of Geo-Information **4** (4), 2267-2291.

NEITZEL, F. & KLONOWSKI, J., 2011: Mobile 3D Mapping with a Low-Cost UAV System. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(1/C22), 39-44.

NEX, FRANCESCO & REMONDINO, F., 2014: UAV for 3D Mapping Applications: A Review. Applied Geomatics, **6**(1), 1-15.

NEX, F., GERKE, M., REMONDINO, F., PRZYBILLA, H.-J., BÄUMKER, M. & ZURHORST, A., 2015: ISPRS Benchmark for Multi-Platform Photogrammetry. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences **2** (3/W4), 135-142.

PUENTE, I., GONZÁLEZ-JORGE, H., MARTÍNEZ-SÁNCHEZ, J. & ARIAS, P., 2013: Review of Mobile Mapping and Surveying Technologies. Measurement, **46**(7), 2127-2145.

REMONDINO, F., BARAZZETTI, L., NEX, F., SCAIONI, M. & SARAZZI, D., 2011: UAV Photogrammetry for Mapping and 3D Modelling - Current Status and Future Perspectives. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(1/C22), 25-31.

REMONDINO, F., SPERA, M. G., NOCERINO, E., MENNA, F. & NEX, F., 2014: State of the Art in High Density Image Matching. The Photogrammetric Record **29** (146), 144-166.

REMONDINO, F., NOCERINO, E., TOSCHI, I. & MENNA, F., 2017: A Critical Review of Automated Photogrammetric Processing of Large Datasets. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **42**(2-W5), 591-599.

ROTHERMEL, M., WENZEL, K., FRITSCH, D. & HAALA, N., 2012: SURE: Photogrammetric Surface Reconstruction from Imagery. Proceedings LC3D Workshop.

SCHARSTEIN, D., HIRSCHMÜLLER, H., KITAJIMA, Y., KRATHWOHL, G., NEŠIĆ, N., WANG, X. & WESTLING, P., 2014: High-Resolution Stereo Datasets with Subpixel-Accurate Ground Truth. GCPR 2014, Pattern Recognition, Springer, Cham, 31-42.

SCHÖNBERGER, J. L. & FRAHM, J.M., 2016: Structure-from-Motion Revisited. IEEE Conference on Computer Vision and Pattern Recognition, 4104-4113.

SCHÖNBERGER, J. L., ZHENG, E., FRAHM, J.M. & POLLEFEYS, M., 2016: Pixelwise View Selection for Unstructured Multi-View Stereo. Computer Vision – ECCV 2016, Springer, Cham, 501-518.

SCHÖPS, T., SCHÖNBERGER, J. L., GALLIANI, S., SATTLER, T., SCHINDLER, K., POLLEFEYS, M. & GEIGER, A., 2017: A Multi-View Stereo Benchmark with High-Resolution Images and Multi-Camera Videos. IEEE Conference on Computer Vision and Pattern Recognition, 3260-3269.

SEITZ, S. M., CURLESS, B., DIEBEL, J., SCHARSTEIN, D. & SZELISKI, R., 2006: A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. IEEE Conference on Computer Vision and Pattern Recognition, 519-528.

STANEK, H. & EUGSTER, H., 2017: 3D-Bildbasierte Infrastrukturplattformen – Eine Neufokussierung aus Anwendersicht. 19. Internationale Geodätische Woche Obergurgl 2017, Herbert Wichmann Verlag, Berlin/Offenbach.

WENZEL, K., ROTHERMEL, M., HAALA, N. & FRITSCH, D., 2013: SURE – The ifp Software for Dense Image Matching. D. Fritsch (Hrsg.) Photogrammetric Week ’13, Wichmann, Berlin & Offenbach, 59-70.

ZHENG, E., DUNN, E., JOJIC, V. & FRAHM, J.M., 2014: PatchMatch Based Joint View Selection and Depthmap Estimation. IEEE Conference on Computer Vision and Pattern Recognition, 1510-1517.

Blended Learning in Photogrammetrie und Bildverarbeitung für regionale KMUs – Ein Erfahrungsbericht

ANSGAR BRUNN[1]

Zusammenfassung: Blended-Learning etabliert sich immer stärker als anerkannte Lehr- und Lernform. Die Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt (FHWS), Labor für Photogrammetrie hat von 2015-2017 (Projektlaufzeit 26 Monate) das Projekt „Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen (DiPhoBi4KMU)" mit dem Ziel der Stärkung regionaler kleiner und mittlerer Unternehmen (KMU) durchgeführt, einerseits durch Wissensvermittlung, andererseits durch Netzwerkbildung. Erstmalig verfolgte der Studienbereich Geo an der FHWS den Ansatz des Blended-Learnings in einem Projekt mit regionalen Unternehmen. Ergebnisse und Erfahrungen werden sowohl aus der technisch-fachlichen als auch aus der didaktischen Sicht präsentiert und bewertet. Nach dem erfolgreichen Abschluss des Projekts DiPhoBi4KMU wird die FHWS das Projektformat in weiteren Themengebieten weiter fortführen.

1 Einleitung

Die Methoden der Vermessungsingenieure und Geomatiker haben in den letzten Jahren und wenigen Jahrzehnten einen rasanten Wandel erfahren. Nach dem Umbruch durch die sukzessive Einführung digitaler Vermessungsgeräte in den Arbeitsprozess, in den 1990er Jahren vollzieht sich ein Übergang zur berührungslosen Vermessung mit bildgebenden Flächensensoren, der bei der Integration der neuen Verfahren in die bisherigen Vorgehensweisen große Herausforderungen an die in der Praxis stehenden Vermessungsingenieure und Geomatiker stellt.
Zu den neuen Messgeräten gehören sowohl kleine fliegende mit Kameras bestückte Sensorplattformen (Kleinflugzeuge, Kleinhelikopter, Unmanned Aerial Vehicles (UAV) mit montierten Digitalkameras) als auch Tachymeter mit integrierten digitalen Kameras und Multikamera-Rover, die durch Tachymeter getrackt und koordiniert werden.
Die neuen Verfahren erfordern grundlegende Kenntnisse in der digitalen Photogrammetrie und Bildverarbeitung, die bisher nur am Rande der Berufs-, Hochschul- und Universitätsausbildung standen und deren Techniken selbst in den letzten Jahren rasanten Entwicklungen unterworfen waren. Die verlustfreie Einbindung der neuen Verfahren in die bisherigen Messprozesse der Unternehmen steht erst am Anfang. Gerade kleine und mittlere Unternehmen (KMU) stehen hier vor großen Aufgaben, die aufgrund der Personalstruktur und der geringen Mitarbeiterzahl in den Unternehmen allein nur schwer erfolgreich zu lösen sind.
Hier setzt das Projekt „Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen (DiPhoBi4KMU)" an. DiPhoBi4KMU liefert Informations- und Weiterbildungsmaterialien in den engverknüpften, innovativen Themenfeldern der digitalen Photogrammetrie

[1] Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt,
Labor für Photogrammetrie und Fernerkundung, Labor für Laserscanning, Röntgenring 8,
D-97070 Würzburg, E-Mail: Ansgar.Brunn@fhws.de

und digitalen Bildverarbeitung für Praktiker und Entscheider im Berufsfeld der Vermessung und Geoinformatik.
Der Berufsalltag der Praktiker erfordert eine hohe Flexibilität und Anpassungsfähigkeit an die Rahmenbedingungen, z. B. die Auftragssituation und die aktuellen Wetterbedingungen. Das Angebot von DiPhoBi4KMU konnte nur erfolgreich wirksam werden, wenn es die in geringem Maße vorhandenen Freiräume nutzte. Daher wurde das Blended-Learning, eine Kombination aus Online- und Präsenzangeboten, als zentrale Lehr- und Lernmethode eingesetzt.
Das Projekt wurde von Juli 2015 bis August 2017 (26 Monate) durch die Virtuelle Hochschule Bayern innerhalb des ESF der Europäischen Union an der Hochschule für angewandte Wissenschaften (FHWS) gefördert und ist mittlerweile abgeschlossen. Die Ergebnisse und Erfahrungen des Projekts werden in diesem Beitrag präsentiert.
Im folgenden Kapitel werden zunächst die Problemstellung und der Hintergrund des Projekts aufgezeigt und analysiert. Im Abschnitt 3 werden die Ergebnisse unter verschiedenen Perspektiven dargestellt. Dieses ist die Grundlage für eine eingehende Darstellung und Diskussion der aus dem Projekt resultierenden Erfahrungen in Kapitel 4. Der Beitrag schließt mit einem kurzen Fazit und einem Ausblick auf weitere Aktivitäten.

Abb. 1: Bayerische Fördergebietskulisse (Landratsämter mit abnehmender Bevölkerungszahl sind blau dargestellt, Graphik angelehnt an VHB 2016) sowie Verteilung der Projektpartner auf das Landesgebiet (Unternehmen in Grün, FHWS in Orange).

2 Problemstellung

2.1 Hintergrund

Innerhalb des bayerischen ESF-Programms 2014-2020 „Perspektiven in Bayern – Perspektiven in Europa“ ist DiPhoBi4KMU ein Projekt zur „Netzwerktätigkeit zwischen Hochschulen und Unternehmen“. Ziel des Programms ist die Förderung der KMU in Gebieten mit abnehmender

Bevölkerung (Fördergebietskulisse, Abb. 1). Wissensnetzwerke unter der Leitung der Hochschulen sollen entstehen, die während der Projektlaufzeit Wissen aufbauen und über die Projektlaufzeit hinaus Bestand haben. Einzelheiten zu den Projektrahmenbedingungen sind in MEYER & BRUNN 2016 ausgeführt. Der virtuellen Hochschule Bayern (vhb) ist die Projektträgerschaft übertragen, die gleichzeitig die Digitalisierung der Hochschullehre in Bayern im Rahmen der bayerischen Digitalisierungsstrategie (STMWI 2014) verfolgt.

2.2 Projektstruktur

Die Projektgruppe baute die FHWS als Projektleiter vor Beginn der Projektphase mit zehn teilnehmenden Unternehmen auf, vorwiegend Ingenieurbüros der technischen Vermessung. Während der Projektlaufzeit kamen drei weitere Unternehmen (ein UAV-Flugunternehmen, ein Architekturbüro und ein weiteres vermessungstechnisches Ingenieurbüro) kamen hinzu. Motiviert wurde die Erweiterung durch das Interesse der Unternehmen, aufgrund ihrer öffentlichen Wahrnehmung des Netzwerks, und aus dem Wunsch der Kompetenzerweiterung der Gruppe. Geographisch konzentrierten sich die Unternehmen auf den Norden und Nordwesten von Bayern. Alle Unternehmen blieben bis zum Ende der Laufzeit im Projekt (Abb. 2). Das Alter der Teilnehmer reichte von 20 bis 60 Jahren gleichmäßig über alle Lebensaltersdekaden. Der bisher erlangte persönliche Ausbildungsabschluss reichte von einer Berufsausbildung bis zum Universitätsstudium. Die Wissensvermittlung im Rahmen der Projektaktivitäten fand während des Projektzeitraums offline (real) und online mit der Methode des *Blended-Learnings* statt. Die online Wissensvermittlung erfolgte mit einem E-Learning-Angebot auf einem Moodle-Server (MOODLE 2018) der FHWS mit ständiger Verfügbarkeit und frei entscheidbarem Zugriff durch die Teilnehmer.

Abb. 2: Anzahl der teilnehmenden Unternehmen

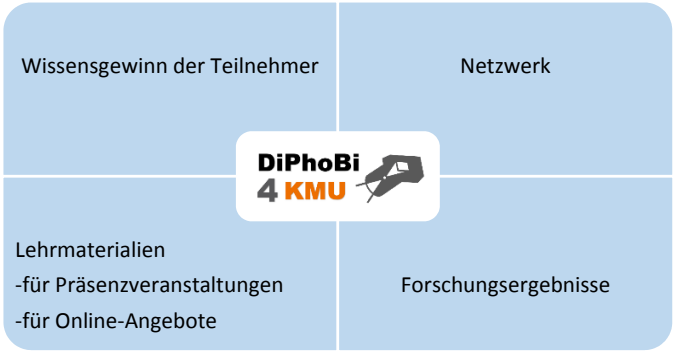


Abb. 3: Ergebnismatrix

Real organisierte die FHWS, nach einem Auftakttreffen, fachliche Präsenztreffen zu den Themen

- Grundlagen der Nutzung von Photogrammetrie und Bildverarbeitung in der terrestrischen Vermessung mit Anwendungsbeispielen,
- Praktische Anwendung von Photogrammetrie und Bildverarbeitung in der Vermessung,
- Anwendungen von Photogrammetrie und Bildverarbeitung und
- UAV-Anwendungen

Alle Präsenztreffen hatten Grundlagen- und Anwendungsanteile.

Das Projekt endete mit einem Abschlussworkshop, der auch der Evaluation diente. Öffentliche Fachveranstaltungen an der FHWS wurden zusätzlich für Treffen der Projektteilnehmer genutzt. Die Präsenztreffen dienten zunächst der Wissensweitergabe von der Hochschule an die Unternehmen. Die Unternehmen beteiligten sich ihrerseits durch die Vorstellung ihrer Erfahrungen, i.d.R. anhand von praktischen Projekten. Ein offener Austausch zwischen allen Teilnehmern prägte die Projekttreffen. Nach dem Abschluss des Projekts wurden an die aktiven Teilnehmer Teilnahmezertifikate vergeben. Neben der Vermittlung und des Austausches von Spezialwissen ist die Netzwerkbildung von großer Bedeutung. Angelehnt an den Begriff *Blended-Learning* wurde ein Konzept des *Blended-Networking* entwickelt, das eine Netzwerkbildung sowohl durch reale Treffen als auch gemeinsame oder gleichgerichtete Online-Angebote ermöglicht. Der Offline-Teil der Netzwerkbildung fand in den realen Treffen zu den Projektworkshops und weiteren Veranstaltungen statt. Förderlich waren die offenen Gespräche bei den Treffen aller Teilnehmer. In dem Online-Teil der Netzwerkbildung wurden Kommunikationsformen über soziale Medien und Online-Aktivitäten mit gegenseitig informativem Charakter genutzt (z. B. offene Abfragen und offene Abstimmungen).

3 Ergebnisse

Das Projekt intendierte sowohl eine Weiterbildung der Teilnehmer als auch eine Netzwerkbildung. In beiden Intentionsbereichen wurden unterschiedliche Aktivitäten durchgeführt. Dadurch sind Projektergebnisse in mehreren Kategorien vorhanden (Abb. 3): Wissensgewinn der Teilnehmer, das Netzwerk, die Lehr- und Lernmaterialien sowie die wissenschaftlichen Ergebnisse. Daneben bleiben Erfahrungen in dem Angebot von Blended-Learning- und Blended-Networking-Angeboten.

3.1 Wissensgewinn der Teilnehmer

Der Wissensgewinn der Teilnehmer war ein wichtiges Ziel des Projekts. Neben dem Wissensgewinn sind die Ebenen Zufriedenheit, Verhaltensänderung und Auswirkungen auf die tägliche Arbeit nicht minder wichtig. Methoden der Wissensmessung sind klassische Prüfungen aber auch moderne Verfahren, die die Selbsteinschätzung der Teilnehmer nutzen (z.B. Vorher-Nachher-Untersuchungen). Der Vorteil der Selbsteinschätzung ist, dass keine Prüfungssituation entsteht. Dennoch bleibt der Wissensgewinn in dem Projekt schlecht qualitativ messbar. In DiPhoBi4KMU erfolgte die Bewertung der eigenen Kenntnisse während des Auftaktworkshops und eine verbale abschließende Beurteilung gegen Ende des Projekts, in der alle Teilnehmer einen deutlichen Wissensgewinn bei sich feststellten und ebenso eine allgemeine Zufriedenheit sowie Auswirkungen auf die tägliche Arbeit identifizierten.
Ergänzend zur Evaluation der FHWS hat die VHB (VHB 2018) eine einige Evaluation durchgeführt. Die Rücklaufquote dieser Evaluation war 33%, 5 von 15 Teilnehmern haben den Fragenbogen zurückgesandt. In Bezug auf die Anzahl der Unternehmen ist die Quote 5 von 13, also ca. 38%. Trotz der schwachen statistischen Situation soll diese Evaluation dennoch im Folgenden genutzt, um Meinungstendenzen der Projektteilnehmer zu beurteilen.
In der VHB-Evaluation beurteilten 60% der Testgruppe den Lernerfolg auf einer Skala „nein, überhaupt nicht“, „eher nicht“, „teils/teils“, „eher ja“ und „ja, sehr viel“ mit „eher ja“ und 40%

den Lernerfolg mit „ja, sehr viel“. Förderlich sahen die Teilnehmer v.a. die theoretischen Grundlagen der Photogrammetrie und Bildverarbeitung („Was passiert überhaupt im Hintergrund der Software?“), Tipps für den Workflow, Kenntnisse über neue Anwendungen, Genauigkeiten und über den Angebotsmarkt. Die Teilnehmer der Umfrage waren zu 80% Hochschulabsolventen (Universität oder Hochschule für angewandte Wissenschaften) und zu 20% Absolventen einer Fachakademie. Die Menge der Lerninhalte wurde allgemein als angemessen oder als eher zu wenig (1x) angesehen.

3.2 Netzwerk

Ein nachhaltiges Netzwerk der Teilnehmer ist für den Erfolg des Projekts sehr wichtig. Unternehmensnetzwerke sind gerade für die KMU notwendige Strukturen, um eine große Angebotsbreite lokal bzw. in ihrer Region an den Markt zu bringen, ohne die dafür notwendige Infrastruktur selbst vorhalten zu müssen. Die mit den neuen Techniken verbundenen massiven Investitionen sind insbesondere für die KMU ein Hinderungsgrund, ihre Angebotspalette zu erweitern (z. B. in der UAV-Vermessung, der Vermessung mit Multikamera-Rovern, dem terrestrischen Laserscanning und auch den flugzeug-basierten Messverfahren wie Luftbildphotogrammetrie und Airborne Laserscanning). Das im Projekt entstandene Netzwerk zeichnet sich durch eine offene fachliche Kommunikation aus. Aber auch die Verteilung der Prozesskette eines Aufnahmeverfahrens auf mehrere Unternehmen, z. B. bei durch die Trennung der Befliegung von der Datenveredlung, sind entstandene Formen der Zusammenarbeit unter den Teilnehmern. Einige Teilnehmer haben die Bedeutung der Zusammenarbeit mit der FHWS für den Alltag erkannt und arbeiten seitdem an der Definition solcher Projekte und anderer Projekte mit der FHWS aktiv mit.

Für die Hochschule ist das Netzwerk ebenfalls von hoher Bedeutung, da zusätzliche Unternehmen für die Zusammenarbeit mit der Hochschule auch für die Lehre gewonnen wurden. So sind aus dem Projekt nicht nur gemeinsame Projekt- und Bachelorarbeiten entstanden, sondern auch neue Praktikumsmöglichkeiten für die Studierenden aufgebaut worden. Die Hochschule wird daher das Netzwerk des Projekts mit den Teilnehmern fortführen und weitere Veranstaltungen der themenspezifischen Kommunikation bieten. Eine Einbindung des Netzwerks in den nationalen und internationalen Kontext ist vorgesehen.

3.3 Lehrmaterialien

Im Rahmen der Wissensvermittlung des Blended-Learning-Angebots sind Lehr- und Lernmaterialien sowohl für Präsenzveranstaltungen als auch Online-Materialien entstanden. In den Präsenzveranstaltungen wurden die Themen vertieft, zu denen vorher den Teilnehmern Online-Dokumente zur Verfügung gestellt wurden.

3.3.1 Lehrmaterial für Präsenzveranstaltungen

Die Präsenzveranstaltungen fanden als für Hochschulen typischer seminaristischer Unterricht statt. Die Wissensvermittlung nutzte neben Vorlesungsanteilen mit intensiver Rückkopplung zu den Teilnehmern moderne didaktische und aktivierende Strategien (z.B. Think-Pair-Share (TPS)). Grundlage bildeten Folienpräsentationen, die später online zur Verfügung gestellt wur-

den und ein Teil des E-Learning-Kurses bleiben. Übungsmaterial für praktische Übungen an Photogrammetriesystemen stehen nun ebenfalls zur Verfügung.

3.3.2 Lehrmaterial für Onlinelehrveranstaltungen

Die nachhaltigen Onlinematerialien teilen sich in die Bereiche Digitale Photogrammetrie und Digitale Bildverarbeitung (s. Abb. 4). Das Diagramm zeigt das Grundgerüst, das orientiert, an den Bedürfnissen und Interessen der Teilnehmer, lokal mit Informationen gefüllt wurde. Der Teil „Digitale Photogrammetrie“ beinhaltet die wichtigsten Grundlagen und aktuellen Verfahren der modernen Digitalen Photogrammetrie für verschiedenste Anwendungen im Bereich terrestrischer Vermessung.

Der Teil „Digitale Bildverarbeitung“ beinhaltet die wichtigsten Grundlagen und aktuellen Verfahren der modernen Digitalen Bildverarbeitung für verschiedenste photogrammetrische Anwendungen im Bereich terrestrischer Vermessung, sowie Luftbildauswertungen.

Das Informationsangebot an die KMU reicht von der aufbereiteten Darstellung der theoretischen Grundlagen bis hin zu erläuterten Praxisbeispielen. Die digitalen Angebote werden durch Teile der eigenen Wissenskontrolle ergänzt. Auf dieser Basis werden Handlungsabläufe dargestellt, die eine reibungsfreie Integration der neuen Techniken in die tägliche Arbeit ermöglichen.

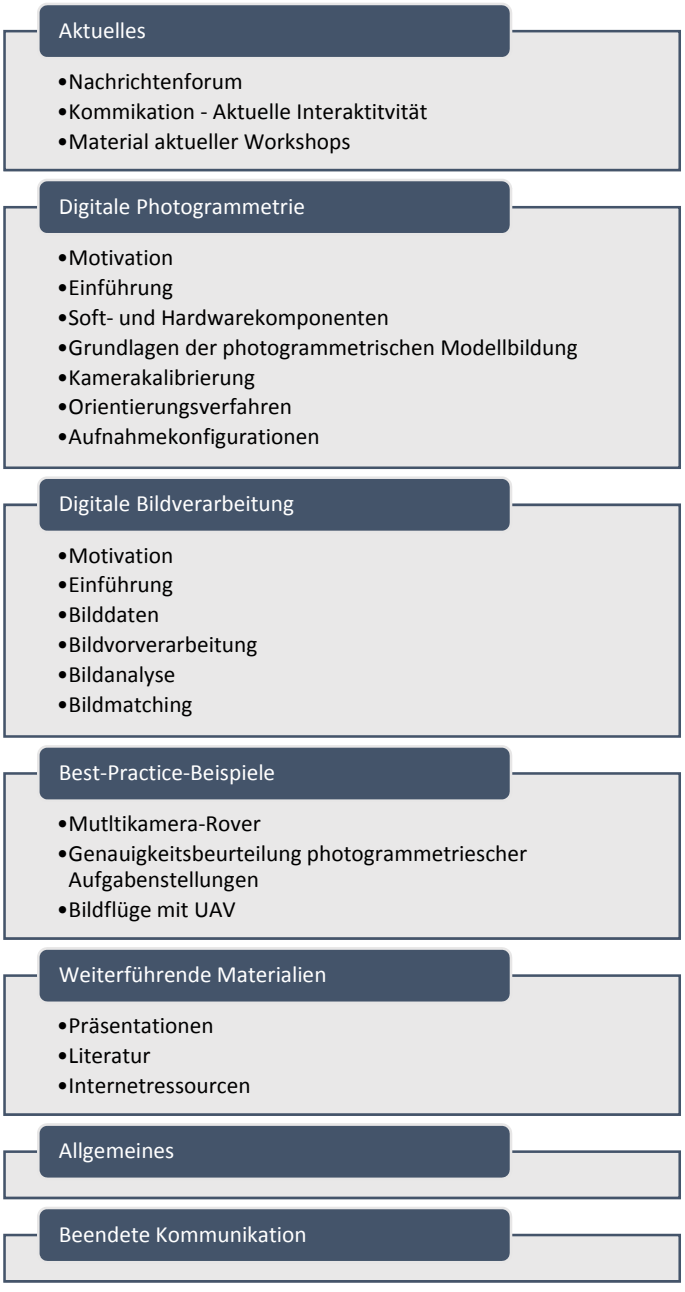


Abb. 4: Struktur des E-Learning-Kurses.

Das entstandene Lehrmaterial soll in weiteren Schulungsangeboten (z. B. in der Weiterbildung), unabhängig von DiPhoBi4KMU, eingesetzt und zu diesem Zwecke auch beständig erweitert und aktualisiert werden. Einige Teile werden in die studentische Ausbildung einfließen.

3.4 Wissenschaftliche Ergebnisse

Innerhalb des Projekts wurden, durch die Teilnehmer motiviert, einige wissenschaftliche Untersuchungen durchgeführt, die teilweise auch publiziert wurden. Zu Beginn des Projekts stand eine Marktanalyse für UAV. Die Kalibrierung und das Genauigkeitspotentials eines Multikamera-Rovers wurden untersucht (BRUNN & MEYER 2016). Darüber hinaus wurden Aufnahmekonfigurationen von UAV-Befliegungen analysiert und verschiedene Softwareprodukte und deren Einfluss auf das Auswerteergebnis (z. B. 3D-Punktwolke und digitales Höhenmodell) getestet. Beispiele für Anwendungen wurden mit den Partnern realisiert (z. B. KARL & BRUNN 2016). Die fachlichen Details sind nicht Thema dieses Beitrags und finden sich in den entsprechenden Publikationen.

4 Erfahrungen

4.1 Die Teilnehmersicht

Die Rückmeldungen der Teilnehmer während des Projektverlaufs waren sehr positiv. Die positive Wirkung der Netzwerktreffen wurde herausgehoben. Gemeinsame Projekte der Teilnehmer, außerhalb der Förderung, sind entstanden und wurden gemeinsam erfolgreich durchgeführt. Die durch die FHWS aufbereiteten Fachinformationen helfen den Unternehmen, ihr Verständnis für die täglich angewandten Sensoren und Algorithmen zu erweitern. Gleichzeitig berichten die Unternehmen bereits von einer Steigerung ihrer Marktpräsenzen in und außerhalb von Bayern.
Im Rahmen des Abschlussworkshops fand eine mündliche Evaluationsrunde statt. Insgesamt äußerten sich die Teilnehmer sehr zufrieden mit dem Projekt und der durch das Projekt vorhandenen Plattform des Austausches. Gerade die Verknüpfung der aufbereiteten theoretischen Grundlagen mit den praktischen Anwendungen und Einzelprojekten wurde als der richtige Weg erachtet, diese neuen Techniken in den täglichen Arbeitsprozess zu integrieren. Mit Bedauern wird das Ende der Projektlaufzeit aufgenommen. Es ergeht der dringende Wunsch der Teilnehmenden an die FHWS, die Plattform fortzuführen.
Die Evaluation der VHB (VHB 2018) bestätigt diesen Eindruck und zeigt konkretisierend, dass

- ein Großteil der Teilnehmer bereits vor dem Projekt Erfahrung mit Online-Lernangeboten hatte,
- der Zeitaufwand für die Teilnehmer unter 30 Minuten pro Woche lag und dass
- die Bearbeitung teilweise in der Arbeitszeit aber auch während der Freizeit erfolgte.

Offenbar haben an dem Projekt, Unternehmen und Mitarbeiter teilgenommen, die aktiv mit den zur Verfügung stehenden Medien versuchen, ihre Fachkompetenz zu erweitern und ihre Marktpräsenzen ihrer Unternehmen zu stärken, trotz einer Vollauslastung in der aktuellen Auftragslage. Das Projekt DiPhoBi4KMU ist dabei nur eine der genutzten Möglichkeiten. Aufgrund der Technikaffinität der Teilnehmer (aus Vermessung und Geoinformatik) ist die Nutzung von Online-Weiterbildungsangeboten scheinbar selbstverständlich.
Durch die Evaluation wird bestätigt, dass die Teilnehmer sich in einem hohen Maße in den Kurs eingebunden gefühlt haben und mit der Kommunikation zufrieden oder sehr zufrieden waren. Deutlich wird hier die hohe Erwartungshaltung an das Ziel der Netzwerkbildung. Gelobt wurde die kleine Gruppengröße, die Kommunikation erleichtert hat. Dennoch werden mehr Kommuni-

kation und mehr Treffen gewünscht. Die Bedeutung des Teilaspekts Online-Weiterbildung selbst wurde von den Teilnehmern indifferent gesehen. Vor- und Nachteile gleichen sich hier aus. Vorteile sind aus Sicht der Teilnehmer:

- Die Unterlagen sind direkt abrufbar (24/7), mit der daraus resultierenden Zeitersparnis (Anfahrt, Parkplatzsuche...).
- Die Inhalte können vertieft oder auch nachgeholt werden, falls das Vorort-Treffen nicht wahrgenommen werden konnte.
- Zeiten für Weiterbildung können nur so in den Arbeitsablauf und Arbeitsplanung reibungsfrei eingebunden werden, auch wenn und weil es Zeiten gibt, bei denen die eigentliche Arbeit im Unternehmen Vorrang hat.

Als Nachteil werden die notwendige Selbstdisziplin und Selbstorganisation erachtet. Bei Präsenzveranstaltungen ist der Austausch mit Kursteilnehmern "vor Ort" natürlicher und intensiver. Häufig wurde das Wissen an andere Kollegen weitergegeben (Multiplikatoreffekt), in dem Kursunterlagen weitergereicht, Hilfe bei Fragestellungen gegeben wurden oder einfach ein Austausch von erlangtem Wissen bei praktischen Anwendungen erfolgte. Dreiviertel der vermittelten Theorie sind als Basiswissen direkt einsetzbar. 80% der vorgestellten Fallbeispiele sind in die eigene Arbeit übertragbar. Erfahrungen des Workflows und teilweise auch Fehler müssen nicht selbst gemacht werden. Die Teilnehmer werden die Kursmaterialien auch weiterhin nutzen.

4.2 Die Hochschulsicht

Die Hochschule hat Erfahrungen sowohl in der Wissensvermittlung im Rahmen der Weiterbildung, in Erstellung von E-Learning-Materialien im Kontext des Blended-Learnings und in der Netzwerkbildung mit Auswirkungen auf die Hochschullehre und folgende Projektaktivitäten gewonnen.

In dem Projekt sind in einer kleinen Gruppe technikaffine Fachkollegen mit unterschiedlicher Vorbildung und unterschiedlichen Stufen der Berufsausbildung zusammengekommen. Wichtig für den Erfolg einer solchen Maßnahme ist das persönliche Interesse der Teilnehmer am Thema und an ihrer eigenen persönlichen Entwicklung, da nur so im Sinne der Weiterbildung die Prioritäten gesteuert werden. Das Interesse muss durch aktivierende Lehr- und Lernmethoden aufrechtgehalten werden. Das Blended-Learning bietet dazu hervorragende Möglichkeiten, da im E-Learning-Angebot eine Binnendifferenzierung entsprechend der persönlichen Fähigkeiten der Teilnehmer und gleichzeitig durch die Präsenzveranstaltungen ein direkter Kommunikationsweg zur Vertiefung und Klärung mit den Lehrern und Kollegen möglich ist.

Viele der bisher erstellten Materialien folgen der Idee eines Online-Buches für den Praktiker, die aufgrund der Fülle des gesetzten Themengebiets gewählt wurde. Zur Reduktion von Zugangsbarrieren werden die einzelnen Themen abschnittsweise auf der Hauptseite verlinkt. Das Konzept ist von den Teilnehmern positiv aufgenommen worden. Dennoch bietet sich für zukünftige Weiterbildungsprojekte an, die Interaktivität in den Bereichen Wissenskontrolle zu erhöhen. Dazu kann die Dokumentenstruktur zugunsten eines Wikis und Glossars aufgelöst werden, um so noch einen weiteren Schritt auf dem didaktischen Weg „Shift from teaching to learning" (BACHMANN 2014) zu gehen. Die Sammlung von Best-Practice-Beispielen folgt bereits jetzt diesem didaktischen Konzept.

Die Netzwerkbildung ist für die Teilnehmer von größter Bedeutung. Traditionell verfolgten die Hochschulen in den letzten Jahrzehnten primär die Wissensvermittlung. Seit wenigen Jahrzehnten widmen sie sich verstärkt auch der Forschung. Die FHWS hat in den letzten Jahren die Weiterbildung mit dem Campus Weiterbildung beständig ausgebaut. Soziale und fachliche überregionale Netzwerke bilden bis dato die Berufsvertretungen. Zur Erfüllung des regionalen Bedarfs an fachlicher Netzwerkbildung sind zusätzlich von den Hochschulen Impulse in die Region gewünscht und erforderlich, dessen Herausforderungen sich die Hochschulen stellen sollten. Nutznießer sind alle, die Unternehmen unabhängig von einer eventuellen Konkurrenzsituation am Markt.
Für die Netzwerkbildung sind die persönlichen „Vorort" Treffen an der Hochschule sehr wichtig. Daher sollte das Einzugsgebiet trotz aller Digitalisierung über das Internet eher regional als überregional sein. Innerhalb des Projekts hat sich gezeigt, dass die Integration von Unternehmen in großer Distanz innerhalb von Bayern sich schwieriger gestaltete. Um diesen Unternehmen mehr Integrationsmöglichkeiten zu geben, bietet sich das „Blended-Networking" an, in dem mehr Kommunikationsangebote online zur Verfügung stehen. Dazu müssen die technischen Möglichkeiten der E-Learning-Plattform deutlich ausgeweitet werden (z.B. Videochats, Diskussionsforen, Blackboards).
Die FHWS will in wenigen Jahren seinen Studierenden 30% der Lehrveranstaltungen digital/online anbieten. Die Projekterfahrungen zeigen, dass Veranstaltungen, die alleine das Konzept des E-Learnings nutzen, zu kurz greifen. Das Blended-Learning in Kombination mit dem Blended-Networking erfüllt auch für die Vollzeitstudenten die Bedürfnisse in besserem Maße.
Die Erfahrungen aus diesem Projekt werden in der Literatur bestätigt. SCHUTTER et. al. (2017) sieht durch Blended-Learning eine Möglichkeit der Überwindung von räumlichen und zeitlichen Grenzen und dadurch auch einer Expansion von Lernzeiten. Vorlesungsaufzeichnungen, asynchrone Kommunikationsmedien, z.B. Foren und Wiki-Umgebungen sowie selbst organisierte Lernprozesse und Nutzung von Kleingruppen zur (kollegialen) Onlineberatung sind die einzusetzenden Techniken. Virtuelle Labore, virtuelle Exkursionen und virtuelle Anwendungsfälle sind die direkte Konsequenz. Eine Verkopplung des Aufenthaltsorts mit dem (online/virtuellen) Lernort vereinfacht die Komplexität der Lernorganisation. GOTTSCHALK et. al. (2017) sind durch die Digitalisierung grundsätzlich eine Veränderung der Eingangskompetenzen der Studienanfänger und gleichzeitig auch veränderte Anforderungen an die Beschäftigungsbefähigung (Employability).
Die Hochschulen müssen sich veränderten und neuen Aufgaben stellen, u. a. der Lehre im Kontext der allgemeinen Digitalisierung der Gesellschaft und der fachlichen Koordinatorenfunktion in der regionalen Wirtschaft, auch nach dem Studium.

5 Fazit und Ausblick

Die Erfahrungen in der Durchführung und der Kommunikation mit den Teilnehmern des Projekts „DiPhoBi4KMU" stehen in diesem Beitrag im Vordergrund. Dazu wurden die Ergebnisse und Erfahrungen zusammengestellt.
Im Rahmen der Digitalisierungsstrategien auf verschiedenen Ebenen der Gebietskörperschaften, des Bundes, der Länder und Kommunen, den Hochschulen und in der Industrie konnten wertvolle Erkenntnisse für zukünftige Angebote gesammelt werden, u.a., dass die Kombination von E-

Learning-Angeboten und von Präsenzangeboten richtig und wichtig ist und dass die Netzwerkbildung gerade durch Präsenzveranstaltungen entsteht, die zukünftig durch ein Blended-Networking gestärkt werden sollten. Um die Teilnahme an Präsenzveranstaltungen überhaupt möglich zu machen, sind daher eher regionale als national oder internationale Gruppen für Blended-Learning/Networking-Angebote zu wählen, da Präsenszeiten und die damit verbundenen Reisezeiten auch zunächst verfügbare Produktivzeit reduzieren und so im Konflikt zum Tagesgeschäft der Unternehmen stehen.
Das Format dieses Weiterbildungsangebots hat sich als erfolgreich herausgestellt. Die FHWS wird Projekte dieser Art, mit einer stärkeren Fokussierung auf das Blended-Networking, weiterverfolgen.

6 Danksagung

Dieser Beitrag wurde innerhalb des Projekts DiPhoBi4KMU durch die Europäische Union und den Freistaat Bayern im Rahmen des ESF-Programms 2014-2020 „Perspektiven in Bayern – Perspektiven in Europa“ in Zusammenarbeit mit der Virtuellen Hochschule Bayern (vhb) unterstützt.

7 Literaturverzeichnis

STMWI, 2014: Zukunftsstrategie Bayern Digital, Bayerisches Staatsministerium für Wirtschaft und Medien, Energie und Technologie, https://www.bayern.de/wp-content/uploads/2014/09/2015-07-27-Zukunftsstrategie-BAYERN-DIGITAL.pdf (letzter Zugriff: 15.1.2018).

BRUNN, A., 2017: Abschlussbericht Förderrunde 1/2014-2020 Projekt: Digitale Photogrammetrie und Bildverarbeitung für KMU (DiPhoBi4KMU), Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt.

BRUNN, A. & MEYER, TH., 2016: Calibration of a multi-camera rover. Int. Arch. Photogramme. Remote Sens. Spatial Inf. Sci., **41**(B5), 445-452, doi:10.5194/isprs-archves-XLI-B5-445-2016.

Gottschalk, A., Karl, M. & Bolsinger, H., 2017: Digitalisierung in der Lehre als Prozess einer strategischen und operativen Organisationsentwicklung, Didaktik-Nachrichten-DiNa, 11/2017, 3-10.

KARL, G. & BRUNN, A., 2016: Erfassung von äußeren Schäden an Windkraftanlagen. VDV-Magazin-Zeitschrift für Geodäsie und Geoinformatik, 1/2016, 32-39.

MEYER, TH. & BRUNN, A., 2016: Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen durch „Blended Learning“. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 25, T. Kersten (Hrsg.), 36. Wissenschaftlich-Technische Jahrestagung der DGPF, Dreiländertagung der DGPF, der OVG und der SGPF in Bern, Tagungsband auf CD-ROM, 436-448.

MOODLE, 2018: Moodle - Von der Community vorangetrieben, weltweiter Support, https://moodle.org (letzter Zugriff: 15.01.2018).

SCHUTTER, S., LÜFTL, K. & LÖHE, J., 2017: Studieren, wo der Mensch zählt. Die neue Hochschule – DNH - Hochschule der Region, 5/2017, 12-15.

VHB, 2016: Netzwerkgestützter Wissenstransfer von Hochschulen in Unternehmen, http://esf.vhb.org/ (letzter Zugriff: 15.01.2018).

VHB, 2018: DiPhoBi4KMU – Ergebnisse der vhb-ESF-Evaluation. Bericht der virtuellen Hochschule Bayern (VHB).

Impulse aus der Praxis –Bereicherung von Forschung und Lehre aus berufsbegleitenden GI-Bildungsangeboten

CHRISTOPH TRAUN[1]

Zusammenfassung: Der Interfakultäre Fachbereich Geoinformatik an der Universität Salzburg bietet neben einem umfassenden Ausbildungsangebot in der Präsenzlehre seit bald 25 Jahren international akkreditierte, berufsbegleitende Geoinformatik-Fernstudien unter der Marke UNIGIS an. Dieser Beitrag geht der Frage nach, wie sich die damit adressierte Zielgruppe der ‚life-long-learners' in die akademische Forschung und Lehre eingliedert.

1 Einführung

Der Fachbereich Geoinformatik an der Universität Salzburg ist bewusst keiner Fakultät zugeordnet. Mit über 80 MitarbeiterInnen „integriert" er jedoch eine Reihe von Fächern unterschiedlichster fakultärer Zuordnung, die Geoinformatik (GI) zur Lösung domänenspezifischer Fragestellungen benötigen. Die damit einhergehende horizontale Breite (von Archäologie bis Zoologie) wird durch eine vertikale Differenzierung der angebotenen Aus- und Weiterbildungsniveaus und den damit verbundenen Zielgruppen um eine zusätzliche Dimension erweitert: Vom primär für Schüler konzipierten, fix installierten „Ideas Lab", über ein umfangreiches Lehrkontingent für sämtliche GI-Belange im Bachelorstudium Geographie, Aktivitäten in der Lehrer-Aus- und Weiterbildung, das Angebot eines international ausgerichteten Masterstudiengangs „Applied Geoinformatics" bis hin zum Doktorratskolleg „GI-Science" werden sämtliche Altersgruppen und Bologna-Niveaus adressiert (BLASCHKE et al. 2011).
Zusätzlich zu diesen Präsenzangeboten wird mit den berufsbegleitenden UNIGIS-Studiengängen (STROBL 2011) eine Zielgruppe angesprochen, die wichtige Beiträge zur Bereicherung von Forschung und Lehre leistet. Studierende im UNIGIS Master- oder Zertifikatsprogramm haben häufig bereits ein Studium in einer Anwendungsdomäne abgeschlossen und sind in der Regel direkt in geoinformatische Arbeitskontexte eingebunden. Vielfach können sie auf mehrjährige GI-Erfahrung zurückblicken, die sektoral durchaus tief und häufig sehr praxisorientiert ausfällt. Damit unterscheiden sie sich wesentlich von typischen Vollzeitstudierenden, die lebenszyklusbedingt kaum facheinschlägige Erfahrung in das Studium mit einbringen.

2 Impulse durch berufsbegleitende Fernstudien

Basierend auf mittlerweile 2500 AbsolventInnen (davon über 1500 im deutschen Sprachraum) der Salzburger UNIGIS Fernstudien lassen sich rückblickend folgende positive Rücklauf-Aspekte für den Fachbereich festmachen:

[1] Universität Salzburg, Interfakultärer Fachbereich Geoinformatik - ZGIS, Hellbrunnerstrasse 34, A-5020 Salzburg, E-Mail: Christoph.Traun@sbg.ac.at

- Sicherung von Qualität und Marktkonformität der Lehre,
- Beiträge zur Weiterführung universitärer Forschungsthemen, sowie eine
- verbesserte Vernetzung mit Wirtschaft und Verwaltung.

Im Folgenden werden diese Aspekte näher ausgeführt und anhand von Fallbeispielen illustriert.

2.1 Sicherung von Qualität und Marktkonformität der Lehre

UNIGIS Studierende sind in der Regel hochgradig motiviert. Das liegt nicht nur an den Kosten und Mühen, die für ein berufsbegleitendes Studium in Kauf genommen werden, sondern auch an der sich bald einstellenden Erfahrung, Gelerntes unmittelbar im Arbeitskontext einsetzen zu können. Gleichzeitig beobachten wir eine erhöhte Erwartungshaltung an die Qualität der Lehre und eine kritische Auseinandersetzung mit Inhalten, nicht zuletzt hinsichtlich Aktualität und Praxiskonformität.

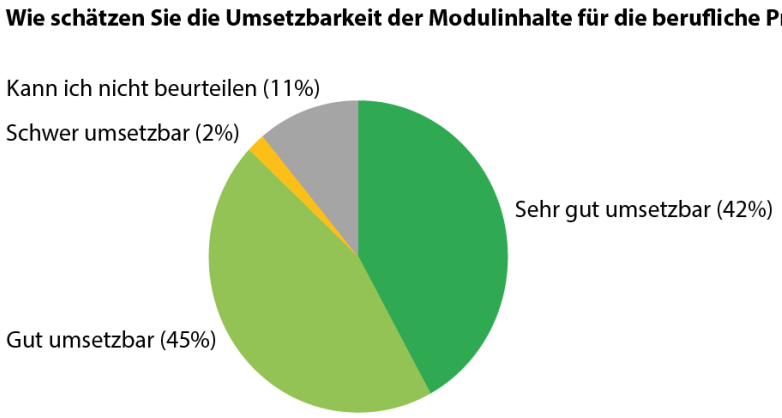


Abb. 1: Aggregierte Evaluationsergebnisse aller Pflichtmodule des UNIGIS professional Zertifikatsstudiums für das Jahr 2017. n= 137

Im Gegensatz zur turnusmäßig durchgeführten Pflichtevaluation von Präsenzlehrveranstaltungen, die primär auf Vergleichbarkeit quantifizierbarer Indikatoren und weniger auf inhaltliches Feedback ausgerichtet ist, werden UNIGIS Studierende dazu angehalten jedes Lernmodul freiwillig und anonym, dafür jedoch möglichst umfassend zu evaluieren. Neben geschlossenen Fragen, die beispielsweise auf die angesprochene Praxisrelevanz der Inhalte abzielen (Abb.1), stellt vor allem die Auswertung der oftmals umfangreichen Antworten auf offene Fragen zu Inhalten und deren Vermittlung einen erheblichen Mehrwert für die kontinuierliche Weiterentwicklung der Studienmaterialien dar. Einzelne Studierende, die in der jeweiligen Thematik selbst über Expertenwissen verfügen, generieren dabei durchwegs interessante, inhaltliche Impulse für den Revisionsprozess. Der weitaus größere Anteil der Anregungen bezieht sich jedoch auf Didaktik und Aufbau oder reflektiert den persönlichen Lernfortschritt. Im Zuge der iterativen Überarbeitung der Studienmaterialien wird mit derartigen Hinweisen ein beachtlicher didaktischer „Feinschliff“ erreicht, der in dieser Qualität bei einer alleinigen Begutachtung durch FachkollegInnen wohl kaum möglich wäre. Da UNIGIS-Lehrende Auszüge der solcherart entwickelten Fernstudienmaterialien auch in den eigenen Präsenzlehrveranstaltungen am Fachbereich einsetzen, kommt das

hohe Qualitätsniveau von Skripten oder Übungsbeispielen wiederum den lokalen Studierenden zugute.

Fallbeispiel
Der seit 2008 seitens der Universität Salzburg vergebene Preis für hervorragende Lehre, kürt in einem kompetitiven, mehrstufigen Auswahlverfahren die jeweils besten vier der über 4000 Präsenzlehrveranstaltungen eines Jahres. Zweimal (2008: *Entwurf und Entwicklung von Geoinformatik-Anwendungen*, 2015: *Praxis der Geo-Kommunikation*) wurden dabei bereits Lehrveranstaltungen prämiert, die auf UNIGIS Studienmaterialien basieren und von UNIGIS-Lehrenden abgehalten wurden.

2.2 Beiträge zur Weiterführung universitärer Forschungsthemen

Aus der vorwiegend drittmittelfinanzierten Forschung am Fachbereich resultieren immer wieder wichtige Fragestellungen, die in eigenen Projekten nicht weiterverfolgt werden können. Im Blog *https://zgis-theses.blogspot.com* werden derartige Fragestellungen an Studierende kommuniziert bzw. als Abschlussarbeitsthemen angeboten. Bei UNIGIS Studierenden, die kein Thema aus ihrem eigenen Arbeitsbereich einbringen können oder wollen, erfreut sich diese Themenbörse großer Beliebtheit. Ein oft breiterer Horizont gepaart mit Erfahrung im eigenständigen (auch wissenschaftlichen) Arbeiten führt dabei durchwegs zu sehr guten Ergebnissen und eine über das Studium hinausgehende Zusammenarbeit stellt keine Seltenheit dar.

Fallbeispiel
Die Bearbeitung des optionalen UNIGIS-Moduls „Spatial Simulation“ weckt das Interesse eines Hamburger UNIGIS Studenten mit 20 Jahren Berufserfahrung im IT-Bereich (Projektleitung, Programmierung, IT-Systemintegration) tiefer in den Forschungsbereich der agentenbasierten Simulation einzusteigen. Die diesbezügliche Forschungsgruppe am Fachbereich wird von der Modulautorin bzw. UNIGIS-Lehrbeauftragten geleitet. Gemeinsam wird ein entsprechendes Thema im Bereich des Nationalpark-Monitoring entwickelt, das sich für eine Masterthesis eignet und sowohl den Interessen des Studenten als auch denen der Forschungsgruppe entgegenkommt. Während der Masterarbeit, die aufgrund der hervorragenden Basiskompetenzen (akademische Vorbildung, Arbeitshaltung, Selbstorganisation, Programmierung) im Jänner 2018 sehr erfolgreich abgeschlossen wird, entwickelt sich ein intensiver fachlicher Austausch. Der Student wird schließlich in die Forschungsgruppe aufgenommen und nimmt per Videokonferenz an deren Treffen teil. Gemeinsame Projekte werden lanciert und durch das Netzwerk des Studenten geeignete Projektpartner gefunden. Die im Rahmen der Masterarbeit erworbene Expertise geht in gemeinsame Publikationen ein und wird im Rahmen einer zehntägigen Winterschool *„Spatial Simulation Modelling“*, in welcher der Student eine aktive Rolle als Lehrender einnimmt, an die nächste Studierendengeneration weitergegeben.

2.3 Verbesserte Vernetzung mit Wirtschaft und Verwaltung

Universitäten erkennen immer mehr, dass die Beziehungspflege zu AbsolventInnen abgesehen von fallweiser Projektkooperation vor allem auch laufende Kontakte zu Wirtschaft und Verwaltung etabliert und so das Netz-Werken, Rückmeldungen aus der Praxis und die Sichtbarkeit des

Faches außerhalb der Wissenschaftscommunity fördert. Alumni-Veranstaltungen zielen darauf ab, dass erfolgreiche Studierende mit ihren akademischen Wurzeln verbunden bleiben. Im berufsbegleitenden Studienkontext sind derartige, oft beschränkt erfolgreiche Aktionen weniger erforderlich, da die Mehrzahl der Studierenden bereits während des Studiums in Wirtschaft und Verwaltung eingebettet ist und damit von Beginn an ein aktives Netzwerk bildet. UNIGIS verschränkt vom ersten Studientag an aktive Studierende mit Alumni und dem Fachbereich und etabliert so langfristig tragfähige Beziehungen, die in vielen Fällen auch bestehen bleiben.

Fallbeispiel
Ein engagierter, 27-jähriger Mitarbeiter eines Ingenieurbüros absolviert 2004 zur Ergänzung seiner Ausbildung als Vermesser das einjährige UNIGIS professional Zertifikatsstudium und wird in Folge im Unternehmen mit anspruchsvolleren Aufgaben im GIS Bereich betraut. In den folgenden Jahren macht er regen Gebrauch der zum fachlichen Austausch unter UNIGIS-Studierenden und Alumni zur Verfügung gestellten Email-Verteilerliste, die als Diskussionsplattform für konkrete GI-Fachfragen aus dem Arbeitsumfeld dient. Dabei tritt der nunmehrige Alumni immer öfter auch in der Rolle des Problemlösers auf. Um seine Karriere weiter voranzutreiben, entschließt er sich das UNIGIS MSc-Studium in Angriff zu nehmen, das er 2012 mit ausgezeichnetem Erfolg und einer vom österreichischen Dachverband für Geoinformation (AGEO) preisgekrönten Masterarbeit abschließt. Sein Erfolg im GI-Projektbereich hat nicht nur dazu beigetragen, dass 2016 ein Arbeitskollege das UNIGIS MSc-Studium ebenfalls erfolgreich abschließen konnte, sondern auch dazu, dass Netzwerkaktivitäten immer weiter ausgedehnt wurden: 2018 hat der ehemalige Mitarbeiter aus dem Vermessungsbereich einen GI-Fachvortrag auf dem vom Fachbereich ausgerichteten AGIT-Symposium eingereicht und sein Arbeitgeber ist auf der AGIT-Expo als Aussteller vertreten.

3 Fazit

Der gesellschaftliche Wandel bzw. die damit einhergehende Forderung zu lebenslangem Lernen bietet den Universitäten nicht nur die Chance die Weiterbildung zu betonen, sondern insbesondere auch Kontakt mit praxisrelevanten Zielgruppen zu halten. Als Akteure in einer Querschnittsmaterie verfügen viele GI-Fachvertreter über eine hohe Vernetzungskompetenz. Damit bisher ungenutzte Potentiale zu heben, sollte also nicht schwerfallen.

4 Literaturverzeichnis

BLASCHKE, T., STROBL, J., CAR, A. & TRAUN, C., 2011: Geoinformatik und GIScience in allen Stufen des Bologna-Systems – eine Bilanz an der Universität Salzburg. Publikationen der Deut. Gesellschaft für Photogrammetrie, Fernerkundung u. Geoinformation e.V., Band 20, Seyfert, E. (Hrsg.), Mainz, 9-12.

STROBL, J., 2011: UNIGIS – Networked Learning over a Distance. In: Unwin, D. J., Foote, K. E., Tate, N. J. & DiBiase, D. (Hrsg.): Teaching Geographic Information Science and Technology in Higher Education. John Wiley & Sons Ltd.

E-Learning und Learning Analytics in der universitären Ausbildung der Geowissenschaften

MIKE TEUCHER[1], CORNELIA GLÄßER[1] & VERA SCHREINER[1]

Zusammenfassung: Im Projekt „Prolearn" wird der Einsatz von E-Learning und E-Teaching Instrumenten in der geowissenschaftlichen Hochschullehre untersucht. Dabei stehen Lernmodule und Learning Analytics Methoden im Fokus, um den Lernerfolg/-fortschritt der Studierenden zu analysieren.
Für die Erstellung der Online-Inhalte, die ergänzend zur Veranstaltung angeboten werden, wird ein System aus 3 Komponenten, bestehend aus der Plattform ArcGIS Online, einem projekteigenem CMS und dem universitätseigenen Learning-Management-System ILIAS verwendet. Durch kurze Tests, kombiniert mit übergeordneten Erhebungen wird der Lernerfolg für die Dozierenden sicht- und durch die direkte Rückmeldung für die Teilnehmer quantifizierbar. Übergeordnete Zielstellung ist, ergänzend zu den fachlich-inhaltlichen Aspekten die Integration eines Empfehlungssystems zur individuellen Reduzierung von Wissensdefizite.

1 Einleitung

Das Angebot digitaler Lehre an deutschen Hochschulen wächst seit den 1990er Jahren stetig. Auch der Stellenwert digitaler Lehre an Hochschulen wird überdurchschnittlich eingestuft (WANNEMACHER 2016), weshalb E-Learning-Systeme zunehmend an Bedeutung für Bildungseinrichtungen und Unternehmen gewinnen (WANNEMACHER 2016; PRINZ et al. 2017). Die Fachgruppe Geofernerkundung und Kartographie des Instituts für Geowissenschaften und Geographie der Martin-Luther-Universität Halle-Wittenberg weist mit den Projekten WEBGEO und GEOVLEX im Bereich Blended Learning seit mehr als 10 Jahren Erfahrungen in diesem Bereich auf (THÜRKOW 2009, 2014).
Als konsequente Weiterentwicklung der bestehenden Konzepte, untersucht die Fachgruppe im Projekt „Prolearn" den Einsatz von prototypischen E-Learning und E-Teaching Instrumenten in der geowissenschaftlichen Hochschullehre. Dabei stehen Lernmodule und Learning Analytics Methoden im Fokus, welche detaillierte Informationen über das Lernerlebnis von Studierenden, u.a. beim Übergang vom Bachelor- zum Masterstudium liefern. Ziel ist es das studentische Vorwissen zu analysieren, Wissenslücken zu identifizieren und neue Lernwege aufzuzeigen.

2 Problemstellung

Das Projekt „Prolearn" ist in seiner zweijährigen Projektzeit in mehrere Phasen gegliedert. In einer ersten Phase wird aktuell für eine ausgewählte Veranstaltung im 1. Fachsemester des Masterstudiums, die durch eine sehr heterogene Zusammensetzung hinsichtlich Studiengänge

[1] Martin-Luther-Universität Halle-Wittenberg, Institut für Geowissenschaften und Geographie, Fachgruppe Geofernerkundung & Thematische Kartographie, Von-Seckendorff-Platz 4, D-06120 Halle/Saale, E-Mail: [mike.teucher, cornelia.gläßer, vera.schreiner]@geo.uni-halle.de

und Vorwissen charakterisiert ist, die Präsenzlehre durch interaktive E-Learning Module ergänzt, bei denen die Studierenden das theoretische Wissen der Vorlesung in modularen Online Lerneinheiten vertiefen können. Voraussetzung für die Erstellung dieser ist die Erfassung des Wissenstandes der Teilnehmer zu Beginn des Semesters. Hierfür wurden in 2 Vorab-Befragungen allgemeine Angaben über die Teilnehmer (u.a. bisherige Studienrichtung, besuchte themenverwandte Veranstaltungen, EDV-technische Ausstattung und Zugangsmöglichkeiten) erfasst und eine Überprüfung der fachlichen Kenntnisse in den Bereichen Geoinformationssysteme, Fernerkundung und Kartographie durchgeführt. Die Teilnehmer wurden zunächst aufgefordert ihren eigenen Wissensstand in Bezug auf die Bereiche GIS, Fernerkundung und Kartographie einzuschätzen. Diese Einschätzung wurde im Verlauf des Fragebogens durch jeweils fünf Fragen aus den Fachgebieten überprüft.
Für die Erstellung der jeweiligen Online Lerneinheiten, die parallel zur Vorlesung angeboten werden, wird ein System aus 3 Komponenten eingesetzt. Für die thematische Bearbeitung der Inhalte wird zum Großteil die Plattform ArcGIS Online der Firma ESRI, als Teil des Campus Systems der Universität genutzt (ESRI 2018). Für die Präsentation der Inhalte ist ein projekteigenes Content Management System (CMS) im Institut erstellt wurden, auf dem alle Inhalte zentral vernetzt sind. Um im Laufe des Projektes den individuellen Lernerfolg/-fortschritt der Teilnehmer entsprechend des ECLASS Ansatzes (GERSON 2000) erfassen zu können, wird jede Lerneinheit durch quiz-artige Tests abgeschlossen. Dieses Testsystem wird im universitätseigenen Learning-Management-System ILIAS betrieben. Durch kurze Tests zum Ende einer jeden Lerneinheit, kombiniert mit zweimaligen übergeordneten Erhebungen im Laufe des Semesters wird der Lernerfolg für die Dozierenden sicht- und durch die direkte Rückmeldung der Lernerfolg für die Teilnehmer quantifizierbar. Die übergeordneten Erhebungen beinhalten Fragen sowohl aus der Präsenzveranstaltung, wie auch aus den Themenbereichen der Online Lerneinheiten. Konkret wurden zum Zeitpunkt der ersten übergeordneten Erhebung, etwa zur Mitte des Semesters, 25 Fragen unterschiedlichen Niveaus aus den drei Bereichen gestellt. In allen Bereichen fanden sich Fragen, die bereits in identischer, oder ähnlicher Form in der Vorab-Befragung beinhaltet waren, um den Lernerfolg der Teilnehmer analysieren zu können.
Weiterhin wird durch diesen Blended-Learning-Ansatz eine individuellere Betreuung der Studierenden durch selbständiges Lernen anhand der entwickelten E-Learning-Module ermöglicht. Der etablierte Ansatz und seine einzelnen Komponenten sind in Abbildung 1 veranschaulicht.

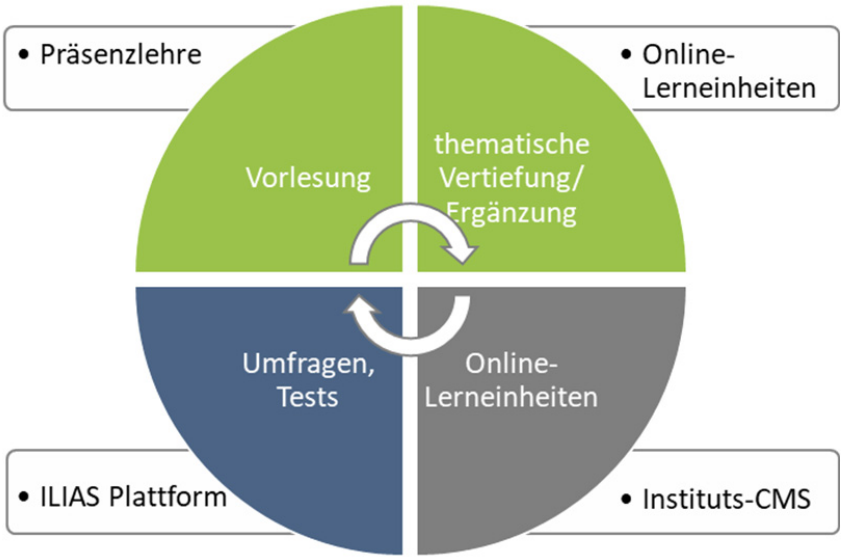


Abb. 1: Konzept Double-Blended Learning

3 Fazit & Ausblick

Die Ergebnisse der Vorab-Befragung liefern wichtige Informationen über den aktuellen Wissenstand der Teilnehmer und die Bereitschaft zur Nutzung der Online-Lerneinheiten. Abbildung 2 zeigt als Ergebnis der Vorab-Befragung die Einstellung der Teilnehmer gegenüber Online Lernumgebungen. 74% der Befragten sind an einer Evaluation des eigenen Lernerfolgs/-fortschritt während der Veranstaltung interessiert.

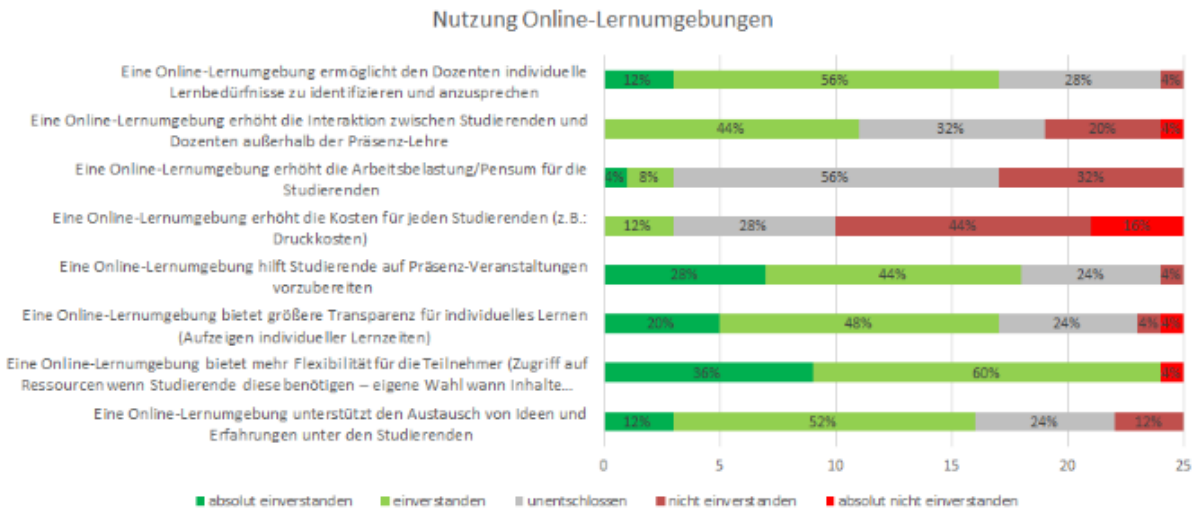


Abb. 2: Erwartungen und Einstellung der Befragten gegenüber Online Lernumgebung

69,23% der Befragten schätzten Ihren Wissensstand zum Thema GIS als „Anfänger(-in) mit geringen Vorkenntnissen“ ein, 26,92% wählten die Kategorie „Fortgeschrittene(r) mit sicherer Beherrschung der Grundfunktionen“ und lediglich 3,85% sind laut eigener Einschätzung „Anfänger(-in) ohne Vorkenntnisse“. Die Überprüfung dieser Einschätzung mittels 5 Detailfragen ergab, dass im Bereich GIS Befragte der Kategorie „Anfänger(-in) mit geringen Vorkenntnissen“ durchschnittlich 55,56% und 51,19% der erreichbaren Punkte durch die Teilnehmer der Kategorie „Fortgeschrittene(r) mit sicherer Beherrschung der Grundfunktionen“ erzielten. Die Aufgabenschwierigkeit nach BORTZ & DÖRING (2006) aller Detailfragen deckte den Bereich von 0 bis maximal 0,88 ab, wobei 53% der Fragen auf den Bereich von 0,2 bis 0,8 entfallen, 40% auf den Bereich von 0,8 bis 1,0 und lediglich 7% den unteren Bereich betreffen. Die gemittelte Schwierigkeit über alle Fragen beträgt 0,5. Für den Bereich GIS beträgt die durchschnittliche Schwierigkeit der Fragen 0,57, mit einem Minimalwert von 0,36 und einem Maximalwert von 0,88. Für die Bereiche Fernerkundung und Kartographie liegen die durchschnittlichen Werte bei 0,31 (0,00-0,88), bzw. 0,57 (0,44-0,88). Tabelle 1 gibt eine Übersicht der gewählten Kategorien für die Bereiche Fernerkundung und Kartographie.

Tab. 1: Einschätzung des Wissensstandes in den Bereichen Fernerkundung und Kartographie

Kategorie	Fernerkundung		Kartographie	
	Selbst-einschätzung	∅ Erreichte Prozent in Detailfragen	Selbst-einschätzung	∅ Erreichte Prozent in Detailfragen
Anfänger(-in) ohne Vorkenntnisse	80,00%	32,00%	-	-
Anfänger(-in) mit geringen Vorkenntnissen	20,00%	28,00%	56,00%	59,92%
Fortgeschrittene(r) mit sicherer Beherrschung der Grundfunktionen	-	-	44,00%	52,53%
Experte/Expertin: Sie beherrschen auch anspruchsvollere Funktionen	-	-	-	-

Die Ergebnisse der ersten übergeordneten Erhebung, die etwa zur Mitte des Wintersemesters 2017/18 durchgeführt wurde, weisen eine gemittelte Schwierigkeit aller Fragen von 0,52 auf, wobei 84% der Fragen im Bereich zwischen 0,2 und 0,8 liegen und jeweils 8% auf den unteren, bzw. oberen Bereich entfallen. Die Erhebung, bestehend aus 25 Fragen, gliedert sich in je 7 Fragen aus dem Bereich GIS, bzw. Fernerkundung und 11 Fragen aus dem Bereich Kartographie. Es zeigt sich, dass es eine Verbesserung im Bereich Fernerkundung gegeben hat, da die Aufgabenschwierigkeit im Vergleich zur Vorab-Befragung von 0,31 auf den 0,52 gestiegen ist. Dies bedeutet, dass gegenüber der Vorab-Befragung die Teilnehmer der ersten übergeordneten Erhebung nun 52,00% der Fragen richtig gelöst haben. Für die Bereiche GIS und Kartographie liegen die Werte bei 0,45 (-0,12), bzw. 0,58 (+0,01). Eine genauere Analyse der Ergebnisse aus dem Bereich GIS ergab, dass vor allem Fragen zum Basiswissen Geographischer Informationssysteme Probleme bereiteten, wohingegen Detailfragen zu den behandelten Inhalten der Präsenz- und Online-Lerneinheiten überwiegend gute Ergebnisse erzielten. Dies ist eine

mögliche Erklärung für den Rückgang in diesem Bereich im Vergleich zur Vorab-Befragung. Ähnliche Tendenzen sind für den Bereich Kartographie zu erkennen, da auch hier die Detailfragen bessere Ergebnisse lieferten als Fragen zum Basiswissen Kartographie. Im Gegensatz hierzu wurde deutlich, dass im Bereich Fernerkundung primär die Fragen zum Basiswissen im Zwischentest bessere Ergebnisse erzielten, als Detailfragen, die die Inhalte der Veranstaltung thematisierten.
Zusätzlich zur Vorab-Befragung und einer ersten übergeordneten Erhebung wurden alle Online-Lerneinheiten mit kleineren quiz-artigen Test beendet. Abbildung 3 zeigt die Beteiligung (orange Linie) und die jeweils erreichten Prozent der Teilnehmenden (blaue Säulen).
Es wird ersichtlich, dass über alle Online-Lerneinheiten ca. 56% der Punkte erreicht wurden. Die Maximalwerte von 86%, bzw. 83% wurden bei den Online-Lerneinheiten zu den Themen „Landsat“ und „Räumliche Analyse und Interpolation“ erzielt, wohingegen die niedrigsten Ergebnisse im Themenkomplex „Veränderungsanalysen mit Fernerkundungsdaten“ erzielt wurden. Auch hier zeigt sich, dass im Bereich Fernerkundung grundlegendes Wissen bessere Ergebnisse erzielten als spezifische Anwendungen. Eine weitere Erklärung für die niedrigen Werte der Online-Lerneinheit 6 liegt wohlmöglich in der (technischen) Komplexität der verwendeten Anwendung selbst, da in dieser Einheit die Teilnehmenden das Thema mit Hilfe einer extern vorkonfigurierten Webanwendung erschließen sollten.

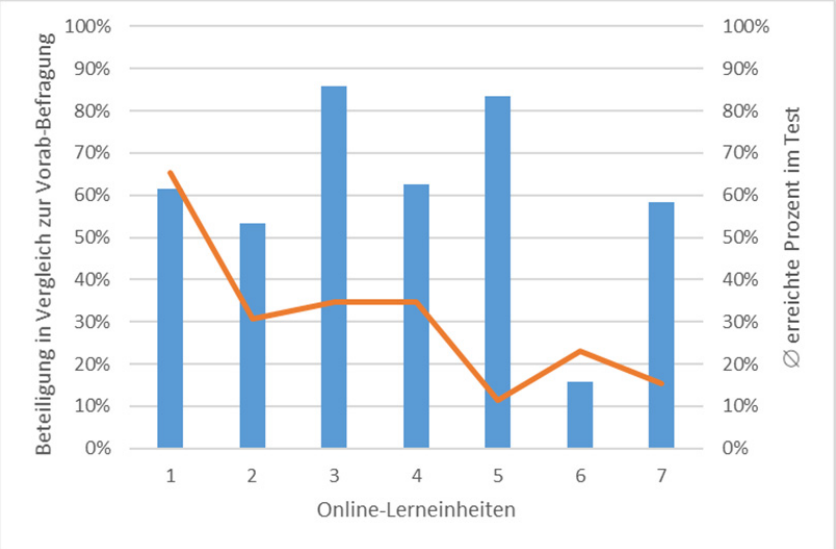


Abb. 3: Ergebnisse der Online-Lerneinheiten

Zusammenfassend lässt sich festhalten, dass anhand der oben dargestellten Ergebnisse das System eine individuelle Identifizierung vorhandener Wissensdefizite und eine Differenzierung des Wissensstands der Teilnehmenden in die Kategorien „Basiswissen“ und „erweiterte Kenntnisse“ ermöglicht. Die Ergebnisse zeigen aber auch, dass eine solide Bewertung des Lernerfolgs nur bei einer ausreichend hohen und konstanten Beteiligung möglich ist. Diesem Problem konnte im ersten stattgefundenen Durchgang noch nicht entgegnet werden, stellt aber einen zentralen Fokus für die weiteren Projektschritte dar. Eine abschließende Evaluation des

Gesamtsystems hinsichtlich Nutzerfreundlichkeit und Funktionalität steht zum momentanen Zeitpunkt noch aus und wird zum Ende der Veranstaltung durchgeführt werden.
Eine weitere übergeordnete Zielstellung des Projektes „Prolearn“ ist ergänzend zu den fachlich-inhaltlichen Aspekten, die eingangs aufgezeigt werden, die technische Integration einer umfangreichen Funktion zur Visualisierung des individuellen Lernerfolg/-fortschritt im System ILIAS selbst. Die Möglichkeiten dieser Integration werden in den kommenden Projektschritten analysiert und bewertet. Vorstellbar zum aktuellen Zeitpunkt ist eine „Dashboard“-ähnliche Darstellung des individuellen Fortschrittes. Als zentrale Funktionalität dieses System ist die Erstellung eines semi-/automatisierten Empfehlungssystems von Lerneinheiten oder Literatur zur individuellen Reduzierung der Wissensdefizite der Teilnehmenden.
Weitere Ergebnisse sind zum Ende des Semesters im Februar zu erwarten. Diese werden dann im Rahmen der PFGK18: Photogrammetrie - Fernerkundung - Geoinformatik - Kartographie – 2018 Tagung präsentiert. Zugleich soll mit diesem Beitrag eine Diskussion über die Bündelung bestehender Online Angebote von Mitgliedern der DGPF angeregt werden.

4 Literaturverzeichnis

BORTZ, J. & DÖRING, N. 2013: Forschungsmethoden und Evaluation. Berlin, Heidelberg, Springer-Verlag (Hrsg.).

ESRI, 2018: ArcGIS Online. Unter: http://www.arcgis.com/features/index.html, letzter Zugriff 22. Januar 2018.

GERSON, S. M., 2000: E-CLASS: Creating a Guide to Online Course Development for Distance Learning Faculty. Online Journal of Distance Learning Administration, **3**(4).

PRINZ, C.; KREIMEIER, D. & KUHLENKÖTTER, B., 2017: Implementation of a Learning Environment for an Industrie 4.0 Assistance System to Improve the Overall Equipment Effectiveness. Procedia Manufacturing, **9**, 159-166.

THÜRKOW, D.; GLÄßER, C.; SCHEUER, S. & SCHIELE, S., 2009: Visualisation of Hydrological Processes with GEOVLEX: Introduction of an integrated methodical-technical Online Learning Approach. Tools and Techniques for E-Learning. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(6/W7), König, G. &Lehmann, H. (Hrsg.), Proceedings of the ISPRS working group VI/1 – VI/2, June 17-19, 2009 Potsdam, Berlin.

THÜRKOW, D.; 2014: H2o EXTREM - interaktives Lernen. Moderne Lehre gestalten, anwenden - motivieren - vernetzen. Tagungsband zur Jahrestagung für multimediales Lehren und Lernen. Zentrum für multimediales Lehren und Lernen der Martin-Luther-Universität Halle-Wittenberg, 07.11.2014, Halle (Saale).

WANNEMACHER, K., 2016: Organisation Digitaler Lehre in den Deutschen Hochschulen. Arbeitspapier Nr. 21, Berlin, Hochschulforum Digitalisierung (Hrsg.).

Immersive 3D-Visualisierung von Bauwerken mit Virtual Reality in der Hochschulausbildung

FELIX TSCHIRSCHWITZ[1], MAREN LINDSTAEDT[1] &THOMAS P. KERSTEN[1]

Zusammenfassung: Virtual Reality (VR) hat sich in den letzten vier Jahren als interessantes Thema in vielen Fachdisziplinen etabliert. Auch wenn VR als Teilgebiet der Informatik eine lange Tradition hat, so sind die Erkenntnisse und Anwendbarkeit dieser Technologie erst durch die Entwicklung der erforderlichen Hardware im Low-Cost-Bereich stark verbreitet worden. Sie hat heute den Endkundenmarkt erreicht und ist auch für die Geodäsie interessant geworden.
Am Beispiel der Veranstaltung 3D-Visualisierung im Masterstudienprogramm Geodäsie und Geoinformatik der HafenCity Universität Hamburg wird aufgezeigt, wie selbstorganisierte Gruppenarbeit zur praktischen Kompetenzfindung eingesetzt werden kann. Der Ausgangspunkt der Lehrveranstaltung waren Grundlagenvorlesungen zum Thema VR, eigene praktische Vorarbeiten sowie eine fordernde Aufgabenstellung, die aus der Praxis inspiriert ist. Am Beispiel erhaltener und rekonstruierter historischer Bauwerke wurde von den Studierenden in einer Game Engine eine virtuelle Umgebung passend zu dem zu visualisierenden Objekt erstellt, in dem das als CAD-Modell zur Verfügung gestellte Gebäude aufbereitet, texturiert und in VR mit Interaktionsmöglichkeiten begehbar gemacht wurde.

1 Einleitung

Virtual Reality (VR) beschreibt eine „hochwertige Schnittstelle zwischen Nutzer und Computer, die Echtzeitsimulation und Interaktionen über mehrere sensorische Kanäle erlaubt“ (BURDEA & COIFFET 2003). VR nutzt dabei das menschliche Verlangen aus, dem Wahrgenommenen zu trauen und täuscht so dem Nutzer eine Erfahrung vor, die real nicht stattfindet. Indem der Nutzer vor seinen eigenen Augen Bewegungen sieht und akustische Informationen erhält, durch Körper- und Kopfbewegungen den audiovisuellen Blickwinkel verändern und durch seine eigenen Hände mit der virtuellen Welt interagieren kann, wird die ihm vorgeführte Virtuelle Realität real wahrgenommen. Dieses Konzept und sogar die aktuelle technische Umsetzung sind dabei nicht neu. Bereits in den 1960er Jahren wurde Hardware entwickelt und prototypisch hergestellt, die mit Hilfe zweier vor den Augen des Nutzers angeordneten Anzeigen virtuelle Inhalte visualisiert haben (RHEINGOLD 1991; SUTHERLAND 1965). Die erzielbare visuelle Qualität bzw. zeitliche Synchronität zwischen den Sensoren und Displays ist allerdings erst seit wenigen Jahren so günstig zu produzieren, dass eine kommerzielle Nutzung der Technologie auch im Endnutzerbereich möglich ist. Der Umstand, dass die virtuelle Umgebung angenommen und das Reale dabei teilweise ausgeblendet wird, kann vielfältig eingesetzt werden. So haben beispielsweise FURMAN et al. (2009) VR als Ablenkung eingesetzt und währenddessen zahnmedizinische Eingriffe durchgeführt. ANDERSON et al. (2013) ersetzen durch VR den realen Stimulus bei einer Expositionsstudie und erzielen vergleichbare Resultate zu den realen Stimulus. So ist es nicht weiter

[1] HafenCity Universität Hamburg, Labor für Photogrammetrie & Laserscanning, Überseeallee 16, D-20457 Hamburg, E-Mail: [Felix.Tschirschwitz, Maren.Lindstaedt, Thomas.Kersten]@hcu-hamburg.de

verwunderlich, dass sich architektonische Bauwerke oder historische Stätten geradezu aufzwingen in VR begehbar gemacht zu werden, um ähnliche Eindrücke wie der reale Besucher vor Ort zu erhalten. Gerade Orte und Bauwerke, die weit entfernt, schwer erreichbar oder gar nicht mehr vorhanden sind bzw. in der Vergangenheit liegen, können durch VR einer breiten Öffentlichkeit zugänglich gemacht werden (GAITATZES et al. 2001). Beispiele für die Begehung historischer Gebäude mit weiteren semantischen Informationen liefert etwa das Virtuelle Museum Alt-Segeberger Bürgerhaus (DEGGIM, 2017a) oder das Projekt Segeberg 1600 (DEGGIM, 2017b), bei dem eine nicht mehr existierende historische Stadt und Burg wieder erschaffen und durch VR erlebbar gemacht wurden. Forschungsarbeiten z.B. von STEINICKE (2010) oder INTERRANTE (2006) zeigen allerdings, dass die aktuelle virtuelle Darstellung noch vom Ideal der Deckungsgleichheit der echten Erfahrung abweicht und so immer nur einen Eindruck vermittelt, nicht aber die reale Erfahrung, die ohnehin durch viele Sinne erlebt wird, ersetzen kann.
In diesem Beitrag wird die technische Umsetzung beschrieben, die im Rahmen der Ausbildung im Masterstudienprogramm Geodäsie und Geoinformatik an der HafenCity Universität Hamburg realisiert wurde, um den Studierenden nicht nur theoretische Grundlagen zu vermitteln, sondern um ihnen sowohl die VR-Technologie zu erläutern, als auch praktische Erfahrungen bei der Implementierung von echten Geodatensätzen in eine Game Engine zu ermöglichen.

2 Methodik und Didaktik der Lehrveranstaltung

Das Modul Visualisierung im Masterstudienprogramm Geodäsie und Geoinformatik (bis Sommersemester 2017 Geomatik) findet im dritten Semester statt und hat eine definierte Arbeitsbelastung von 10 CP bzw. 300 Stunden pro Student. Das Modul teilt sich in die zwei Lehrveranstaltungen Geovisualisierung (25%) und 3D-Visualisierung (75%) und deckt unterschiedliche Aspekte der Visualisierung von Geodaten ab. Während in der Geovisualisierung Geodaten aller Skalen und Themen behandelt werden, wird in der 3D-Visualisierung explizit auf Visualisierung von geometrischen dreidimensionalen Inhalten, sprich 3D-Modellen, eingegangen und setzt somit auf die Erfassung und Modellierung aus vorangegangenen Lehrveranstaltungen auf (siehe Kapitel 3).
Seit der Kickstarter Kampagne von Oculus Rift in 2012, spätestens durch die Übernahme der Firma Oculus durch Facebook im Jahr 2014, ist VR der (erwartete) Hype der letzten Jahre. Nicht zuletzt durch massiven Verkauf von sog. (Hardware) Development Kits (Oculus Rift DK1, Oculus Rift DK2) wurde der Markt für die Möglichkeiten der Technologie vorbereitet und Softwareentwicklungen wurden ermöglicht. So war zum Verkaufsstart (HTC Vive Anfang April 2016 / Oculus Rift Ende März 2016) großes Interesse und Akzeptanz vorhanden und Softwareanwendungen standen bereits zur Verfügung. Die ehemals Experten vorbehaltene Technologie VR ist nun dem breiten Anwenderkreis (Massenmarkt) zugänglich. Sie eignet sich ideal für die Visualisierung von 3D-Modellen, da ihre native Darstellungsform der menschlichen bestmöglich entspricht. Durch ihr breites Wissen über raumbezogene Daten sind Geodäten eine geeignete Anwendungsgruppe dieser sich rasant entwickelnden Technologie.

2.1 Lehrveranstaltung als Projekt

Betrachtet man das Lernen und Lehren in einer Hochschule oder Universität, so lassen sich mehrere Akteure und Sichtweisen identifizieren. Die zwei prominentesten sind dabei die direkten Akteure der Studierenden und Lehrenden. Dem Humboldt'schen Bildungsideal folgend sollte diese die Dualität von Lehre und Forschung wahren und nur der Wissensgenerierung verpflichtet sein. Heutzutage ist allerdings mit dem Arbeitsmarkt, für den ausgebildet wird, ein dritter Akteur sehr sichtbar geworden, der nicht nur positiv bewertet wird (FOHRMANN 2016). Die Lehrenden analysieren und bewerten den Markt und abstrahieren Kompetenzen, die gefordert sind, um ihre Absolventen für das angestrebte Berufsfeld zu qualifizieren. Sie müssen dabei die Qualifikationen der Studierenden über geeignete Methoden evaluieren und vergleichbar machen durch eine qualitative Bewertung. Im Kontext der Wirtschaftlichkeit soll diese Bewertung möglichst einfach aber aussagekräftig erfolgen. Der Studierende hat eine ähnliche, aber entgegengesetzte Sichtweise. Er möchte sich ebenfalls für den Markt qualifizieren und dabei Kompetenzen ökonomisch erwerben, allerdings diese ebenfalls möglichst positiv nachweisen können.

Die Ökonomisierung des Lernens, speziell in einem sehr neuen, wenig disziplinär beschränkten Arbeitsfeld erfordert auch ein Umdenken in bestehenden Lehr- und Lernmustern und führt zu einer Beschleunigung der Informationssuche und -verarbeitung. Gerade im Kontext der omnipräsenten Sozialen Medien, der medialen Kommunikation abseits klassischer Sender-Rezipienten-Modelle (Web 2.0), werden immer leichter und selbstverständlicher Blogs, YouTube und andere Wissensaustauchformen abseits von klassischen Studienquellen wie Lehrbüchern und Fachartikeln verwendet. So ist der Rückkanal ein wesentliches Kriterium von Web 2.0 und wird von RIEMANN et al. (2007) als wichtiger Baustein in neuen Lernmustern aufgeführt, in der nicht nur Wissen abgefragt, sondern auch reproduziert dargestellt werden soll. Leider fordert dieses einen großen Mehraufwand für die Studierenden und Lehrenden, der meistens gescheut wird.

Die durchgeführte Lehrveranstaltung besteht aus einführender Vorlesung in die Thematik sowie in einer Projektbearbeitung, die sich anhand realer, der Praxis entsprechender Anforderungen orientiert. Bearbeitet werden dabei sechs individuelle Projekte in Gruppen von vier bis fünf Studierenden. Sind die grundlegenden Anforderungen durch die Visualisierung eines historischen Gebäudes klar umrissen, ist die explizierte Ausgestaltung den einzelnen Gruppen überlassen. Sie können dabei selbst den Schwierigkeitsgrad mitbestimmen, das Projekt fachlich definieren und durchführen. Neben der fachlichen Bearbeitung werden auch die immer wichtiger werdenden Aspekte des Gruppen- und Projektmanagements erprobt.

2.2 Evaluation von Gruppenarbeit bei selbst gestellter Aufgabe

Die Bearbeitung von sechs individuellen Projekten mit ähnlichen, aber doch projektspezifischen Anforderungen und Herausforderungen führt die Studierenden zu einer Gruppen-Leistung, die sich im Vergleich zu den anderen Gruppen relativ gut beurteilen lässt. Die selbstgestellte Herausforderung gibt dabei den Schwierigkeitsgrad vor, der durch die Ausführung multipliziert wird. Der Nachweis der Leistung erfolgt dabei über das erstellte Projekt sowie über einen individuellen Projektbericht je Studierenden, der die Einzelleistung von der Gruppenleistung entkoppelt und daher zu einer individuellen Gesamtbenotung führt.

Die Anfertigung eines qualifizierten Projektberichtes ist zwar eine notwendige Qualifikation für jeden Ingenieur, sie ist allerdings sowohl für den Studierenden als auch für den Lehrenden sehr

arbeitsintensiv. Ein weiteres, in zukünftigen Veranstaltungen mögliches Bewertungsformat ist dabei das Projektkolloquium, wo jeder Studierende seine Projektbearbeitung frei darstellen kann und durch gezielte Fragen in Form einer mündlichen Prüfung die erlangten Qualifikationen (Fähigkeiten und Wissen) evaluiert werden können. Dies geschieht in Abgrenzung der Bewertungsmethode Präsentation, da über die Aufbereitung des Wissens hinaus auch die praktische Umsetzung sichtbar dargestellt wird, d.h. neben dem Wissen werden auch die handwerklichen Qualifikationen erkennbar.

3 Verwendete 3D-Modelle

Als Bestandteil der Hochschulausbildung haben Studierende im Rahmen von Lehrveranstaltungen, Diplom-, Bachelor und Masterarbeiten sowie von Projekten im Laufe der letzten 17 Jahre viele 3D-Modelle historischer Gebäude erstellt. Die Datenerfassung der Gebäude erfolgte in der Lehrveranstaltung (Wahlpflichtmodul) Architekturphotogrammetrie im 6. Semester des Diplom- oder Bachelorstudienganges Geomatik durch digitale Photogrammetrie oder durch terrestrisches Laserscanning (TLS) im gleichnamigen Modul des Masterstudienganges Geomatik. Nach der Aufnahme der Gebäude wurden die Bilder durch Bildtriangulation mit entsprechender Kamerakalibrierung orientiert und die Laserscans zueinander registriert. Die orientierten Bilddaten und die TLS-Punktwolke bildeten die Grundlage, um die Gebäude anhand von manuellen Punktmessungen in AutoCAD zu konstruieren (KERSTEN et al. 2004, KERSTEN et al. 2005, KERSTEN et al. 2015). Viele der erstellten Modelle wurden später texturiert und anhand von Videosequenzen visualisiert, die in der Lehrveranstaltung “Projekt Visualisierung“ im Masterstudiengang Geomatik (altes Studienformat) durch verschiedene Softwarepakete wie Cinema 4D, Autodesk Maya und 3DS Max, Blender und Lumion 3D erzeugt wurden. So konnte den Studierenden neben der Wiederverwertung der Daten die gesamte Prozesskette durch praktische Arbeiten von der Erfassung, über die Konstruktion der Gebäude durch 3D-CAD-Modellierung bis hin zur Texturierung und Visualisierung gezeigt werden. Eine große Auswahl dieser durch Studierende erstellten Videos ist in dem YouTube-Kanal HCUHamburgGeomatics[2] anzuschauen. Für die Erstellung einer jeweiligen VR-Applikation mit einer Game Engine wurden in diesem Lehrprojekt folgende 3D-Gebäudemodelle verwendet, die auch Innenbereiche umfassten (Abb. 1/Abb. 2): Landdrostei Pinneberg, Westerturm Duderstadt, ausgewählte Gebäude der Sternwarte Hamburg-Bergedorf, Salomonischer Tempel sowie Almaqah-Tempel und Grat Be'al Gebri (beide Yeha, Äthiopien).

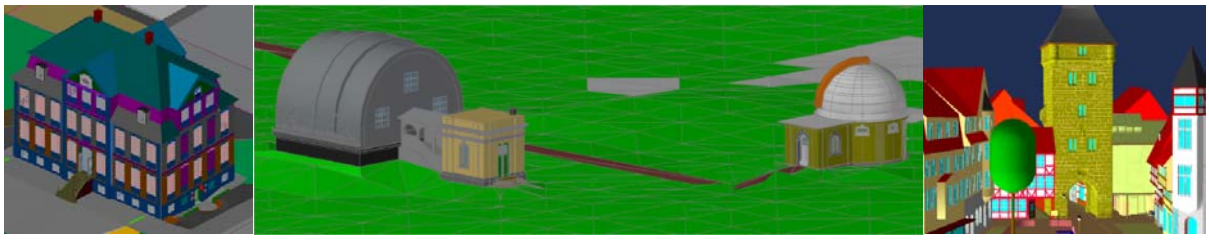


Abb. 1: 3D-CAD-Modelle von der Landdrostei Pinneberg (links), vom Äquatorial- und Meridianhaus der Sternwarte Bergedorf (Mitte) und vom Westerturm Duderstadt

[2] www.youtube.com/user/HCUHamburgGeomatics

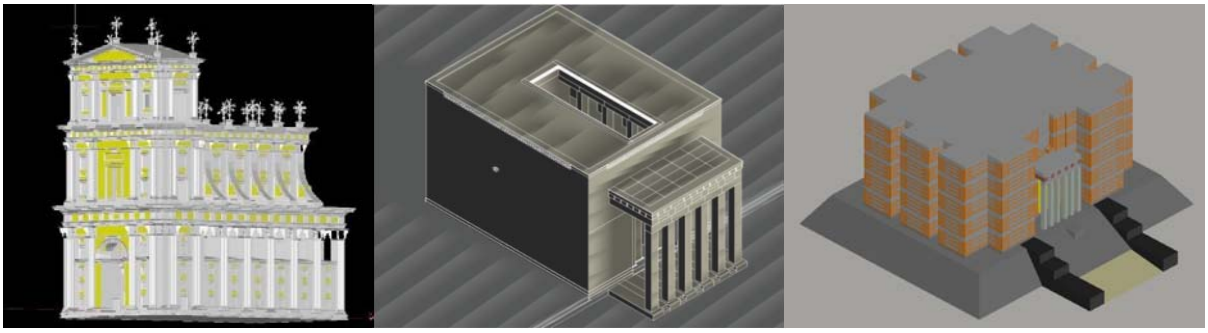


Abb. 2: 3D-CAD-Modelle vom Salomonischen Tempel (links), vom Almaqah-Tempel (Mitte) und vom Grat Be'al Gebri in Yeha (Äthiopien)

Studierende haben die Landdrostei Pinneberg, den Westerturm in Duderstadt und die Gebäude in der Sternwarte Hamburg-Bergedorf im Rahmen der oben erwähnten Lehrveranstaltungen generiert (Abb. 1). Das 3D-Modell vom Hauptgebäude des Salomonischen Tempels, welcher von 1680 bis 1692 in Hamburg angefertigt wurde und heute im Museum steht, wurde in einem Projekt mit dem Museum für Hamburgische Geschichte erstellt. Das Holzmodell des Gebäudes wurde von außen und nach Zerlegung des Gebäudes in deren Einzelteile auch von innen mit einer digitalen Spiegelreflexkamera Nikon D800 aufgenommen. Aus den Bildsequenzen wurden skalierte Punktwolken generiert, mit denen anschließend die detaillierte 3D-Konstruktion des Tempelgebäudes in AutoCAD erfolgte (KERSTEN et al. 2018). Von dem Almaqah-Tempel (oder Großer Tempel) und dem Monumentalbau Grat Be'al Gebri, beide in Yeha im Norden von Äthiopien gelegen, stehen heute nur noch Ruinen in der Ortschaft. Aus Punktwolken vom terrestrischen Laserscanning wurden vermaschte 3D-Modelle der beiden Ruinen berechnet, die als Grundlage für eine virtuelle Rekonstruktion der beiden Bauten durch einen Architekten dienten (LINDSTAEDT et al. 2011). Das Projekt in Äthiopien wurde in Zusammenarbeit mit der Außenstelle Sana'a der Orient-Abteilung des Deutschen Archäologischen Instituts bearbeitet.

4 Workflow der Implementierung

Die in vorherigen Lehrveranstaltungen, in Projekten oder in Abschlussarbeiten generierten Modelle bedürfen in den meisten Fällen einer mehr oder weniger aufwändigen Überarbeitung, um diese für VR einsetzbar zu machen. Die Modellierung, die zu großen Teilen aus Datenerfassung durch Photogrammetrie und TLS stammt, ist meist auf Volumenkörpern aufgebaut. Diese werden bei der Übernahme nach 3ds Max bzw. in die Engine in Dreiecke und Quads unterteilt, was die Datenmenge erstmal vergrößert. Grundsätzlich erfordern mehr Dreiecke bzw. Quads mehr Rechenleistung, was gerade bei VR leistungsstarke PCs mit guter Grafikkarte verlangt. So besteht generell das Bestreben, die Datenmenge einerseits gering zu halten und andererseits eine optisch ansprechende Darstellung des Objektes zu erzielen. Daher kann es zur Datenreduktion sinnvoll sein, geometrische Details durch entsprechende Texturen zu ersetzen, um gewünschte Details trotzdem darzustellen. Des Weiteren sind viele weitere Korrekturen in den Modellen zu berücksichtigen, wie z.B. verdrehte Flächennormalen, doppelte Wände oder Modellierungsungenauigkeiten, die in einem CAD-Programm gut zu beheben sind.

Ist das Modell geometrisch überarbeitet bzw. angepasst (korrigiert), besteht der nächste Schritt in der Texturierung der Objektelemente. Diese kann grundsätzlich vor der Datenübernahme in die Game Engine passieren, z.B. in 3ds Max, oder in der Engine selbst. Die Verwendung von Programmen wie 3ds Max, die viel in der Unterhaltungsbranche eingesetzt werden, bietet den Vorteileiner hohen Funktionalität für die Texturierung, was im Gegensatz aber eine anspruchsvolle Handhabung bewirkt und eine gewisse Einarbeitungszeit in diese Software unerlässlich macht. Als Ergebnis sind jedoch Texturen von hoher Qualität zu erwarten.
Die Erstellung der VR erfolgt mit einer Game Engine. Game Engines oder auch Spiele-Engines sind eine Sammlung von Funktionen, die zur Erstellung von interaktiven Spielewelten notwendig sind und zu diesem Zweck gesammelt zur Verfügung gestellt werden. Die Funktionen umfassen die visuelle Darstellung von texturierten 3D-Modellen, das Abspielen von Sound, Möglichkeiten zur Erstellung von logischen Abläufen durch Skripte, Netzwerkfunktionen und Physik bzw. Kollisionsberechnungen. O'FLANAGAN (2014) und LAWSON (2016) geben einen Überblick über beliebte Game Engines. Für diese Lehrveranstaltung wurde sich für die Unreal Engine 4 von Epic Games entschieden. Sie bietet neben sehr ansehnlicher grafischer Qualität, eine gute Implementierung von VR sowie mit Hilfe einer visuellen Programmiersprache (Blueprint) die Möglichkeit, Reaktionsverhalten der Umgebung und Interaktionen zu programmieren, ohne Quelltext verfassen zu müssen. Gerade Blueprints ermöglichen allen Studierenden trotz unterschiedlichem Fachwissen die Generierung einer Virtuellen Realität. Weiterhin ist die Engine für nicht kommerzielle Anwendungen kostenlos, sodass sie sich ohne weiteres in der Lehre einsetzen lässt.
Als Virtual Reality System kommt die Vive von HTC zum Einsatz (Abb. 3 links). Die HTC Vive wurde im April 2016 veröffentlicht und bietet eine raumfüllende VR-Erfahrung durch ein von außen getracktes Head Mounted Display (HMD) (Lighthouse-Technologie) sowie Bewegungscontrollern zur Interaktion mit der virtuellen Welt. Die Lighthouse-Technologie erlaubt eine hochpräzise und zeitliche hochaufgelöste Positionsbestimmung des Nutzers innerhalb der 4,6 m × 4,6 m großen Interaktionsfläche (Abb. 3 rechts). Die visuelle Darstellung erfolgt mit einer Auflösung von 1080×1200 Pixel und einer Bildwiederholrate von 90 Hz.

Abb. 3: Komponenten (links) und schematischer Aufbau (rechts) des VR-Systems HTC Vive (Quelle: www.next-gamer.de / www.tomsguide.com)

Nach Sichtung und Vorbereitung aller notwendigen Daten erfolgt der Import in die Game Engine. Als Grundlage dient den Studierenden eine vom Hersteller bereitgestellte Vorlage, das VR-Template. Darin enthalten sind bereits die wichtigsten Funktionen zur Bewegung in der virtuellen Welt, wie die Teleportation zur Navigation und das Greifen von Objekten.

Ein anderer wesentlicher Aspekt der VR ist die Umgebung des zu visualisierenden Gebäudes/Bauwerkes. Ist ein reales Gelände in Form eines digitalen Geländemodells (DGM) vorhanden, so gibt es Möglichkeiten, dieses in Form eines höhencodierten Graustufenbildes, der sog. Heightmap in die Engine einzufügen. Für die Erstellung von Gelände wird das Landscape-System innerhalb der Game Engine eingesetzt. Landscape ermöglicht eine intelligente und daher schnelle Darstellung durch adaptive Auflösung des Geländes (Level of Detail) inklusive darauf erstellter Füllobjekte wie z.B. Bäume. Da in der Engine selbst jedoch generell nicht mit bekannten, geodätischen Koordinatensystemen gearbeitet wird, besteht hier Anpassungsbedarf, z.B. in Bezug auf die Skalierung des Gebietsausschnitts sowie der Höhendaten. Trotzdem stellt sich gerade bei Studierenden der Geodäsie der Anspruch, alle Geodaten lagerichtig zueinander und in den richtigen Dimensionen darzustellen. Dies ist bei einer VR, in der sich alle Größen im Maßstab 1:1 zum Betrachter darstellen, von großer Wichtigkeit. Zu dem Gelände sind auch umgebende Objekte wie weitere Gebäude etc. bei der Modellierung zu berücksichtigen. Hier muss eine sorgfältige Abwägung über den Detailgrad getroffen werden, um einerseits das Objekt nicht isoliert darzustellen und andererseits aber auch nicht vom eigentlichen Hauptobjekt abzulenken.
Neben den in Kapitel 5 genannten Interaktionen wird außerdem die Beleuchtungssituation der Szene generiert. In der Engine wird dabei unterschieden zwischen statischen und dynamischen Beleuchtungen. Statische Beleuchtungen lassen sich bereits während der Beschreibung voll berechnen und sind daher sehr performant. Sie lassen sich allerdings nur auf unbewegte Objekte anwenden, sodass immer bekannt sein muss, welche Objekte mobil und welche fest platziert (stationär) sind. Die zweite Beleuchtungsmöglichkeit wird durch die Verwendung von dynamischen Lichtern erzeugt. Da diese in Echtzeit mitberechnet werden müssen, sind sie sehr gezielt einzusetzen. Ein sehr wichtiger Bestandteil in der Projektbearbeitung liegt im Testen der erstellten Simulation, um die Intuition der Interaktionen, die Steuerung der VR-Applikation und das Leistungsverhalten des Systems laufend zu überprüfen. Unterschiedlich gute PC-Systeme standen den Studierenden dafür zur Verfügung, um einen Eindruck davon zu erhalten, welche Berechnungszeiten durch die visuelle Qualität der dargestellten Szenerie ausgelöst werden. Als Abschluss der Projektbearbeitung wird eine ausführbare Datei generiert. Dazu werden alle unnötigen Debug und Analyseelemente der Unreal Engine aus der VR herausgelöst und die Unabhängigkeit von der Entwicklungsumgebung geschaffen.

5 Generierte Virtuelle Realitäten

In diesem Kapitel werden die angefertigten virtuellen Realitäten kurz vorgestellt und einige Besonderheiten dazu erwähnt. Für alle Modelle wurden durch die Studierenden vielfältige Möglichkeiten der Interaktion und Animation innerhalb der VR-Applikation erarbeitet und umgesetzt. Dafür war es wichtig, sich mit der Geschichte und Bedeutung des Objektes näher auseinanderzusetzen. Bei den beiden Beispielen aus Äthiopien ging es darum, aus den Annahmen und Forschungsergebnissen der Archäologen und Architekten umsetzbare Elemente zu kreieren. Für den Almaqah-Tempel wurde deshalb dem Opfer-Altar im inneren des Bauwerkes eine zentrale Rolle zugedacht (Abb. 4 Mitte). Dem Besucher des virtuellen Tempels wird es ermöglicht, ein historisches Ritual virtuell durchzuführen, in dem eine Ziege geopfert wird, bei der sogar das Blut spritzt (Abb. 4 rechts).

Abb. 4: VR des Almaqah-Tempels in Yeha, Äthiopien; Außenansicht (links) und Innenraum (rechts)

Für das Monumentalbauwerk Grat Be´al Gebri (Abb. 5) sind zur früheren Nutzung noch keine gesicherten Erkenntnisse vorhanden (SCHNELLE 2015), so dass hier viel kreativer Freiraum bestand. Bei der VR-Umsetzung bestand ein wesentlicher Punkt in der richtigen Installation der Beleuchtung, da der Bau sehr wahrscheinlich fensterlos war. Daher wurden einige Interaktionen wie tragbare Fackeln, entzündbare Wandlampen und offene Feuerstellen umgesetzt (Abb. 5). Zusätzlich wurden Fundstücke aus dem Grabungsareal bzw. der näheren Umgebung als 3D-Modelle zur „Möblierung“ des Gebäudes und der Szenerie hinzugefügt.

Abb. 5: VR des Grat Be´al Gebri in Yeha, Äthiopien; Außenansicht (links/rechts) und Innenraum (Mitte)

Die Grundlage für die Konstruktion des CAD-Modells des Salomonischen Tempels war das Holzmodel von 1692, so dass hier nicht viele Hinweise auf die verwandten Materialien des ursprünglichen Bauwerkes vorhanden waren. Daher konnten sich die Studierenden bei der Umsetzung der VR-Applikation einerseits an vorhandener Literatur z.B. von PRICE (2013) orientieren, aber anderseits auch eigene Kreativität entwickeln. Allein die Skalierung auf den Maßstab 1:1 war nicht einfach zu ermitteln, da sämtliche überlieferten Größen des ursprünglichen Bauwerks nicht durch gesicherte Einheiten in heutige Maßeinheiten umzurechnen waren. Geometrische Unterschiede zwischen der Literatur und dem zu Grunde liegenden Holzmodell wurden nicht aufgelöst. Abb. 6 zeigt die generierte Szenerie vor dem Hauptgebäude (links), die Frontfassade des Hauptgebäudes mit dem Eingang zum Heiligtum (Mitte) und eine Ansicht vom Inneren des Tempelbauwerks (rechts).

Abb. 6: VR des Salomonischen Tempels: Szenerie vor dem Hauptgebäude (links), Frontfassade mit Eingang des Hauptgebäudes (Mitte) und Ansicht vom Inneren des Tempelbauwerks (rechts)

Abb. 7: VR der Sternwarte Hamburg-Bergedorf – Nachtansicht der beiden Gebäude Äquatorial- (links) und Meridianhaus (Mitte) sowie von einem Fernrohr (rechts)

Für die Erstellung einer VR-Applikation von der Sternwarte Hamburg-Bergedorf wurden die beiden Gebäude Meridiankreis und Äquatorial verwendet (Abb. 7). Diese beiden modellierten Gebäude sind heute nicht mehr aktiv in Benutzung, da ihre Fernrohre nicht mehr in situ vorhanden sind. Daher wurden jeweils in das virtuelle Gebäude zwei nicht originale, aber frei verfügbare Fernrohre integriert. In einem Gebäude kann interaktiv die Kuppel geöffnet werden, um durch das Fernrohr in den Sternenhimmel zu sehen. Analog zur früheren realen Verwendung der Teleskope sind nur ausgewählte Bewegungen möglich, z.B. Nord-Süd-Bewegung beim Meridiankreis.

Abb. 8: VR der Landdrostei Pinneberg – Außenansicht des Gebäudes (links) und Innenansicht eines Raumes (rechts)

Die Landdrostei in Pinneberg ist durch ihre Lage mitten in der Stadt in der Gestaltung der Umgebungstopographie sehr festgeschrieben, da sowohl die umgebenden Gebäude als auch das unmittelbar umgebende Gelände bekannt sind. Dieser Umstand wurde bei der Erstellung der VR-Applikation weitestgehend berücksichtigt (Abb. 8 links). Für die Ausgestaltung des Inneren des

Gebäudes bot die VR-Applikation mehr Freiheiten. So wurde ein Raum sowohl mit heutigem Mobiliar (Abb. 8 rechts), als auch mit historischer Einrichtung von vor ca. 200 Jahren ausgestattet, zwischen denen der Besucher interaktiv beim Begehen des Raumes hin- und herschalten kann.

Abb. 9: Verschiedene Ansichten in der VR-Applikation des Westerturms in Duderstadt

Der Westerturm in Duderstadt (Abb. 9) schließlich lag sowohl von außen als auch innen sehr detailliert modelliert vor, so dass hier in erster Linie die Ausstattung der Umgebung und des Gebäudeinneren, welches als Museum genutzt wird, durch z.B. Infotafeln und Ausstellungsstücke gestaltet werden konnte. Schmale Treppen und Wendeltreppen im Turm selber boten Gelegenheit zu speziellen Bewegungsformen wie bei einer virtuellen Rolltreppe bzw. bei einem virtuellen Aufzug (Abb. 10 links). In der Umgebung wurden zusätzlich zur Belebung der Szene diverse Gegenstände (Abb. 10 rechts) und Personen wie z.B. eine tanzende Frau animiert.

Abb. 10: Treppe im Westerturm (links) und Mobiliar für die VR-Applikation (rechts)

6 Fazit & Ausblick

In diesem Beitrag wurde das Lehrkonzept für die Veranstaltung 3D-Visualisierung im Masterstudiengang Geodäsie und Geoinformatik der HafenCity Universität Hamburg vorgestellt. Es werden die generelle Projektbearbeitung und die Ergebnisse in Form der verschiedenen VR-Applikationen gezeigt. Am Beispiel des hochaktuellen Themenfeldes VR wird demonstriert, wie wichtig neue, digitale Informationsquellen abseits klassischer Fachliteratur sind und wie sie sich über Gruppen- bzw. Projektarbeit zur praktischen Kompetenzaneignung eignen. Die Ergebnisse der erstellen Virtuellen Realitäten verdeutlichen sehr gut, welches Potential einerseits in den

Geodaten liegt, aber wie wenig anderseits die reine Geometrie ohne Kontext erzählt. So können interaktive Visualisierungen die Informationsvermittlung in VR übernehmen. Sie erfordern aber einen hohen Grad an Intuition für die Benutzersteuerung, die während der VR-Erstellung neben der sehr präzise abgestimmten Informationsbereitstellung intensiv zu evaluieren ist.
Das Modul Visualisierung im Masterstudiengang Geodäsie und Geoinformatik an der HafenCity Universität Hamburg zeigt, wie der Lerninhalt und die Lernkompetenz im Laufe der Zeit auf die neuen technologischen Entwicklungen und Potentiale zum Nutzen der Studierenden und Lehrenden angepasst wird. Aus der Ökonomisierung der Lehre ergibt sich der Auftrag an die Lehrenden, noch besser seinen Zeitaufwand einzuplanen und ggf. bessere Formen der Kompetenzevaluation zu nutzen. Für die Studierenden ergibt sich sowohl eine große Herausforderung als auch ein hoher Lerneffekt durch die Größe und Komplexität der Aufgabe und der damit verbundenen Verantwortung in der Lern- und Projektorganisation innerhalb der Gruppe.
Für die interaktive VR-Visualisierung eignen sich dabei nicht nur einzelne Gebäude oder Bauensembles, sondern auch ganze Ortschaften und Städte. Eine gute Quelle für historische Daten sind dabei physische Stadtmodelle, die in Museen oder als Bronzemodelle auf Marktplätzen ausgestellt werden. Eine solche Datengrundlage wird im aktuellen Wintersemester 2017/2018 eingesetzt, um das historische Stadtbild von Stade, Niedersachsen im Jahr 1620 anhand des gescannten physischen Stadtmodells aus dem Stader Rathaus in VR erlebbar zu machen.

7 Literaturverzeichnis

ANDERSON, P. L., PRICE, M., EDWARDS, S. M., OBASAJU, M. A., SCHMERTZ, S. K., ZIMAND, E. & CALAMARAS, M. R., 2013: Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. Journal of consulting and clinical psychology, **81**(5), 751-760.

BURDEA, G. C. & COIFFET, P., 2003: Virtual Reality Technology (Vol. 1). John Wiley & Sons.

DEGGIM, S., TSCHIRSCHWITZ, F. & KERSTEN, T., 2017a: Entwicklung eines virtuellen Museums für ein historisches Gebäude am Beispiel des Alt-Segeberger Bürgerhauses. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 26, Tagungsband der 37. Wissenschaftlich-Technischen Jahrestagung der DGPF, 8.-10. März 2017 in Würzburg (auf CD), 454-463.

DEGGIM, S., KERSTEN, T., TSCHIRSCHWITZ, F. & HINRICHSEN, N., 2017b: Segeberg 1600 – Reconstructing a Historic Town for Virtual Reality Visualisation as an Immersive Experience. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **42**(2/W8), 87-94.

FOHRMANN, O., 2016: Im Spiegel des Geldes, Bildung und Identität in Zeiten der Ökonomisierung. transcript Verlag, Bielefeld, 180 S.

FURMAN, E., JASINEVICIUS, T. R., BISSADA, N. F., VICTOROFF, K. Z., SKILLICORN, R. & BUCHNER, M., 2009: Virtual reality distraction for pain control during periodontal scaling and root planing procedures. The Journal of the American Dental Association, **140**(12), 1508-1516.

GAITATZES, A., CHRISTOPOULOS, D. & ROUSSOU, M., 2001: Reviving the Past: Cultural Heritage meets Virtual Reality. Proceedings of the Conference on Virtual Reality, Archaeology and Cultural Heritage, 103-110.

INTERRANTE, V., ANDERSON, L. & RIES, B., 2006: Distance Perception in Immersive Virtual Environments, Revisited. Proceedings IEEE Virtual Reality (VR 2006), 3-10.

KERSTEN, T., ACEVEDO PARDO, C. & LINDSTAEDT, M., 2004: 3D Acquisition, Modelling and Visualization of north German Castles by Digital Architectural Photogrammetry. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, **35**(B2), 126-132.

KERSTEN, T., BIEBERMANN, M. & SCHNEIDER, M., 2006: 3D-Erfassung und Modellierung des Duderstädter Westerturmensembles durch Kombination von digitaler Architekturphotogrammetrie und terrestrischem Laserscanning. Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage 2006, Th. Luhmann/C. Müller (Hrsg.), Wichmann Verlag, Heidelberg, 254-263.

KERSTEN, T., HINRICHSEN, N., LINDSTAEDT, M., WEBER, C., SCHREYER, K. & TSCHIRSCHWITZ, F., 2015: Baugeschichtliche 3D-Dokumentation des Alt-Segeberger Bürgerhauses durch Photogrammetrie und Terrestrisches Laserscanning. Tagungsband der Denkmäler3.de 2013 - Von low-cost bis high-tech: 3D-Dokumentation in Archäologie & Denkmalpflege, H.-J. Przybilla, T. Kersten & F. Boochs (Hrsg.), LWL Industriemuseum Zeche Zollern Dortmund, 16.-18. Oktober 2013, 30-37.

KERSTEN, T., TSCHIRSCHWITZ, F., LINDSTAEDT, M. & DEGGIM, S., 2018: The Historic Wooden Model of the Solomon Temple – 3D Recording, Modelling and Immersive Virtual Reality Visualisation. Journal of Cultural Heritage Management and Sustainable Development, Emerald Publishing, im Druck.

LAWSON, E., 2016: Game Engine Analysis. https://www.gamesparks.com/blog/game-engine-analysis/, letzter Zugriff 22. Januar 2018.

LINDSTAEDT, M., MECHELKE, K., SCHNELLE, M. & KERSTEN, TH., 2010: Virtuelle 3D-Rekonstruktion des Almaqah-Tempels von Yeha in Äthiopien mit Hilfe terrestrischer Laserscanning-Daten. Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage 2010, Th. Luhmann/Ch. Müller (Hrsg.), Wichmann, VDE Verlag GmbH, Berlin und Offenbach, 50-57.

O'FLANAGAN, J., 2014: Game Engine Analysis and Comparison. https://www.gamesparks.com/blog/game-engine-analysis-and-comparison/, letzter Zugriff 22. Januar 2018.

PRICE, R., 2013: Inner Cubes Handbuch zum Tempel. First edition, Inner Cube, 152 S.

RHEINGOLD, H., 1991: Virtual Reality: Exploring the Brave New Technologies. Simon & Schuster Adult Publishing Group.

REINMANN, G., SPORER, T. & VOHLE, F., 2007: Bologna und Web 2.0: Wie zusammenbringen, was nicht zusammenpasst. eUniversity–Update Bologna, 3, 263-278.

SCHNELLE, M., 2015: Yeha, Äthiopien: Die Restaurierungsarbeiten am Monumentalbau Grat Beʿal Gebri. e-Forschungsberichte des Deutschen Archäologischen Instituts, eDAI-F 2015-3, 7-12.

STEINICKE F., BRUDER G., HINRICHS, K.H. & STEED A., 2010: Gradual transitions and their effects on presence and distance estimation. Computers & Graphics, **34**(1), 26-33.

SUTHERLAND, I. E., 1965: The ultimate display. Proceedings IFIP Congress, 506-508.

Comparison of Power Law Tropospheric Correction for Time Series InSAR Application

NOORLAILA H. ISYA[1 2], ANIKA RIEDEL[1], BJÖRN RIEDEL[1] & WOLFGANG NIEMEIER[1]

Abstract: Radar interferometric technology is nowadays considered as one geodetic method to observe earth surface changes. The use of time series analysis in SAR imageries could figure out slow periodical displacement rates on natural hazard phenomena such as mass movement in landslide prone areas and crustal deformation in tectonic areas. However, this multi-temporal SAR results could be biased due to effect of atmospheric phase screen (APS) when microwave signals are going through tropospheric layer. Power law tropospheric correction is a method that recently used on InSAR application to minimize atmospheric noise both spatially and temporally. The method attempts to find a correlation of phase delay to hydrostatic and wet components in troposphere properties. Thus, we try to estimate the tropospheric delay on individual interferograms and apply Small Baseline Subset (SBAS) on InSAR in different regions. The case study of a small region is represented in Ciloto, West Java, Indonesia and for a large region, we chose an area located at northern Baja California, Ensenada, Mexico. 50 Sentinel-1A ascending SAR Imageries have been proceed in Ciloto Landslide prone area from October 2014 to September 2017. In term of detecting the deformation signal at Baja California, we tested 44 Envisat Descending data from February 2003 to October 2010. Sounding data are needed to calculate tropospheric delay at different altitudes. They are taken from International Upper Air network data which provide variables of atmospheric pressure, temperature and relative humidity. We find the estimation of phase delay using topography correlated by a power law and reference height could improve the displacement's accuracy value. Nevertheless, the attention either under or over estimation due to band-filtering performance and correlated height terrain should be watched out for every interferogram. We then conduct general assessment before and after power law tropospheric correction for small baseline InSAR results and compare them to measurements from Extensometer installed in Ciloto. For the large area in northern Baja California, Mexico, the phase-based power law correction is compared to atmospheric delay maps generated by MERIS, a spectrometer instrument on board Envisat, in condition to free cloud cover 50% and 90% at the study area.

1 Introduction

Interferometry Synthetic Aperture Radar is one promising techniques to measure the earth surface changes on space borne. A considerable number of research studies research have identified seismic activities such as earthquake 's impact (FUNNING et al. 2005), subsidence due to over-exploitation of ground water (CHEN et al. 2016) and oilfields (LIU et al. 2015), the behaviour of volcanoes (HOOPER et al. 2007), and slow movement in landslide prone areas (CASCINI et al. 2010). Permanent Scatter (PS) (FERRETTI et al. 2001) and SBAS techniques

[1] Technische Universität Braunschweig, Bienroder Weg 81, D- 38106 Braunschweig, E-Mail: n.isya@tu-bs.de
[2] Institut Teknologi Sepuluh Nopember, Jl. Raya ITS – 60111 Surabaya, Indonesia

(BERARDINO et al. 2002) are advance InSAR methods to generate real deformation and reduce unwanted signals coming from orbit and topographic errors, atmospheric effects, and noises. Some spatial correlated terms, orbit to height errors, could be removed from interferogram considering to the use of abundant SAR imageries and finding out the stable scatters. Previously, atmospheric terms could be filtered out as well when we consider that atmospheric signals are random in time (e.g., FERRETTI et al. 2001; BERARDINO et al. 2002). However, a recent study proved that they are also temporally correlated to stratified tropospheric delays (DOIN et al. 2009). Thus, removing the atmospheric phase screen especially for tropospheric delay patterns using low-pass filtering might be miss interpretation of certain deformation signals as a noise or otherwise. If we observe some high level deformations of large scale unit (cm to m) at relative small area, we could neglect the effect of atmospheric phase screen due to the same of phase characteristics between pixels. Nevertheless, the monitors of crustal deformation, subsidence and slope movement could be considered as a very low rate displacement (mm to cm) whose InSAR practicals have to pay attention to atmospheric artifacts performing more accurate results of displacements.

Two atmospheric layers which mostly effect microwave signals are propagating through ionosphere and troposphere. Ionospheric delay mostly finds on L band sensor, ALOS PALSAR, and will be neglected on the processing procedure. The focus study is correcting the tropospheric delay using power law function (BEKAERT et al. 2015) on InSAR process and how the correction impacts the results of displacement value both in Indonesia and Mexico areas. For the Indonesia case, we observe slow movement in Ciloto, West Java, where the area has a high risk of landslide hazard using 50 SAR images of Sentinel 1A ascending orbit. Furthermore, we attempt to detect crustal deformation near to Agua Blanca Fault and additionally a subsidence around agriculture area in Manaedero Valley using 44 images of Envisat descending orbit.

2 Study Areas and Method

The troposphere delay correction will be tested on two case studies, which are Ciloto, Indonesia and Baja California, Mexico. Both of those are located on the mountainous terrain where probably having high precipitation on the top terrain so that it may has a different troposphere propagation signal between the flat and hilly area. Hence, these delay variations could cause miss interpretation of interferograms (GOLDSTEIN1995). HANSSEN (2001) distinguish two types of atmospheric signal based on their physical origin. The first one, turbulent mixing, is caused by turbulent process in the atmosphere. Turbulent delay could affects both flat and mountain terrain. It makes phase artifacts on the interferogram due to spatial heterogeneity in the refractivity. Secondly, vertical stratification is mostly correlated to topography and only affects mountainous terrain (MASSONET & FEIGL 1998). It is caused by different vertical refractivity profiles between master-slave SAR acquisitions considered as no heterogeneities within the horizontal layers.

Estimation of tropospheric delay might be described with a mathematical model which is generally divided as systematic and stochastic component. Figure 1 explains shortly the scheme of atmospheric delay that most recent atmospheric studies categorizing the delay estimation.

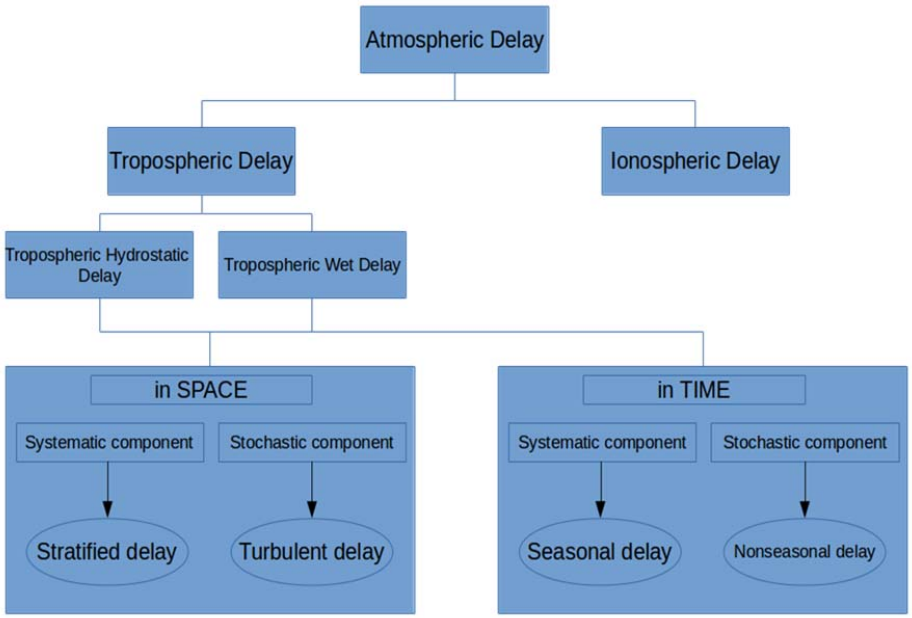


Fig. 1: The scheme of atmospheric delay that might influence interpretation between SAR acqusitions

According to the troposphere layer, it contains of hydrostatic delay and wet delay. The determined parameters of hydrostatic delay are the refractivity from dry air (atmospheric pressures and temperature) and of wet delay from wet air (partial pressure of water vapor and temperature). SMITH & WEINTRAUB (1953) characterized the phase delay through atmosphere using the refractivity (N), with

$$N = \left(k_1 \frac{P}{T}\right)_{hydro} + \left(k'_2 \frac{e}{T} + k_3 \frac{e}{T^2}\right)_{wet} \tag{1}$$

where P = total atmospheric pressure, T = temperature and e = partial pressure of water pavor, k1, k'2, k3 are constants parameter with 77.6 K hPa-1, 23.3 K hPa-1 and 3.75.105 K2 hPa-1, respectively (THAYER 1974). Furthermore, if we consider tropospheric phase delay (ϕtropo) between pixels (p,q) and SAR acquisitions along radar line of sight (HANSSEN 2001), it can be defined as

$$d_{tropo} = 10^{-6}(\cos\theta)^{-1}\left[\int_{h_p}^{h_{top,p}} N(h)\,dh - \int_{h_q}^{h_{top,q}} N(h)\,dh\right] \tag{2}$$

$$\phi_{tropo} = \frac{-4\pi}{\lambda}\, d_{tropo} \tag{3}$$

where d_{tropo} = the tropospheric (slant) delay, θ = the incidence angle, λ = the radar wavelength and h = height coresponds to the integration of the refractivity between h and the top of the troposphere (h_{top}).

Several main approaches to estimate tropospheric delay have been conducted to correct InSAR results (Figure 2). Three main source model are based on numerical weather model, for instance,

(i) high spatial and temporal resolution weather forecast models (WRF) (JUNG et al. 2014), (ii) global atmospheric reanalysis models with low spatial and temporal resolution ERA-1 (~80 km), MERRA2 (~ 55 km) from the European Centre for Medium-Range Weather Forecasts, and (iii) Generic Atmospheric Correction Online Service for InSAR (YU et al. 2018). The second one is from satellite spectrometers, which use observations of atmospheric water vapor from Moderate Resolution Imaging Spectroraidometer (MODIS) and Medium-Resolution Imaging Spectrometer on board Envisat (MERIS) (LI et al. 2009). Another strategy is using a phase based model itself with main purpose to reduce stratified delay correlated with topography variations. A former method of phase based use linear relation between tropospheric phase delay and topography and recently improved by using a spatially variable power law approach (BEKAERT et al. 2015). On this article, we assess only a power law method and test the tropospheric correction at small area in Ciloto and large area in northern Baja California.

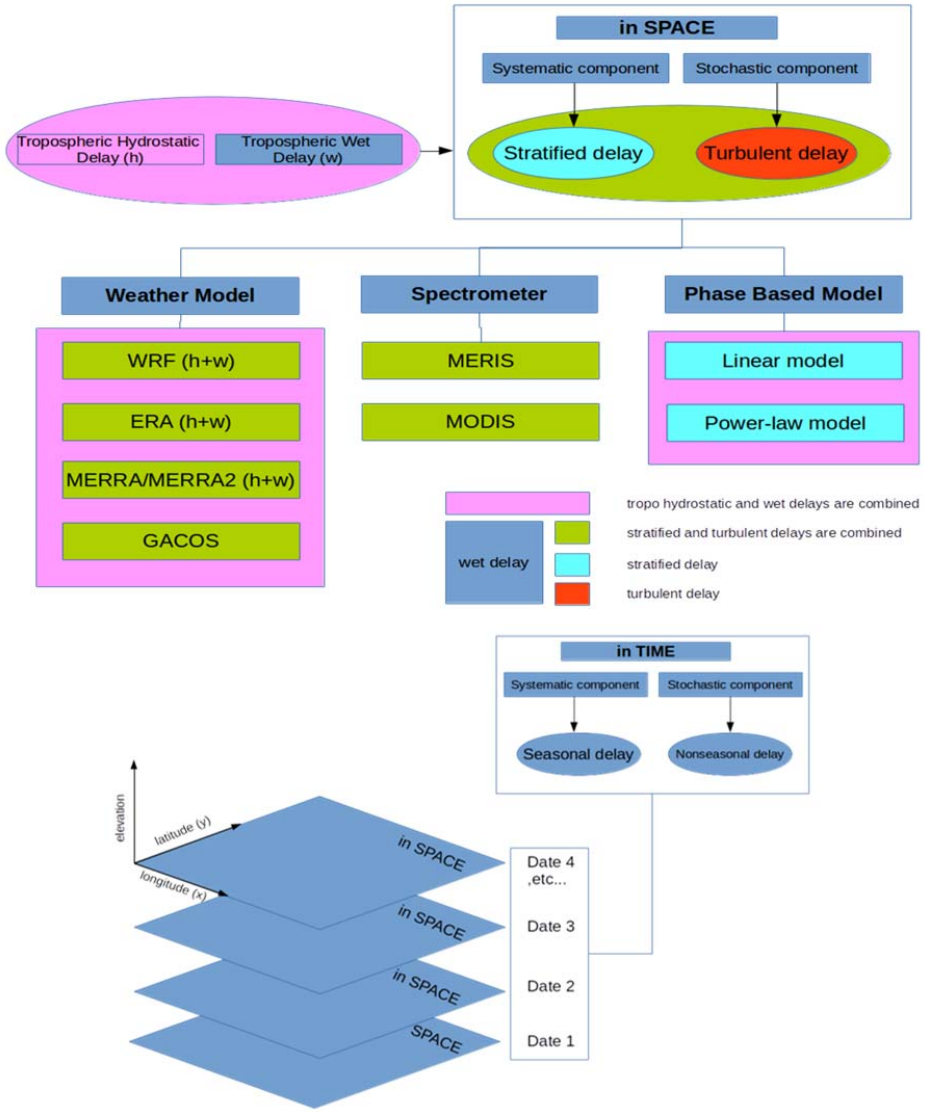


Fig. 2: Different source and strategies for estimating tropospheric delay in space (top) and in time (bottom)

2.1 Ciloto, West Java, Indonesia

Ciloto is located on a hilly ridge to creek valley southern of Cijember River (Figure 3). This area has been built for residence, tourism (Puncak) and agriculture. The main type of rock around Ciloto is Quaternary volcanic rocks consisting of basalt, andesitic tuffs especially breccia, and lava from local vents. In recent years, many evidences for ground displacements have been occurred in Ciloto. Starting from the top of valley's body, ground movements were about 1-2 m located at a resort hotel reported in 2014. The swimming pool of the hotel could not be operated any more since the incident. In the same year, 200 m away to first evidence, 33 families were evacuated from their houses since cracks have been found on several house's walls. Moreover, there was also a total broken house due to miss construction and easily triggered by the ground movements. Therefore, Technical Research Center for Geological Hazard (BPPTKG) has installed extensometer near to the resort hotel since 2014 because of the phenomenon. On next section, we compare the LOS displacement from time series InSAR to graphic displacements from extensometer in the same time overlap.

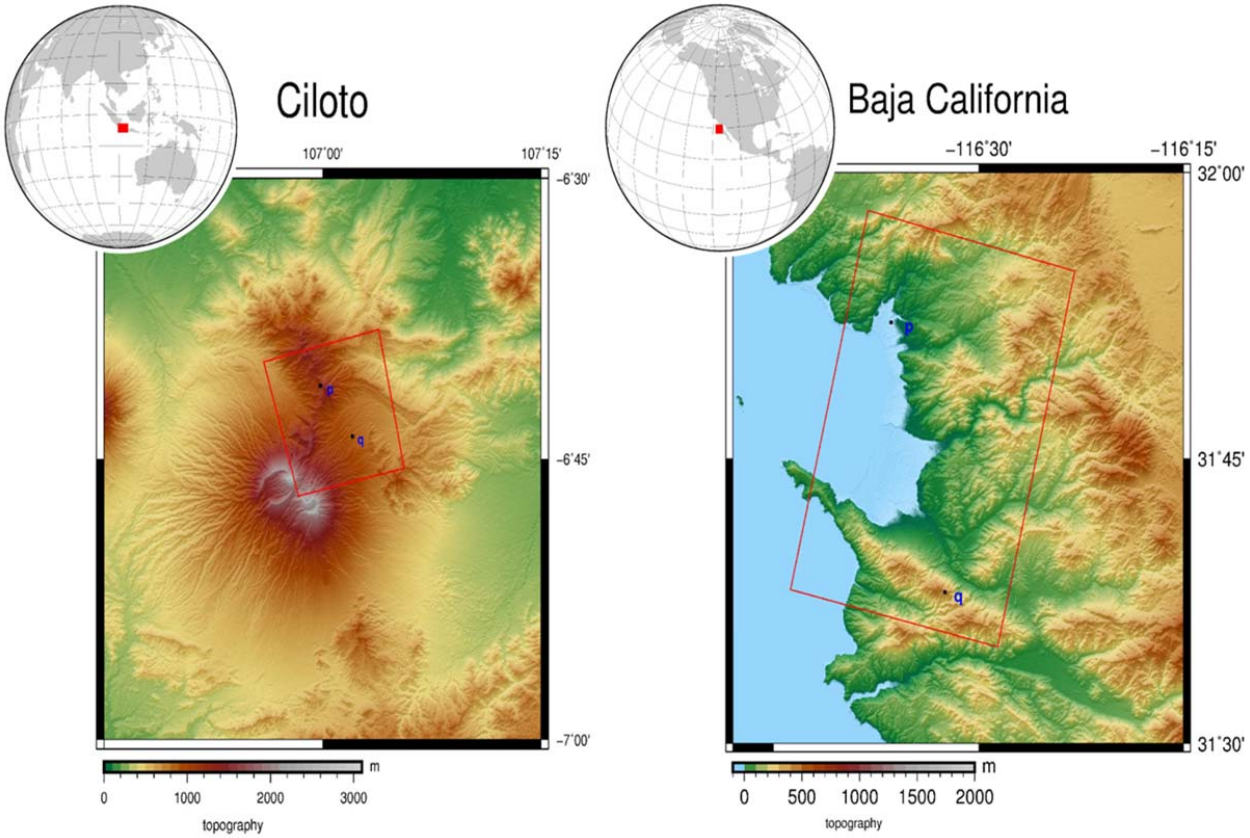


Fig. 3: Study areas: (left) Ciloto, West Java, Indonesia with the red rectangle line showing Sentinel-1A ascending scene borders (Long E:105.5° – 108.5° and Lat S: -5.5° - 7.5°) and (right) Northern Baja California, Ensenada, Mexico with the red rectangle line showing Envisat descending scene borders (Long W: 117.5° – 115.5° and Lat N: 30.5° – 32.5°). Pixel p and q refer to different phase delays showed on Fig. 4 and Fig. 5.

Furthermore, several landslides occurred at Puncak highway, the middle part of study area. *Indonesian National Board for Disaster Management* (BNPB) recorded two big landslide

hazards in Bandung and Bandung Barat, West Java at 23th February 2010 and 25th March 2013 with 33 died, 17 were injured, 936 lost their houses and 14 died, 23 were injured and 185 lost their houses, respectively. Even two landslides periodically were occurred on the same location, Raya Puncak Highway at Km.81, close to Bumi Aki restaurant. The first mass movement was in 2015 placed next to the body of Raya Puncak highway and damaged the local infrastructure and plant field around the area. The second event happened in rainy season, February 2017. This hazard belongs to a small type of mass wasting which the impact area was only ± 25 x 25 m. However, the deposit debris flow right into Raya Puncak highway and disconnected one of the most important access pathways between Bogor and Bandung city.

In general the effect of atmospheric phase screen is visible mostly in large area, such as, crustal deformation monitoring. However, Ciloto is placed on mountainous terrain and has high relative humidity. We demonstrate the temporal variation of the hydrostatic delay and wet delay between two pixel p [106.9976° ; -6.6851°] located on a hilly area and pixel q [107.0343° ; -6.7299°] located on a flat area. The length between the two pixels is about 6.5 km and 380 m elevation differences. Since the SAR scene was cropped to the small coverage to evaluate only at the prone landslide region, we do not attempt to figure out the behavior of tropospheric phase delay in such large area.

The temporal variation of hydrostatic and wet delay is calculated during the Sentinel 1A SAR acquisitions from October 2014 – September 2017 at 11:00 – 12:00 hour using weather model from ERA-1 Interim. Figure 4 describes the effect of slant hydrostatic delay (SHD) and slant wet delay (SWD) between these two pixels. The seasonal component is assumed to a periodic signals described as Fourier series model, a summation of sine and cosine functions with annual and biannual components (FATTAHI et al. 2015) using the following function

$$\text{SHD or SWD} = a_o + \sum_{i=1}^{2} a_i \cos(2\pi i t) + b_i \sin(2\pi i t) \tag{4}$$

$$\text{STD} = \text{SHD} + \text{SWD} \tag{5}$$

where a_o = intercept, a_1,b_1 = the coefficients of annual components, a_2,b_2 = the coefficients of semi-annual components, t = period (every day in SAR acquisitions' duration) and STD = Slant Total Delay. We assume the stochastic components of time series of relative delay are residual of ERA-1 and best fitted seasonal. The relative delay and seasonal component could be plotted as Gaussian distribution in which the standard deviations of SHD, SWD, SHD are 0.0342, 0.4545, 0.4438, respectively. The highest and lowest peak of SHD is -14.6630 cm and -14.4246 cm, for SWD is -6.6920 cm and -3.7747 cm, respectively. The interval delay for SHD and SWD is 0.2385 cm and 2.9173 cm for 3 years. Hence, we could find out the interval between highest and lowest peak of STD about 2.9659 cm.

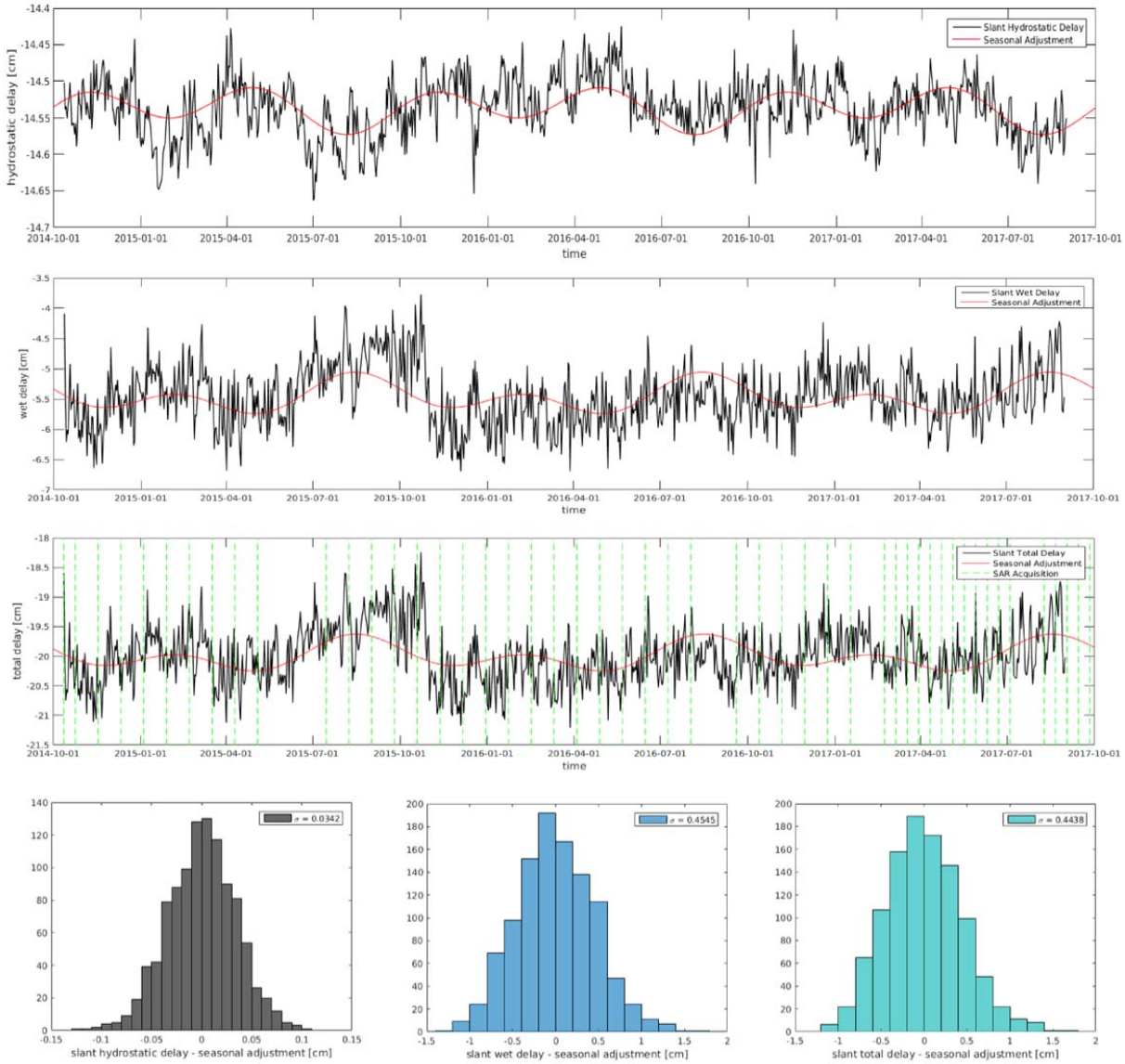


Fig. 4: (The top rows) Time series of relative slant hydrostatic delay (top), slant wet delay (middle) and slant total delay (below) from ERA-1 with seasonal model for each delay between pixels [p,q] in Ciloto region. The green line is Sentinel-1A date acquisitions. (The bottom row) Non-seasonal component with Gaussian distribution.

2.2 Northern Baja California, Ensenada, Mexico

The other second study area is located in an arid coastal region in the northern part of the Peninsula of Baja California, 110 km south from the US-México border line (Figure 3). The region is morphologically characterized by coastal and alluvial flatlands, where the city of Ensenada and the croplands of Maneadero are located. These flatlands are surrounded by the Guadalupe and Ojos Negros intermountain valleys. The Agua Blanca fault is the southern border of our study area and extends from NW to SE as a 120 km length of dextral strike-slip fault. The geological background of this area was explained more at previous works (DIXON et al. 2002; ANIKA et al. 2018).

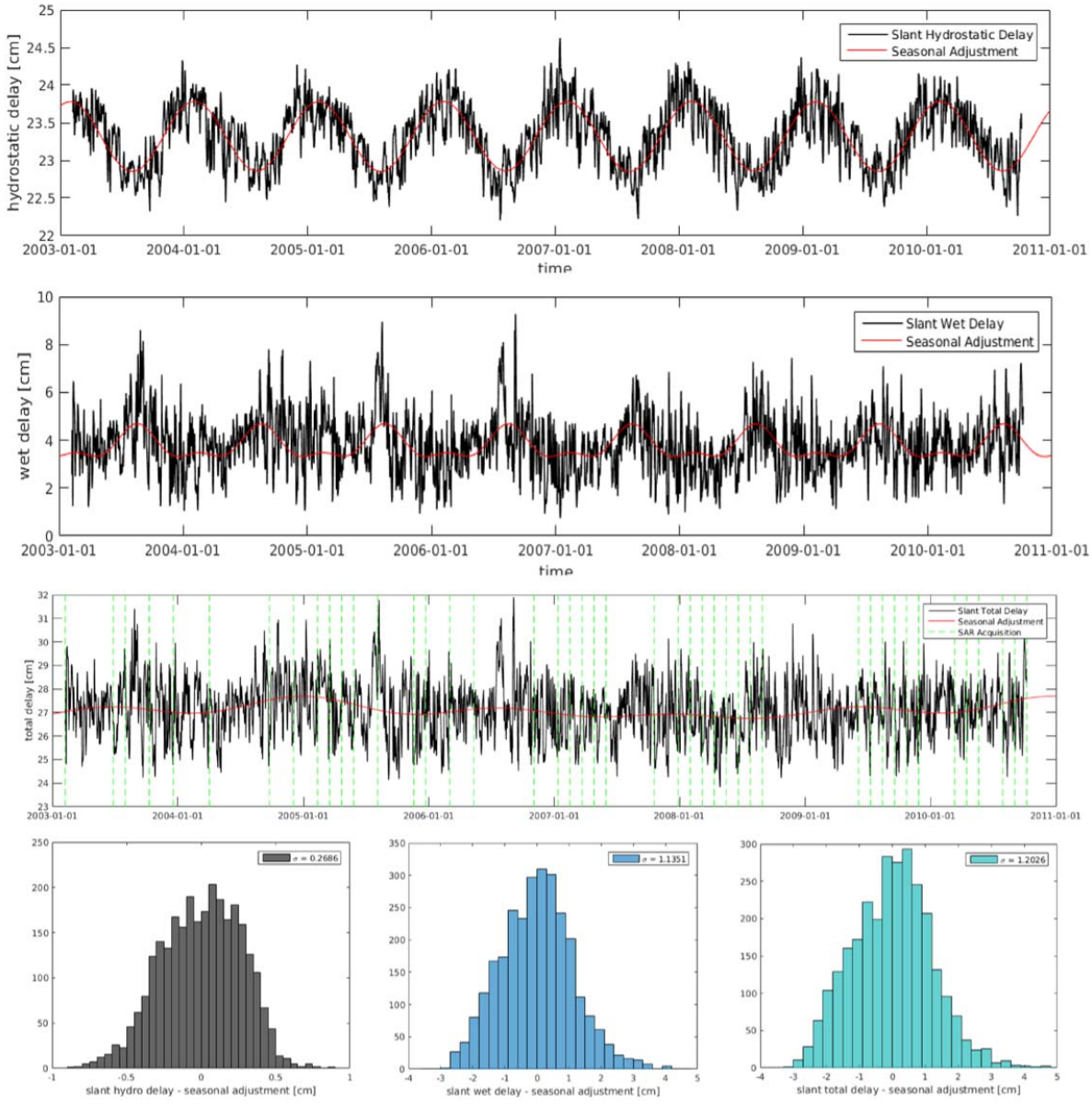


Fig. 5: (The top rows) Time series of relative slant hydrostatic delay (top), slant wet delay (middle) and slant total delay (below) from ERA-1 with seasonal model for each delay between pixels [p,q] in Ensenada region. The green line is Envisat date acquisitions. (The bottom row) Non-seasonal component with Gaussian distribution.

Northern Baja California has a big possibility to observe earth's surface change, for example, crustal deformation due to tectonic activity and subsidence due to over-exploitation through water pumping in agricultural areas like in Maneadero Valley. The region has various terrains from flats in El Bajio and Maneadero Valley to mountainous on Punta Banda Ridge. Therefore, atmospheric phase screen could affect microwave propagating signal especially through tropospheric layers in this region and make a dominant stratified delay in the interferogram.

As the section 2.1, we demonstrate also the temporal variation of the hydrostatic delay and wet delay between two pixel *p* [-116.6069°; 31.8690°] located on flat terrain and pixel *q* [-116.5410 °; 31.6327 °] located on mountainous terrain. The height difference between pixels is 910 m with

29 km length. The calculation of seasonal and non-seasonal delay is the same as what we applied in Ciloto region. The relative delay and seasonal component could be plotted as Gaussian distribution in which the standard deviations of SHD,SWD,SHD are 0.2686, 1.1351, 1.2026, respectively. The highest and lowest peak of SHD is 22.2033 cm and 24.6427 cm, for SWD is 0.7407 cm and 9.2680 cm, respectively. The interval delay for SHD and SWD is 2.4214 cm and 8.5274 cm for 8 years. Hence, we could find out the interval between highest and lowest peak of STD about 8.0610 cm. According to Figure 5, the seasonal model with the same summation of sine and cosine functions is not best fitted to STD since the periodic harmony between SHD and SWD in Ensenada is quite different and clearly shows a greater range interval delay (8.06 cm) than its in Ciloto (2.96 cm). The high peak of seasonal delay for SHD is occurred on every October-November annually. On the other hand, SWD has high peak of it on every January - March annually.

2.3 InSAR Processing

We generate the interfrograms using GMTSAR (SANDWELL et al. 2011) for Sentinel-1A imageries in Ciloto, Indonesia and DORIS (KAMPES et al. 2003) for Envisat imageries in Ensenada, Mexico. Furthermore, we select Persistent Scatter (PS) candidates with Slowly Decorellating Filter Phase (SDFP) algorithm. They are processed by StaMPS, a non-commercial statistical InSAR software, developed by HOOPER (2012). All interferograms are selected by small temporal and perpendicular baseline parameters. We choose master-slave pairs with baseline less than 150 days; 150 m in Ciloto and 365 days; 500 m in Ensenada adding additional pairs to make sure there is no isolated network. The final small baseline configuration could be seen on Figure 6 and StaMPS processing parameters both in study areas are described in Table 1. Additionally, the phase contribution due to oscillator drift is also removed in Envisat processing.

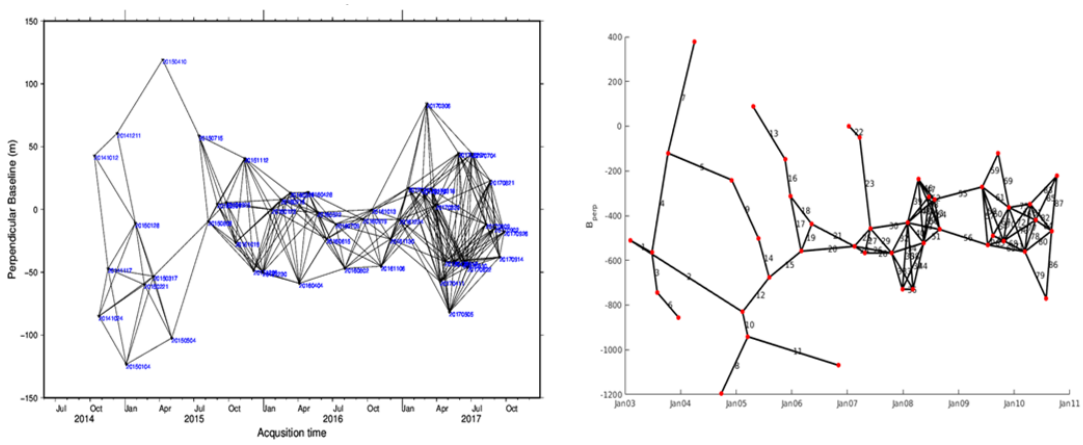


Fig. 6: Small baseline pairs from 50 images Sentinel-1A Ascending in Ciloto, Indonesia (left) and 44 images Envisat Descending in Baja California, Mexico (right).

Tab. 1: InSAR processing Parameters on StaMPS for both study areas

Parameter	Value		Parameter	Value	
	Mexico	Indonesia		Mexico	Indonesia
DEM (SRTM)	30 m	30 m	SCN time win	365	365
Max DEM error	5	5	Merge resample size	100	0
Filter grid size	50	50	Unwrap grid size	200	200
CLAP α	1	1	Unwrap time win	730	730
Spatial density rand	2	2	Unwrap goldstein filter	0.8	0.8

2.4 Power law Tropospheric Delay Correction

The power law method accounts variations spatially and temporally at tropospheric properties and belongs to phase-based tropospheric delays. It corrects only stratified delay and turbulent mixing still remains. How to reduce the effect of turbulent mixing will be explained at Section 3. The interferometric tropospheric phase $\Delta\phi_{tropo}$ could be estimated using the relationship between the interferometric phase and the topography described by BEKAERT, 2015 as

$$\Delta\varphi_{tropo} = K'_{\Delta\varphi}(h_o - h)^{\alpha}, with\ h < ho \tag{6}$$

where $K'_{\Delta\phi}$ = coefficient relating phase to topography with spatial and temporal variation, h_o = the power law reference height, α = power law coefficient. α is an empirically constant calculated from balloon sounding data provided by International Upper Air network data, the Department of Atmospheric Science of the University of Wyoming. We used WIII Jakarta Station (106.65° ; -6.11° ; Number 96749) approximately 78 km away from Ciloto, Indonesia and NKX Sand Diego Station (-117.12° ; 32.85° ; Number 72293) approximately 130 km away from Baja California, Mexico. These sounding measurements provide atmospheric properties, such as, pressure, temperature, relative humidity and wind speed, to calculate the refractivity (equation 1), tropospheric delay (equation 2 & 3) hence the relative tropospheric delay with $\Delta d_{tropo} = d_{tropo}(t_{master}) - d_{tropo}(t_{slave})$.

The power law exponent α could be found from the slope of linear fit by plotting log-log relationship between the mean tropospheric delay d_{tropo} and relative height (h_o-h). h_o is being set to the altitude with estimating no relative delays ($\Delta d_{tropo} \approx 0$) between acquisitions as illustrated on Figure 7. Furthermore, $K'_{\Delta\phi}$ is estimated in local windows using band-filtered phase according to the pixel values between the interferometric filter phase $\Delta\phi_{filt}$ and the power law scaled heights $(h_o\text{-}h)^{\alpha}$. We set the parameter of power law spatial band to 1000 - 1100 m in Ciloto and 4000 – 12000 m in Ensenada assuming the suitable-fit band filters excluding the deformation signal on those areas. The number of windows for local estimation depends on the size area. We set 50 windows with overlap 50% in Baja California case and 2 windows with overlap 10% in Ciloto case because the small location does not have extremely „varying“ topography at certain lengths.

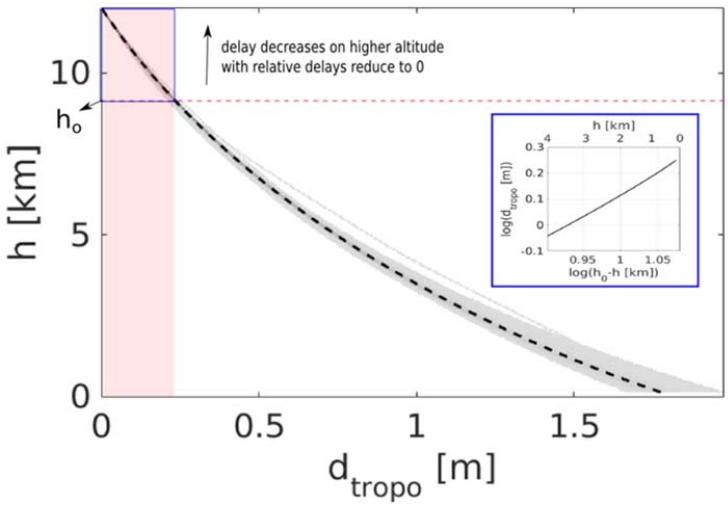


Fig. 7: Tropospheric delay in Northern Baja California, Mexico computed from upper-air sounding data at San Diego Station, 72293 located ~ 130 km away from the observed study area with the blue rectange line described as log-log relationship indicating power law relationship as we assume a linear model. The red colour shows the interferometric delay at this reference height is approximately zero and will be neglected in the tropospheric correction.

3 Result and Discussion

3.1 Corrected interferograms with power law method in Ciloto, Indonesia

SAR images from Sentinel-1A ascending orbit have been processed and generated 269 interferograms. We drop out 19 interferograms for further time series computation since the residuals between small baseline (SB) network and inverted to single master (SM) network are more than 2π in magnitude indicating unwrapping problem (HOOPER et al. 2008). The phase unwrapped continue to be processed on the Toolbox for Reducing Atmospheric InSAR Noise (TRAIN) created by BEKAERT (2015) with power law method fully implemented.

Phase tropo delay based on power law method has been computed on 250 interferograms. The success of applying this method is depending on a relationship between phase and topography assuming APS dominant by stratified delay so that it will be appropriate to power law function. We compare some interferograms which has high and low correlations between phase and elevation. Figure 8a shows an example of interferogram (20150317 – 20150808, ifg 24) where the stratified delay clearly appears on phase unwrapped result. Nevertheless, an interferogram (20141211 – 20150128, ifg 13) does not have a good correlation to topography terrain as shown on Figure 8c. The graphic at the bottom right refers to the power law method not successful for mitigating the phase delay. Moreover, this correction could make the final phase interferogram having another bias than the original interferogram. Overcoming the problem that might be appears in some interferograms, we smooth the corrected phase in time series analysis by low-pass filtering in which we predict the bias as spatially correlated noise due to either turbulent mixing or phase unwrapped error.

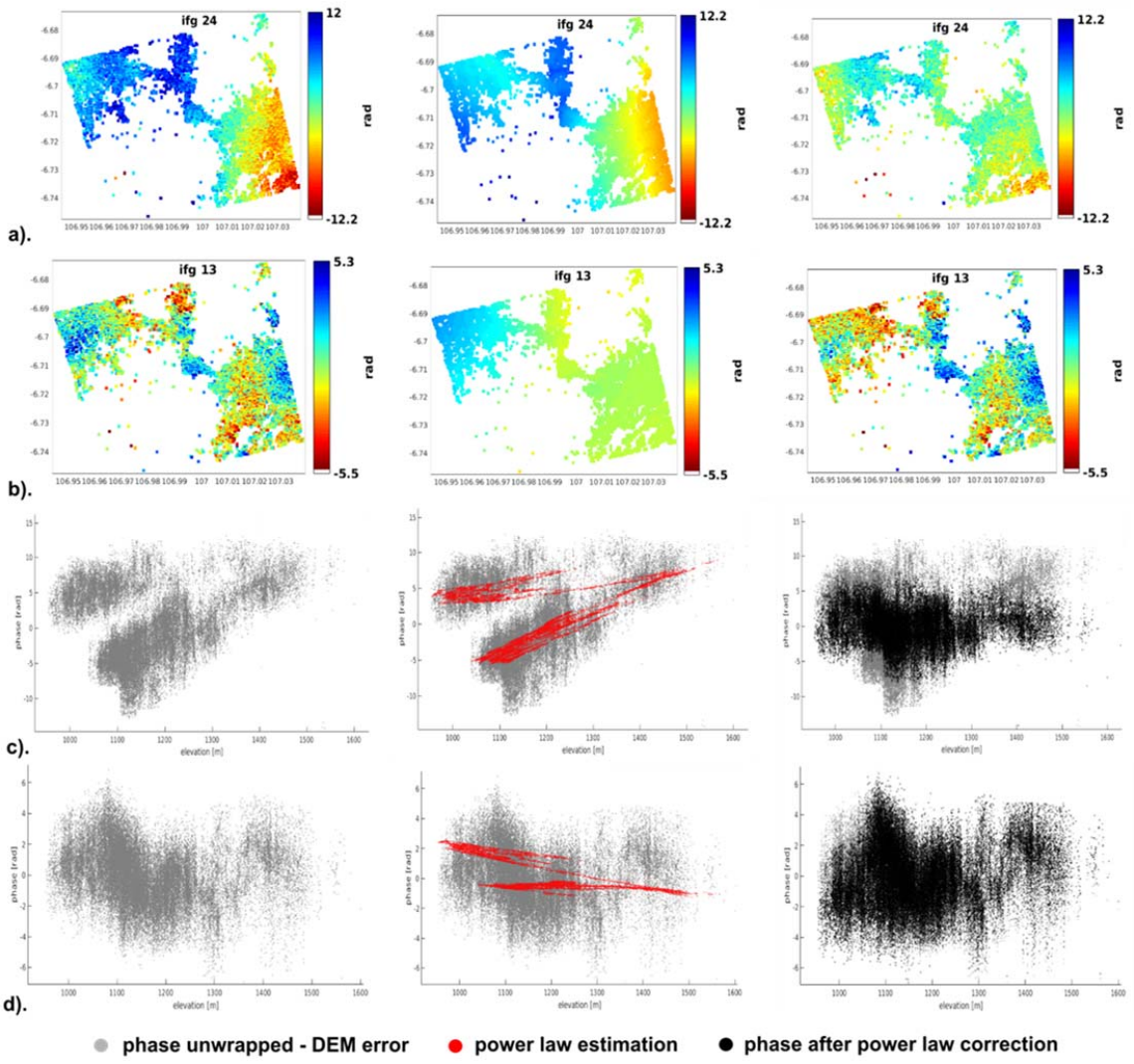


Fig. 8: The left images are original phase unwrapped including DEM error correction, the middles are estimation of power law phase delay and the rights are the corrected phase after power law correction. (a) and (b) refer to sample of success and unsuccessful corrected interferograms with (c) and (d) showing the relationship between phase and topographic height in Ciloto, Indonesia.

3.2 Time series displacement and assessment in Ciloto, Indonesia

250 interferograms with SB network will be inverted using least square to SM network (Schmidt, 2003) in purpose to calculate the time series displacement and mean velocity each PS pixels. The final interferometric phase $\Delta\phi_{depo}$ is PS scatters after substraction from DEM error (δ_{topo}), phase ramp (δ_{orbit}), stratified tropospheric delay ($\delta_{tropo,strat}$) and spatial correlated noise ($\delta_{corr.noise}$), which is described as

$$\Delta\varphi_{defo} = \Delta\varphi - \delta_{topo} - \delta_{orbit} - \delta_{tropo,strat} - \delta_{corr.noise} \quad (7)$$

The mean velocity computed by the general least-squares from October 2014 – September 2017 in Ciloto could be seen on Figure 9. It is calculated with an assumption that the displacement on

every range SAR acquisitions is a linear regression model. Moreover, it accounts only for a depiction of Ciloto's surface change for 3 years aiming to a visualization of PS scatters. It considers that -10 – +10 mm/year slow rate displacement occurs in Ciloto. There are a few PS scatters having high deformation signals in the middle of the case study where the previous geological research (SUMARYONO et al. 2015) described the location as an active landslide zone. It is also near to Puncak highway Km. 81 which in 2015 and 2017 small scale landslides occurred at the area. However, PS results cannot detect clearly the movement's behavior due to limitation of SAR resolution and lack of the exist-able for high correlated pixels.

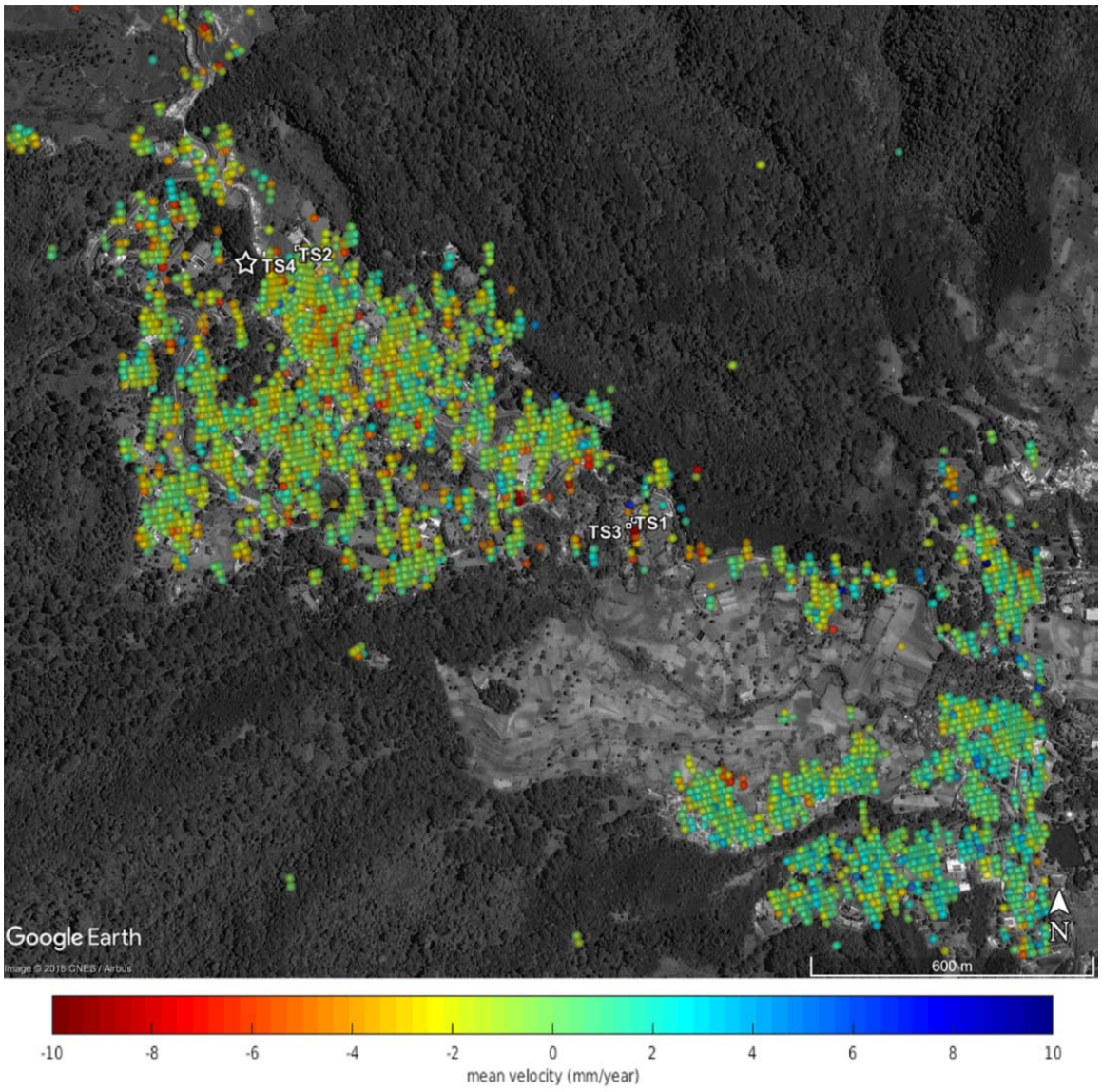


Fig. 9: The mean velocity of slope movement -10 – +10 mm/year in Ciloto landslide prone area. The type of movement is considered as a slow rate displacement with the high signal deformation is found in the middle of study area. TS2, TS4 are a sample of time series displacement near to Extensometer location (the star symbol). TS1, TS3 are located to main body of Landslide Ciloto which recent years a few active landslide occurred.

Viewing on the main body of prone landslide zone, sample points TS1, TS3 have detected ground movement -4.3 mm/year and -1.96 mm/year, respectively. TS1 refers to a sample for regular InSAR processing and TS3 after corrected tropospheric phase delay. Figure 10 shows PS scatters found on radius 25 m from TS1, TS3. There are only a few scatters left on the area since it was difficult to find persistent objects on agricultural area. Simply a linear model to figure out the mean velocity has been conducted using ordinary least square method described as

$$disp_{TS1} = -0.0207t - 3.8072 \pm 0.31$$
$$disp_{TS3} = -0.0223t - 5.7450 \pm 0.39$$

The mean velocity of TS1 and TS2 are -7.5437 ± 2.22 and -8.1615 ± 2.79 mm/year. The model exaggerates the average time series values 0.0087 and 0.0169 /day accumulating to 3.2051 and 6.1959 /year.

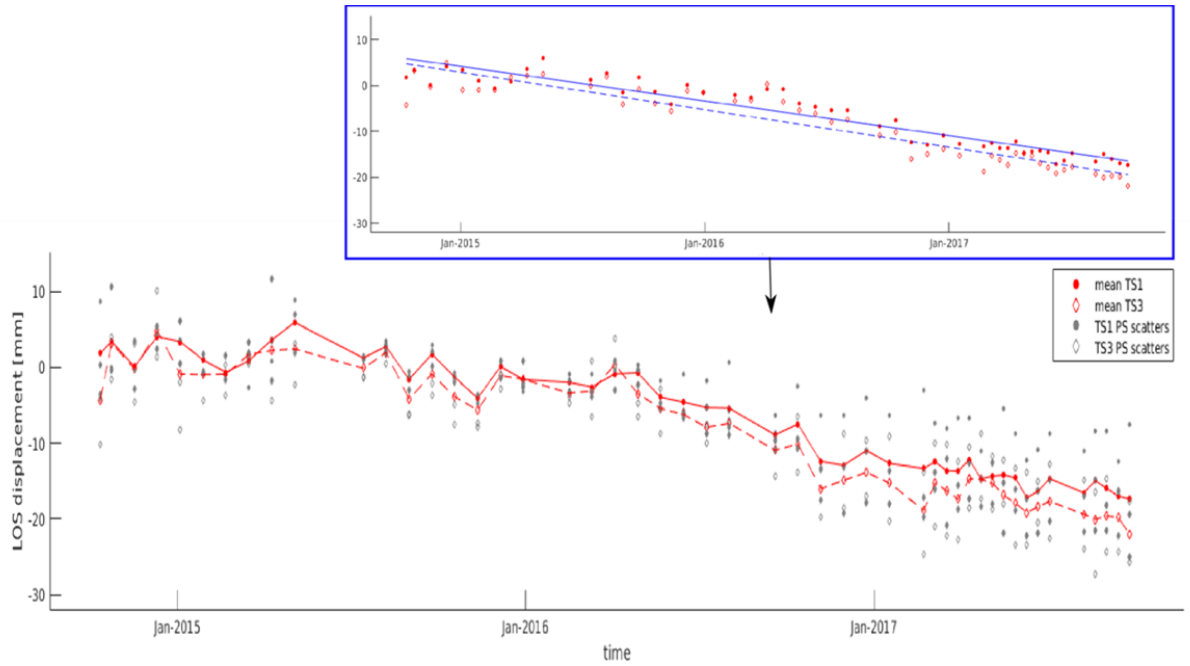


Fig. 10: Time series displacements of TS1, TS3 points (107.0021° E, -6.7119° S) located on Jl. Raya Puncak Km. 81, Ciloto. The blue box indicates the linear model movements both from InSAR results before (solid blue line) and after power law correction (dash blue line).

Furthermore, the result of InSAR processing series need to be compared to ground measurement installed on the study area to quantify before and after corrected tropospheric delay. *Balai Penyelidikan dan Mitigasi Bencana Geologi* (BPPTKG), The Ministry of Energy and Mineral Resources of the Republic of Indonesia has built 3G Track Extensometer, Cipanas Station, located at Kampung Puncak RT06/01 Desa Ciloto, 106°59"41.90" E and 6°42' 26.3" S. The real-time updates of recorded extensometer could be browsed online via BPPTKG's official website (*http://www.merapi.bgl.esdm.go.id/gertan/grafik.php?ids=28*). They installed rain gauge sensor also at the station but unfortunately it was not working thus there are no recorded data from rainfall intensity.

We divide the time of extensometer to be two part based on the reference height of displacement in raw data due to re-installment of the wire. The displacement trend from a first part is corrected to the trend showed on Figure 11a because of an assumption having systematical error + 0.14 cm. Since the verified raw data need to be discussed with BPPTKG directly, we only assess the correlation from second part in which we do not find any systematical error, presuming the wire working properly. TS2, regular InSAR processing, and TS4, corrected tropospheric phase delay InSAR processing, are located at 106.9959° E, -6.707° S. The mean time series displacements for 28 scatters of 50 m around the center could be seen on Figure 11a with ± 50 m the spatial length difference between PS scatters and the extensometer location. Figure 11b indicates a correlation between extensometer and InSAR result before and after troposphere phase delay. It uses linear Pearson's correlation using a Student's *t* distribution for the correlation's transformation to quantify the measurements. The correlation value before tropospheric delay is 0.7674 and after 0.8125. There is a small better accuracy after power law correction, in other words, it increases 0.0451.

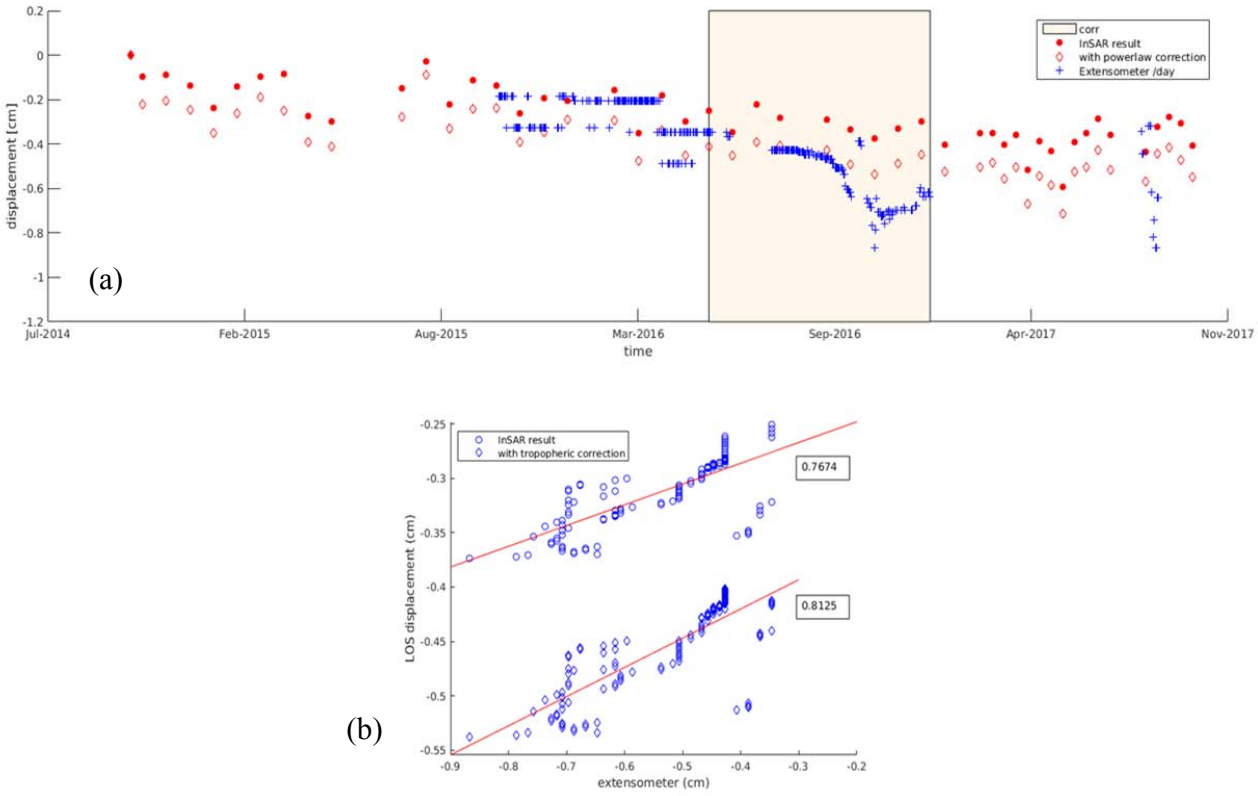


Fig. 11: (a) Time series LOS displacement samples from TS2 point (before power law correction) and TS4 point (after power law correction) compare to extensometer displacement (local geometry). (b) A correlation between extensometer and LOS displacement from 22 May 2016 to 02 Jan 2017 (a corr. rectangle box).

3.3 Corrected interferograms with power law method in Northern Baja California, Mexico

SAR images from Envisat descending orbit has been processed and generated 87 interferograms. As previous case, we correct the tropospheric delay for each interferogram. The terrain in Northern Baja California mostly mountainous area with the flat terrain found near to the coast. The stratified delay is more dominant causing atmospheric phase screen than turbulent mixing due to the varying topographic range and the demonstration of propagating delay from ERA-1 data. An example from an interferogram (18^{th} Nov 2015 – 23^{rd} Dec 2005 with B⊥ -165m) having a high atmospheric artifact could be seen on Figure 12a. The high correlation accounts

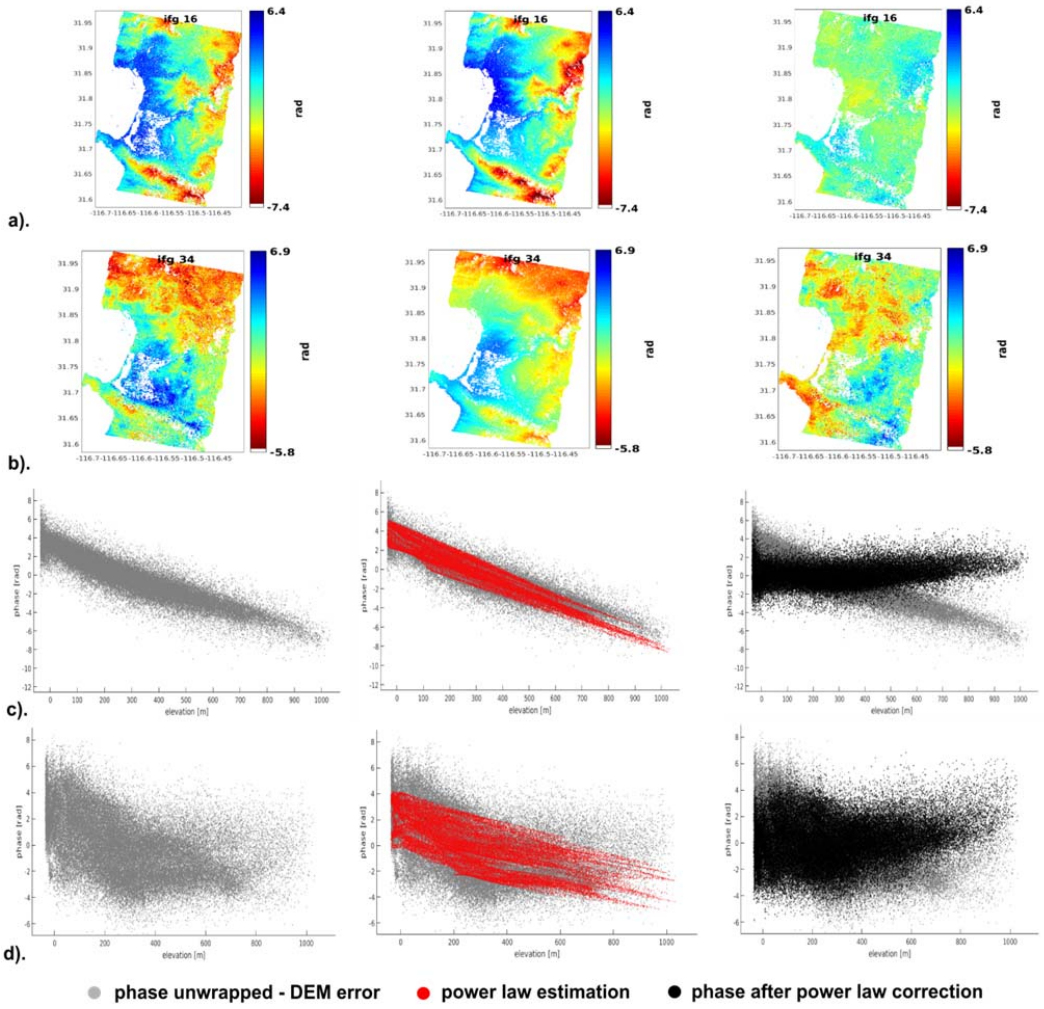


Fig. 12: The left images are original phase unwrapped including DEM error correction, the middles are estimation of power law phase delay and the rights are the corrected phase after power law correction. (a) and (c) refer to sample of success and unsuccessful corrected interferograms with (b) and (d) showing the relationship between phase and topographic height in Northern Baja California, Mexico.

the power law function mitigating the propagation effect properly. Nonetheless, a problem may appear if a region is effected by either subsidence or uplift correlated to topography. Hence, it makes the deformation signal accidentally filter out as well. This is a drawback that we should concern besides paying attention as well for choosing a right band filtering to estimate reliable coefficients $K'_{\Delta\phi}$.

Not all interferograms are properly working to a power law function, some interferograms may have a partial success tropospheric correction or unsuccessful whether phase delay is fully not correlated to topography or the band filter is chosen incorrectly. One solution deciding the right band filter is trying some windows size and compares the result to weather model or non-deformed regions. The bandwith is limited to the spatial extent of the study area for the largest window size and to the resolution of resample size for the smallest one (BEKAERT et al. 2015). We have tried 4 power law spatial bands, 2000 – 8000; 4000 – 10000; 4000 - 12000; 4000 – 14000 m, and selected [4000 12000] m as briefly discussed on section 2.4 with a reason that most estimated tropospheric delays have scale range value similar to original interferograms. Figure 12c refers to a corrected interferogram (19 Oct 2007 – 16 May 2008 with B⊥ 45 m) partially mitigating the propagating delay. Moreover, the left Figure 12c declares clearly that the power law method had an over-estimation phase delay creating a strong artifact in La Bufadora and Punta Banda regions (southwest from the study area).

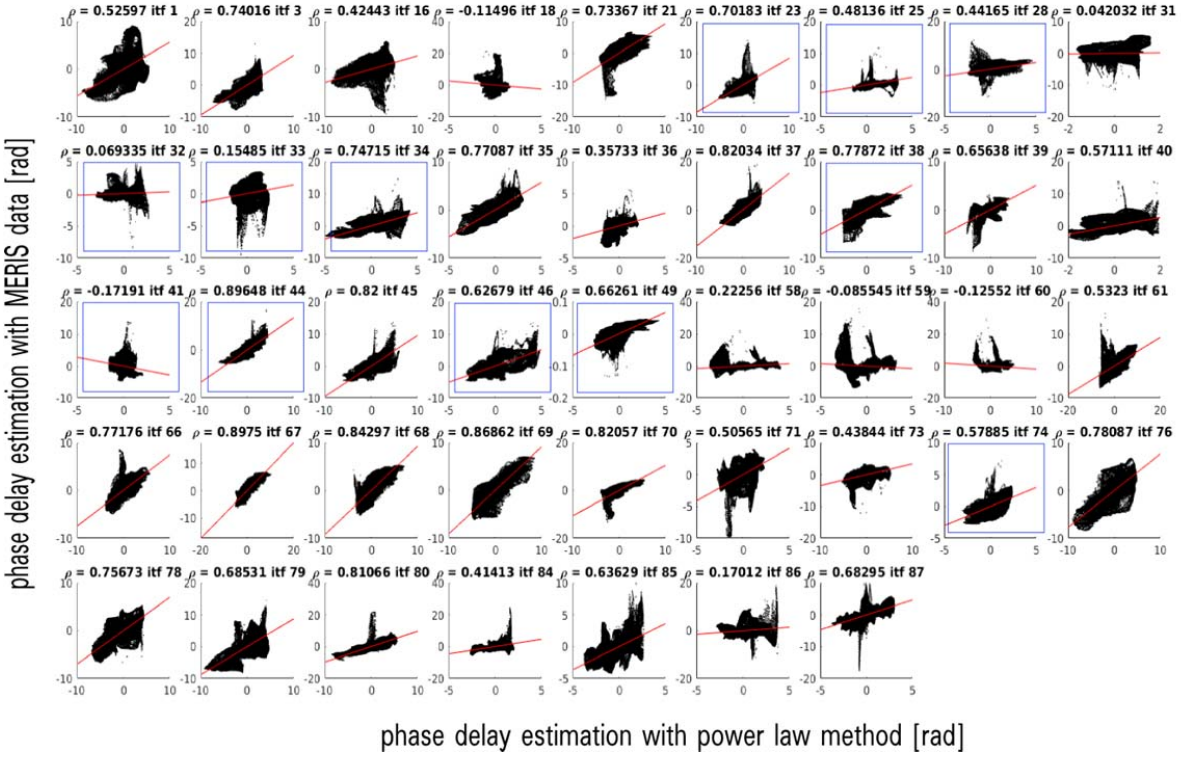


Fig. 13: The correlation of estimated phase delay between phase based - power law and spectrometer - MERIS method with 50% threshold of free cloud coverage while the blue line box refers to 90% threshold of it.

We statistically compared the power law method to MERIS computation. Because ASAR and this spectrometer instrument are operated on the same on-board Envisat Satellite, there is no time delay to estimate an atmospheric delay for each SAR acquisition. However, the drawback to use MERIS data is the daily varying cloud coverage. It makes the success of estimation depending on how much free cloud exist at Mexico region. We have tested 50% and 90% threshold free cloud coverage with the result 29 and 14 out of 44 SAR images, respectively. The mean correlation value between the phase based power law and spectrometer MERIS is 0.52 for 50% threshold and 0.53 for 90% threshold with the fitted number of interferograms showed on Figure 13.

If we consider 15 interferograms from 90% threshold as the appropriate comparison to quantify power law method, 10 out of 15 have a high correlation (mean corr = 0.7) while interferograms 25,28,32,33 and 41 show a low correlation (mean corr = 0.2). In spite of an inaccurate a surface interpolation for the gaps fill of cloud mask, the different correlation values are caused by the type of source delay estimated from spectrometer and phase-based method. The MERIS correction is generated an estimation coming from a wet delay which includes stratify and turbulent components. On the other hand, a phase-based estimation is only from a stratified component but combining hydrostatic and wet delay. A hydrostatic component could be added to MERIS result using the weather model, such as, ERA-1 data. However, the mean correlation after adding the hydrostatic term from ERA-1 does not increase significantly (0.52 for 50% threshold) since a low coarse resolution (~80 km) of ERA1-1 limits the varying hydrostatic delay in spatial. Another reason supposes that some interferograms do not have a high impact from the stratified delay examined by 5 low correlated interferograms. Their correlation to the height terrain encourage the point of view that only interferogram 33 has a high correlation (corr = 0.79) with the other rest numbers 25,28,32,41 showing low correlation -0.44, 0.54, 0.11, -0.58, respectively. Namely, the influence could be from either a turbulent component, dynamic local weather, or unknown noise errors could occur at some acquisition dates.

3.4 Time series displacement and assessment in Northern Baja California, Mexico

The northern part of the Peninsula of Baja California (Mexico), south of Ensenada at the Pacific Ocean has mountain ranges up to 1000 m and arable land in the valleys. The morphology in those areas created some environmental issues either due to concessive natural or human activities. The Agua Blanca fault strikes from northwest to southeast and is a right-lateral strike-slip fault, which was reactivated during the development of the San Andreas Fault. Moreover, the excessive use of groundwater for agricultural, urban, commercial and touristic infrastructure put a big pressure on the aquifers with loss of groundwater supply and quality. Therefore, an investigation needs to be conducted in purpose to support the technical decision policy.

For the first observation, we seek a long-term fault slip along Agua Blanca fault using InSAR time series from Envisat descending data during 8 years. The propagating delay could mask out the sensitivity of InSAR performance for detecting the real signal of deformation, especially for the case of Agua Blanca fault surrounded by mountainous terrain with the most extreme height difference of 1 km. The mean velocity using a general least-squares method and Envisat oscillator correction (MARINKOVIC et al. 2015) has been generated to overview surface's changes

which there are significant artifacts clearly showed by a conventional InSAR SB network on Figure 14a. Although inversion SM (Fig. 14b) network has been done to reduce the noise, it still appears at the southeast observed area. Therefore, on the following paragraphs, we examine the result of corrected tropospheric delay time series (TS) InSAR (Fig. 14c) categorized in three groups. The first group explains a few TS points (TS1; TS2; TS3) along Agua Blanca fault to infer a slip rate

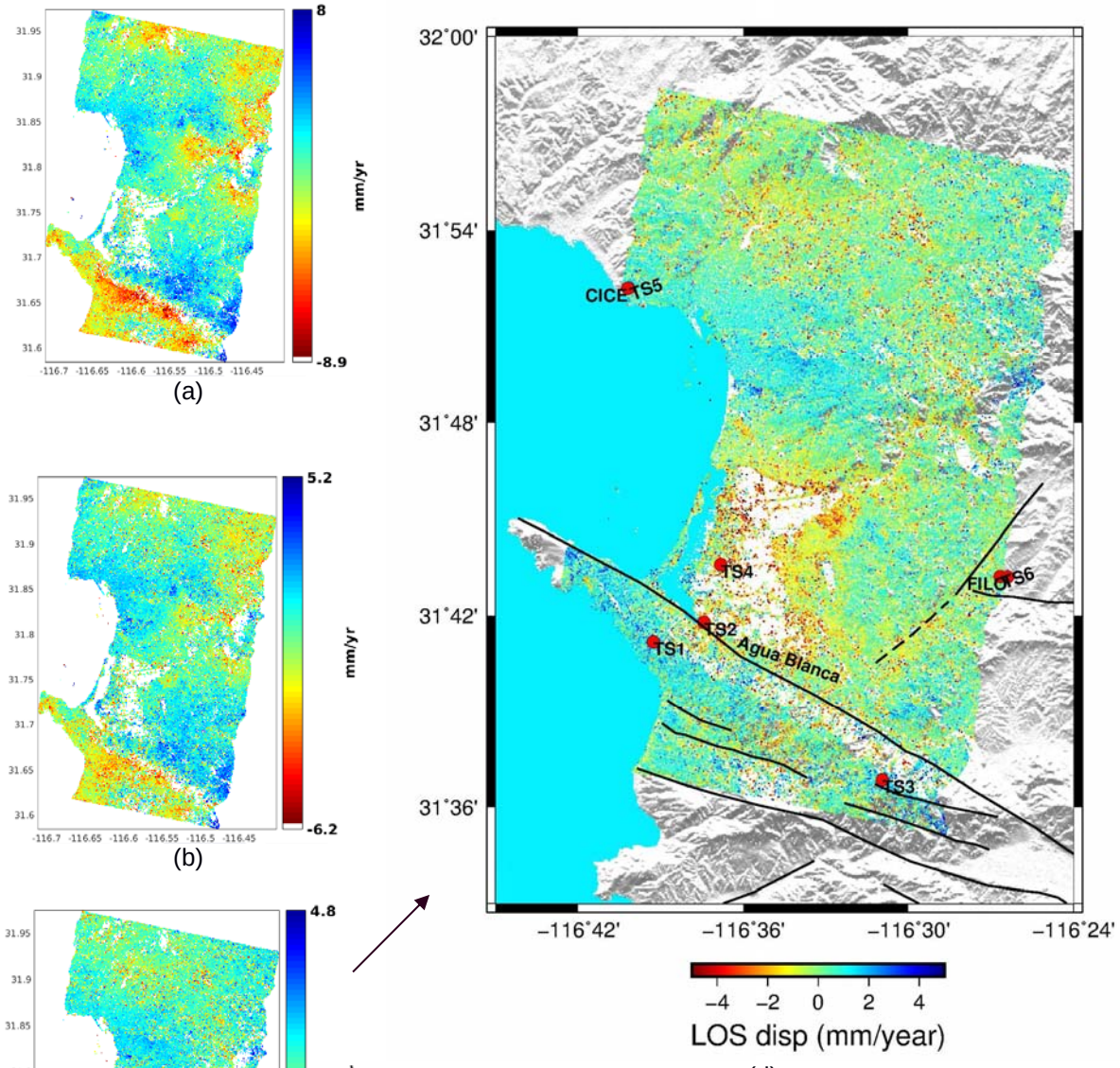


Fig. 14: (a) The mean velocity generated from (a) conventional time series InSAR SB network, (b) with inversion to SM network and (c) after tropospheric delay has been corrected using power law method. (d) The mean velocity for final InSAR results overlaying on SRTM with sample time series points located at Agua Blanca strike-slip fault, Maneadero and GPS locations.

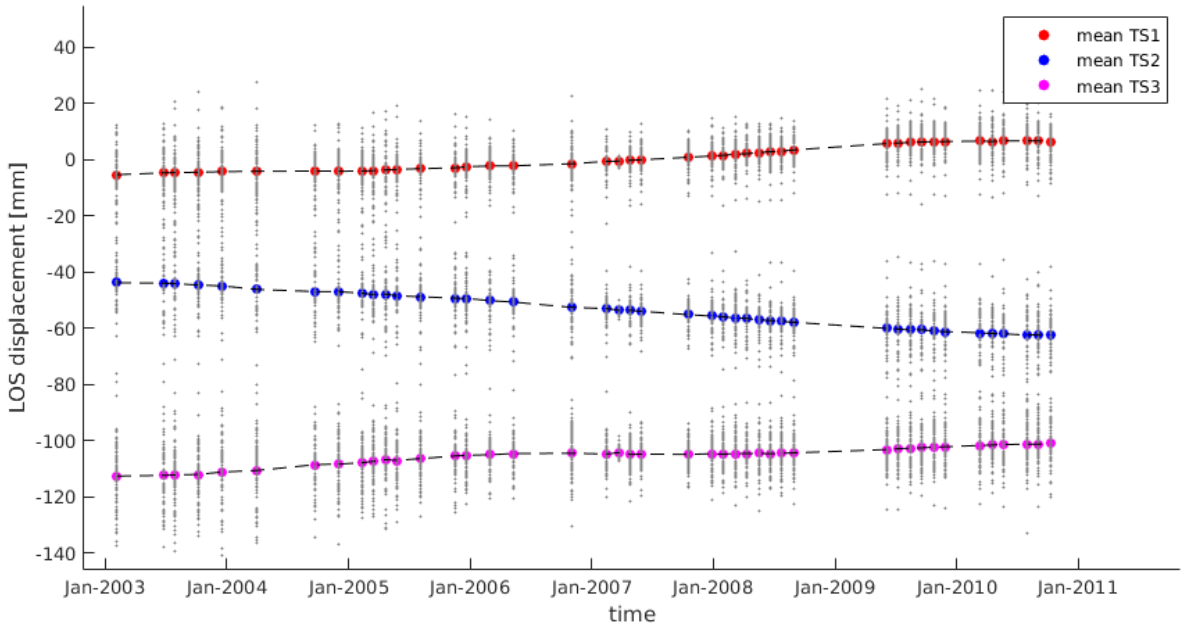


Fig. 15: Time series displacement TS1 (-116.6546°, 31.6862°), TS2 (-116.6239°, 31.6964°), TS3 (-116.5158°, 31.6141°). The average movement is calculated from PS scatters found in radius of 500 m from each point. The Figure has been shifted in y axis aiming for display.

displacement. The second one (TS4) is exploring Manaedero Valley where an environment issue arrised recent years due to over-exploitation water pumping. The last group (TS5; TS6) is a general resemblance between two TS points and GPS points (CICE, FILO) located on the covered area.

According to Figure 15, the mean velocity of TS1, TS2, TS using linear model as written on Section 3.2 are 1.832 ±0.87, -2.688 ±0.43, 1.34 ±1.02 mm/year and considered as a slow rate motion. DIXON et al. 2000 suggested that the fault is active but seismically quiet. There is a similar uplift pattern between TS1 & TS3 that might assume the points placed on „tectonic block“. These results are agree to the previous conducted research from ANIKA et al. 2018 which described the southern part of Aqua Blanca fault laid on the same block and having a similar movement behavior. Although the range TS1 and TS3 is approximately 15 km away agree to a major right lateral strike-slip (ALLAN et al. 1960), we cannot determine whether it is a total offset along the fault because we miss the horizontal components, noise scatters are existing in the middle of Agua Blanca fault and some parts have been masked out due to high decorrelation. On the other hand, TS2 acts differently to TS1, TS3 where the area found depressed or down lift to -2.688 mm/year. Moreover, not only verification to this behavior motion needs to be compared to ground observation but also a geometry relationship between the ground and LOS direction is highly recommended in the future.

We further investigate in general Maneadero Valley where the region has been actively exploited for groundwater to support the agriculture liveliness. Unfortunately, there is not enough PS scatters left on the middle of region due to a decorrelated characteristic for vegetation cover. Near to the coastal dunes, the mean velocity of InSAR reveals that the surface's level around San

Miguel decreases to -2.41 ± 0.76 mm/year using a linear approach of time series displacement for 8 years showed as Figure 16. The graphic presents a down lift movement which is likely occurred because of soil degradation. Wind and water highly contributes to the erosion in fact that ecologically the area belongs to a sand spit landform modified by agriculture and urban areas (MARTINEZ et al. 2007). Since it is salt marshes located close enough to the shoreline, the motion easily gets influence by wind-driven currents, tidal currents, wave action and seasonal weather especially during winter. Regarding to the assumption of sinking land level at a developing salt marsh, there should be an area with the rate of sediment build up (RAFFERTY 2011) and it is seen at southwest from San Miguel (TS4 mark) which PS scatters showing an uplift movement. In addition, the InSAR results are not significantly detecting any subsidence signal in Maneadero City. Hence, we consider the region is more stable sustained by the geological structure of volcanic and sediment rocks than Manaedero Valley with the alluvial material structure.

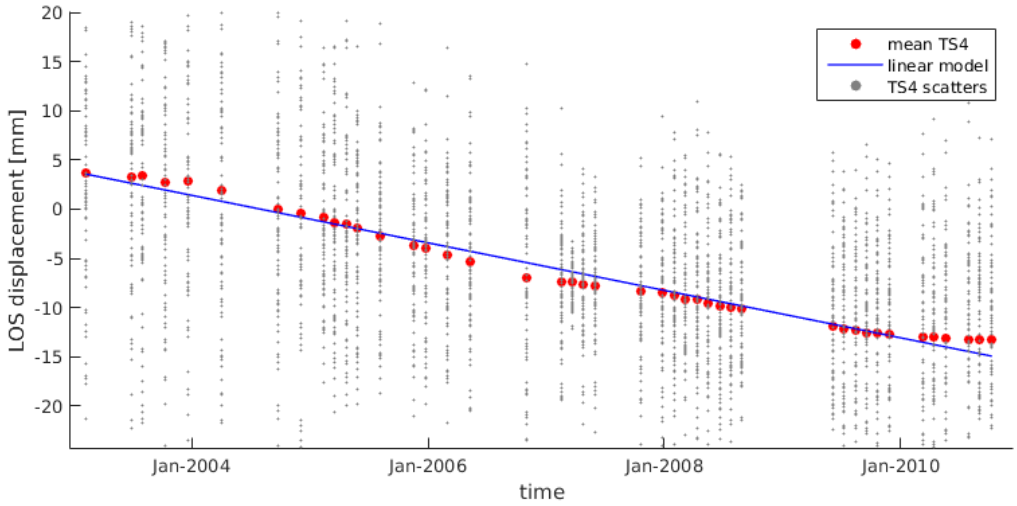


Fig. 16: Time series displacement TS4 (-116.6137°, 31.7264°) located around San Miguel, Baja California. The mean movement is calculated from 81 PS scatters found at radius 500 m from the centre. The level is sinking likely due to soil degradation.

Based on the previous research using GPS observation from 1993 to 1998 (Dixon et al, 2002), the mean velocity is about 5 mm/year across the Agua Blanca and San Miguel Vallecitos faults. Our location only covers the small part of a whole long-term slip rate area across these two faults. Therefore, we only discuss two GPS points, CICE, northwestern direction, and FILO, western direction, at northern Baja California. The sample of TS5 and TS6 represents to GPS locations where 69 and 59 PS scatters found in radius 500 m from the center of TS5 and TS6. We focus only to the vertical components generated from GPS because of unavailability the horizontal component from InSAR results. Since we processed only the descending orbit, the slant range is more sensitive to vertical and west-east than north-south direction. The condition is hardly possible to retrieve the full displacement vector (d_z, d_e, d_n) even if we use both ascending and descending interferograms. The velocities generated both by GPS relative to ITRF-96 from 1993 to 1998 and SBAS InSAR from 2003 to 2011 could be seen on table 2. We projected the LOS displacement to a vertical component using the look angle information on selected PS

pixels. Furthermore, the mean vertical velocity of InSAR results are calculated based on weighted least squares – linear line fit.

Tab. 2: GPS velocities relative to ITRF-96 for CICE and FILO station (Source: DIXON et al. 2002) and InSAR velocities samples for TS5 and TS6 points generated by 44 Envisat descending images.

Point	Latitude	Longitude	Velocity (mm/year)			Duration
			North	East	Vertical	
CICE	31.87°	-116.67°	17.5 ± 0.6	-38.0 ± 0.9	1.5 ± 2.0	1993 - 1998
TS5	31.87°	-116.67°	-	-	0.91 ± 1.8	2003 - 2011
FILO	31.72°	-116.44°	17.3 ± 1.1	-39.5 ± 1.7	1.4 ± 3.2	1993 - 1998
TS6	31.72°	-116.44°	-	-	-0.47 ± 0.6	2003 - 2011

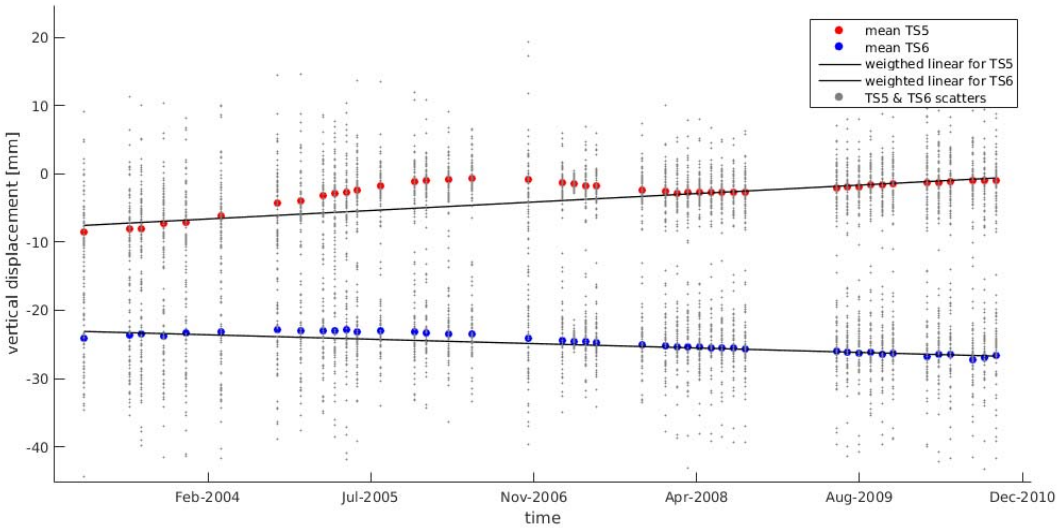


Fig. 17: Time series displacement TS5 (-116.67°, 31.877°) and TS6 (-116.44°, 31.72°) located at CICE and FILO marks from GPS observation. The means are calculated from 69 and 59 PS scatters found at radius 500 m from the centers. The Figure has been shifted in y axis aiming for display.

CICE station has same behavior to TS5 considering that the area was moving up 1.5 ± 2.0 and 0.91 ± 1.8 mm/year. Nevertheless, TS6 acted differently to the movement recorded at FILO station. The FILO records showed uplift to 1.4 ± 3.2 on the contrary TS6 down lift to -0.47 ± 0.6 mm/year. There are a few reason that the vertical components between GPS and InSAR projected results acted differently. In point of geometric reason, GPS is the most insensitive measurement to the vertical component otherwise the slant-range of SAR has the most insensitivity decomposition to north-south component. Hence in fact, the fault belongs to strike-slip type which horizontal displacements are more taking into account than the vertical one. According to the processing strategies, a few concerns are firstly that the band-filter process in

tropospheric delay correction could accidentally remove a deformation signal because of the parallel movement direction to terrain. Secondly, the unsuccessful corrected tropospheric delay and a random noise might still remain. Moreover, FILO station was located at steep terrain which retrieving the side-looking displacement is a challenge due to geometric distortions in the SAR image.

4 Conclusion

We conducted the performance of power law method to mitigate the tropospheric delay over Ciloto, Indonesia and Baja California, Mexico. The correction allows estimating locally varying interferometric phase delay signal in spatial. We tested the method on mountainous terrains at small and large areas. Although the validation to the ground measurement, the extensometer, showed an increased correlation before and after tropospheric delay correction, the performing is way better to be applied at large area in condition to the dominant delay's source from a stratified term. The reason is the spatial scales over these areas change significantly in pressure, temperature and relative humidity. However, the power law method is only successful when the cases have relative delays appropriate to power law function. If it is not, it might introduce incorrect signals because of under or over tropospheric signal estimation subjecting to turbulent variations and local weather. A further tropospheric mitigation needs to use the combination method of phase-based, spectrometer and weather model phenomena with developing an optimization algorithm to select the most accurate strategy being applied to each interferogram.
The quality of slow rate displacement time series InSAR could be improved with stratified tropospheric delay correction. It removes the sampling biases of time series because the stratified delay is systematically presenting seasonal fluctuations and correlated to topography. Especially for the Mexico study area, the hydrostatic delay has to be taken into account when there is more than 10 °C different surface temperature in a year (DOIN et al. 2009). Nevertheless, the displacement's value remains still in slant-range direction which in the next step of our work, we will perform both satellite's orbits for the Ciloto study area and assess the displacement's geometry projected from LOS to the ground surface direction.

5 Acknowledgement

The authors would like to thank European Space Agency (ESA) for providing the Envisat data with Proposal ID 31690 and Copernicus for open-access Sentinel-1 data. This research is funded by Deutsche Akademischer Austauschdienst (DAAD) with Research Grants 91565966. A few images are created by Generic Mapping Tools (GMT).

6 Reference

FUNNING, G.J., PARSONS, B. & WRIGHT, T.J., 2005: Surface Displacements and Source Parameters of the 2003 Bam (Iran) Earthquake from Envisat Advanced Synthetic Aperture Radar Imagery, Journal of Geophysical Research, **110**, B09406.

CHEN, M., TOMAS, R., LI, Z., MOTAGH, M., LI, T., HU, L., GONG, H., LI, X., YU, J. & GONG, X., 2016: Imaging Land Subsidence Induced by Groundwater Extraction in Beijing (China) using Satellite Radar Interferometry, MDPI: Journal of Remote Sensing, **8**, 468.

LIU, P., LI, Q., LI, Z., HOEY, T., LIU, Y. & WANG, C., 2015: Land Subsidence over Oilfields in the Yellow River Delta, MDPI: Journal of Remote Sensing, **7**, 1540-1564.

CASCINI, L., FORNARO, G. & PEDUTO, D., 2010: Advanced Low- and Full-resolution DinSAR Map Generation for Slow-moving Landslide Analysis at Different Scales. Elsevier: Journal of Engineering Geology., **112**, 29-42.

HOOPER, A., SEGALL, P. & ZEBKER, H., 2007: Persistent Scatter Interferometric Synthetic Aperture Radar for Crustal Deformation Analysis with Application to Volcán Alcedo, Galápagos, Journal of Geophysical Research, **112**, B07407.

FERRETTI, A., PRATI, C. & ROCCA. F., 2001: Permanent Scatterers in SAR Interferometry, IEEE Trans. Geosci. Remote Sens., **39** (1), 8-20.

BERARDINO, P., FORNARO, G., LANARI, R. & SANSOSTI, E., 2002: A New Algorithm for Surface Deformation Monitoring Based on Small Baseline Differential SAR Interferograms, IEEE Trans. Geosci. Remote Sens., **40**(11), 2375-2383.

DOIN, M.-P., LASSERRE, C., PELTZER, G., CAVALIE, O. & DOUBRE, C., 2009: Corrections of Stratified Tropospheric Delays in SAR Interferometry: Validation with Global Atmospheric Models, Elsevier: Journal of Applied Geophysics, **69**, 35-50.

BEKAERT, D. P. S., A. HOOPER, & T. J. WRIGHT, 2015: A Spatially Variable Power Law Tropospheric Correction Technique for InSAR Data, J. Geophys. Res. Solid Earth, **120**.

GOLDSTEIN, R., 1995: Atmospheric Limitations to Repeat-track Radar Interferometry, Geophysical Research Letters, **22**(18), 2517-2520.

HANSSEN, R.F., 2001: Radar Interferometry: Data Interpretation and Error Analysis, Dordrecht. Kluwer Academic Publisher, 130-153.

MASSONET, D. & FEIGL, K.L., 1998: Radar Interferometry and Its Application to Changes in the Earth's Surface, Reviews of Geophysics, **36**(4), 441-500.

SMITH, JR. E. K. & WEINTRAUB, S., 1953: The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies, Proceeding of the I.R.E., **41**, 1035-1037.

THAYER, G.D., 1974: An Improved Equation for the Radio Refractive Index of Air, Radio Science, **9**, 803-807.

JUNG, J., KIM D.-J., PARK, S.-E., 2014: Correction of Atmospheric Phase Screen in Time Series InSAR using WRF Model for Monitoring Volcanic Activities, IEEE Trans. Geosci. Remote Sens., **52**(5), 2678-2689.

YU, C., LI, Z. & PENNA, N.T., 2018: Interferometric synthetic aperture radar atmospheric correction using a GPS-based iterative tropospheric decomposition model, Elsevier: Remote Sensing of Environment, **204**, 109-121.

LI, Z., FIELDING, E.J., CROSS, P. & PREUSKER, R., 2009: Advanced InSAR Atmospheric Correction: MERIS/MODIS Combination and Stacked Water Pavour Models, International Journal of Remote Sensing, **30**(13), 3343-3363.

FATTAHI, H. & AMELUNG, F., 2015: InSAR Bias and Uncertainty due to the Systematic and Stochastic Tropospheric Delay, Journal of Geophysical Research: Solid Earth, **120**, 8758-8773.

DIXON, T., DECAIX, J., FARINA, F., FURLONG, K., MALSERVISI, R., BENNETT, R., SUAREZ-VIDAL, F., FLETCHER, J. & LEE, J., 2002: Seismic Cycle and Rheological Effects on Estimation of Present-day Slip Rates for the Agua Blanca and San Miguel-Vallecitos Faults, Northern Baja California, Mexico, Journal of Geophysical Research, **107**(B10), 2226.

RIEDEL, A., RIEDEL, B., SCHÖTMÜLLER, C., NIEMEIER, W. & GERKE, M., 2018: Bestimmung der Erdoberflächenbewegung mittels Radarinterferometrie nahe der Agua Blanca Störung, Baja California. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 27.

SANDWELL, D., MELLORS, R., TONG, X., WEI, M. & WESSEL, P. 2011: Open Radar Interferometry Software for Mapping Surface Deformation, Eos Trans. AGU, **92**(28).

KAMPES, B., HANSSEN, R. & PERSKI, Z., 2003: Radar Interferometry with Public Domain Tools. Proceedings Fringe 2003, **6**.

HOOPER, A., BEKAERT, D., SPAANS, K. & ARIKAN, M., 2012: Recent Advances in SAR Interferometry Time Series Analysis for Measuring Crustal Deformation, Elsevier: Tectonophysics, 514-517, 1-13.

MARINKOVIC, P. & LARSEN, Y., 2015: On Resolving the Local Oscillator Drift Induced Phase Ramps in ASAR and ERS1/2 Interferometric Data - The Final Solution, in: Fringe 2015 workshop (ESA SP-731).

BEKAERT, D.P.S., WALTERS, R.J., WRIGHT, T.J., HOOPER, A.J. & PARKER, D.J., 2015: Statistical Comparison of InSAR Tropospheric Correction Techniques, Elsevier: Remote Sensing of Environment, **170**, 40-47.

HOOPER, A., 2008: A Multi-temporal InSAR Method Incorporating both Persistent Scatterer ans Small Baseline Approaches, Geophysical Research Letters, **35**, L16302.

SUMARYONO., SULAIMAN, C., TRIANA, D.Y., ROBIANA, R. & IRAWAN, W., 2015: Landslide Investigation and Monitoring at Ciloto, West Java, Indonesia, Engineering Geology for Society and Territory, **2**, 1089-1096.

MARTINEZ, M.L. & PSUTY, N.P., 2007: Coastal Dunes: Ecology and Conservation. Ecological Studies, **171**, Springer Berlin Heidelberg, 301-308.

ALLAN, C. R., SILVER, L. & STEHLI, F. G., 1960: Agua Blanca fault, A Major Transverse Structure of Northern Baja California, Mexico, Geol. Soc. Am. Bull., **71**, 457-482.

RAFFERTY, J.P., 2011: The Living Earth: Lakes and Wetlands. New York, N.Y.: Britannica Educational Publishing, Chapter 5, 132-156.

Methoden zur automatisierten Suche von Bodendenkmälern in LiDAR-Daten – neues aus Westfalen

M. FABIAN MEYER[1], INGO PFEFFER[2] & CARSTEN JÜRGENS[1]

Zusammenfassung: Die LiDAR-Technologie hat in den letzten Jahren einige Probleme der archäologischen Prospektion gelöst, doch besteht bei der systematischen und effizienten Auswertung der LiDAR-Daten noch immer Forschungsbedarf.
In einem Forschungsprojekt suchen Geographen und Archäologen aus Bochum und Münster daher seit 2015 nach Möglichkeiten, die großen LiDAR-Datenmengen möglichst effizient auszuwerten. Als Methoden kommen derzeit OBIA und Template Matching zur Anwendung. Am Ende einer Erfassung liegen Verdachtsflächen vor, die der Interpretation bedürfen. Diese wird dadurch erleichtert, dass die Flächen danach sortiert sind, wie sehr sie ihrem jeweiligen Idealtyp entsprechen. Auf diese Weise können die interessantesten Flächen zuerst begutachtet und uninteressante zunächst vernachlässigt werden.

1 Einleitung – archäologische Prospektion

Bodendenkmäler sind laut Denkmalschutzgesetz erhaltenswerte Zeugnisse der Kulturgeschichte. Daher ist es ein grundlegendes Interesse der Archäologie, sie zu schützen, zu erhalten und der Öffentlichkeit zugänglich zu machen.
Die dafür nötige, flächendeckende Erfassung wurde in den letzten 15 Jahren durch die LiDAR-Technologie v.a. dadurch revolutioniert, dass sie das Relief auch in vegetationsreichen Gebieten wie Wäldern sichtbar macht – gerade diese Gebiete sind interessant, da Bodendenkmäler unter der relieferhaltenden Wirkung von Vegetation relativ gut erhalten bleiben und zudem seltener durch landwirtschaftliches Gerät überformt werden. Zumindest oberirdische Bodendenkmäler sind damit gut zu erfassen – also solche, die als Reliefanomalie im Gelände vorliegen.
Um die Sichtbarkeit selbst subtiler Strukturen zu verbessern, werden die Laserdaten in Form verschiedenster Visualisierungen für die manuelle Erfassung mit dem geschulten Auge aufbereitet, wobei der Schummerungsdarstellung eine zentrale Rolle zukommt (Abb. 1). Dieses Verfahren bestätigt das Anfang des Jahrtausends für LiDAR vorausgesagte Potential immer wieder eindrucksvoll, indem es selbst in alten Testgebieten neue Strukturen sichtbar macht.

[1] Geographisches Institut – AG Geomatik, Ruhr-Universität Bochum, Universitätsstraße 104 / Etage 3, D-44799 Bochum, E-Mail [matthias.meyer, carsten.juergens]@rub.de
[2] LWL-Archäologie für Westfalen, Außenstelle Münster, An den Speichern 7, D-48157 Münster, E-Mail ingo.pfeffer@lwl.org

Abb. 1: Landwehr in Dülmen in Luftbild und DGM; Quelle: Bezirksregierung Köln 2016 (Hrsg.): TIM-online, http://www.tim-online.nrw.de

2 Problemstellung – manuelle und automatisierte Suche

Bei allem Erfolg leidet die Schummerungsdarstellung unter derselben Beleuchtungsproblematik wie Luftbilder, die Bodendenkmäler über Schattenwürfe preisgeben: Strukturen, die parallel zur Beleuchtungsrichtung liegen, werfen keine Schatten. Deswegen wurden verschiedenste Visualisierungen entwickelt, die ohne eine simulierte Beleuchtungsquelle auskommen (Abb. 2, eine vergleichende Darstellung von Visualisierungen findet sich z.B. bei HESSE 2016). Am Ende stehen dennoch die manuelle Erfassung und eine Reihe von möglichen Fehlerquellen, angefangen von der Bildschirmdarstellung über das damit verbundene Übersehen von Verdachtsflächen bis

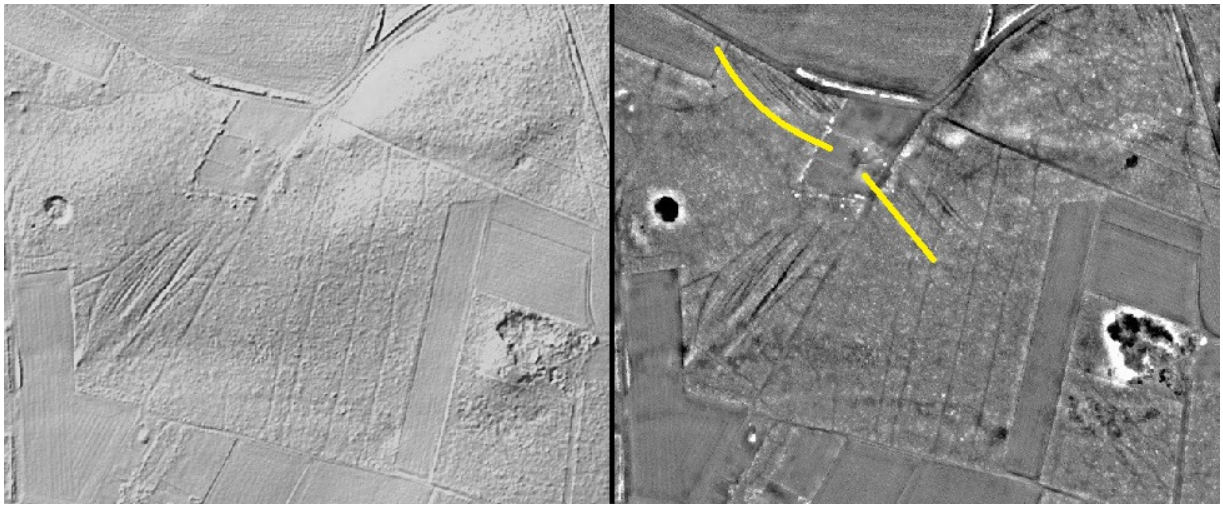


Abb. 2: Vergleich einer herkömmlichen Schummerungsdarstellung (links) und einer *Difference Map* (rechts) nach Hesse: Links fehlen beleuchtungsbedingt die Hohlwege in NW-SO-Richtung. Rechts sind alle Strukturen gut zu sehen (gelb), da die Visualisierung ohne Beleuchtung arbeitet.

hin zum großen Zeitaufwand. Letzteres wird unter Berücksichtigung der großen Zahl zu erwartender Verdachtsflächen einmal mehr deutlich.
Der Wunsch nach Automatisierung besteht schon seit Langem, weswegen seit etwa zehn Jahren verstärkt Forschungsarbeiten zu diesem Thema entstehen. Seit Mitte 2015 sucht auch das Geographische Institut der Ruhr-Universität Bochum in Kooperation mit der LWL-Archäologie für Westfalen in Münster nach Möglichkeiten, die großen LiDAR-Datenmengen möglichst effizient auszuwerten und potentielle Bodendenkmäler weitestgehend automatisiert zu erfassen. Im Zentrum der Entwicklung steht ein Tool für ArcGIS, mit dem bislang Workflows zur Erfassung vier verschiedener Bodendenkmäler entstanden sind.

3 Methodik - Automatische Erfassung von Bodendenkmälern

Im Kontext automatisierter Erfassung bewegt sich die Oberflächenform von Bodendenkmälern zwischen zwei Polen: Auf der einen Seite kann sie so regelmäßig und einfach sein, dass ein Bodendenkmal nicht von natürlichen und anderen Oberflächenformen unterschieden werden kann (z.B. Grabhügel). Auf der anderen Seite kann sie so speziell sein, dass sich das Bodendenkmal nicht mit Automationstechniken erfassen lässt, weil es unmöglich ist, eine allgemeine Form, sozusagen einen ‚kleinsten gemeinsamen Nenner', abzuleiten bzw. diese wiederum so einfach ist, dass sie wieder zu nah am ersten Pol liegt. Letzteres trifft besonders auf Bodendenkmäler zu, die wie z.B. Landwehren aus mehreren einfachen Elementen zusammengesetzt sind: Wird deren gemeinsamer Formenschatz auf die Kombination eines Grabens und eines Walls reduziert, um überhaupt Treffer zu generieren, besteht wiederum eine große Ähnlichkeit zu Entwässerungskanälen auf Feldern.
Für eine automatisierte Erfassung ist daher ideal, wenn ein Bodendenkmal gerade so speziell ist, dass es sich von anderen Formen unterscheidet, sich die verschiedenen Exemplare eines Denkmaltyps untereinander aber möglichst ähnlich und damit gut beschreibbar sind. Ist dieser Sachverhalt nicht gegeben, kann es hilfreich sein, die Exemplare zu gruppieren und getrennt zu klassifizieren. Eine weitere Möglichkeit besteht darin, Bodendenkmäler über die Lage zu unterscheiden und unpassende Lagen im Vorhinein auszuschließen – bspw. ist es unmöglich, dass sich ein Bodendenkmal unter einer versiegelten oder einer Abbaufläche im LiDAR-Datensatz abzeichnet.
Auf die genannten Schwierigkeiten reagieren nicht alle Klassifikationsmethoden in gleicher Weise. Wie sich herausgestellt hat, sind beispielsweise objektbasierte Methoden anfälliger für Störungen als Template Matching-Verfahren (s.u.), da erstere die einzelnen Bestandteile einer Struktur erfassen, während letztere eine Struktur in ihrer Gesamtheit ‚sehen'. Dies ist bei der Wahl der Klassifikationsmethode zu beachten.
Unabhängig von der Klassifikation liegen am Ende der unten beschriebenen Erfassung Verdachtsflächen vor, die der Interpretation bedürfen. Zwar ist diese auf absehbare Zeit nicht automatisierbar (vgl. SEVARA et al. 2016: 496), doch wird sie dadurch erleichtert und beschleunigt, dass die Verdachtsflächen, dort wo es möglich ist, danach sortiert sind, wie sehr sie ihrem jeweiligen Idealtyp entsprechen. Auf diese Weise können die interessantesten Flächen zuerst begutachtet und uninteressante zunächst zurückgestellt werden. Somit ist das Tool als eine Hilfe zur Erfassung zu verstehen, das Archäologen auf Verdachtsflächen aufmerksam macht und diese vorsortiert.

Die Höheninformationen stammen aus dem LiDAR-Befliegungsprogramm des Landes NRW und liegen als Punktdaten in zwei Varianten vor: Entweder als „DGM1“ mit einem gleichmäßigen 1m-Raster oder als „DGM1L“ mit ungleichmäßig verteilen Messpunkten. Die Klassifikationen wurden in Testgebieten in Nordwest- und Ostwestfalen entwickelt.
Zur Kontrolle der Ergebnisse dienen Referenzdaten aus der Fundpunkt-Datenbank der LWL-Archäologie. Treffer, die in diesen nicht vorhanden sind, werden zunächst als „falsche Treffer“ (*false positive*) bezeichnet. Da die Referenzdaten jedoch aufgrund der anfangs dargelegten Schwierigkeiten bei der Auswertung von Geländemodellen möglicherweise unvollständig sind, kommt zusätzlich das ‚geschulte Auge‘ zum Einsatz. So konnten bereits einige neue Bodendenkmäler identifiziert werden.
Der seit 2015 entwickelte Workflow besteht zurzeit aus den in Abb. 3 dargestellten Schritten. Aus den Rohdaten werden zunächst mit einem GIS-Tool verschiedene Geländemodelle berechnet. Diese werden dann je nach Klassifikationsmethode mit eCognition und OBIA (*object-based image analysis*) untersucht oder mit dem gleichen GIS-Tool und Template Matching weiterverarbeitet. Diese Schritte werden im Folgenden näher beleuchtet.

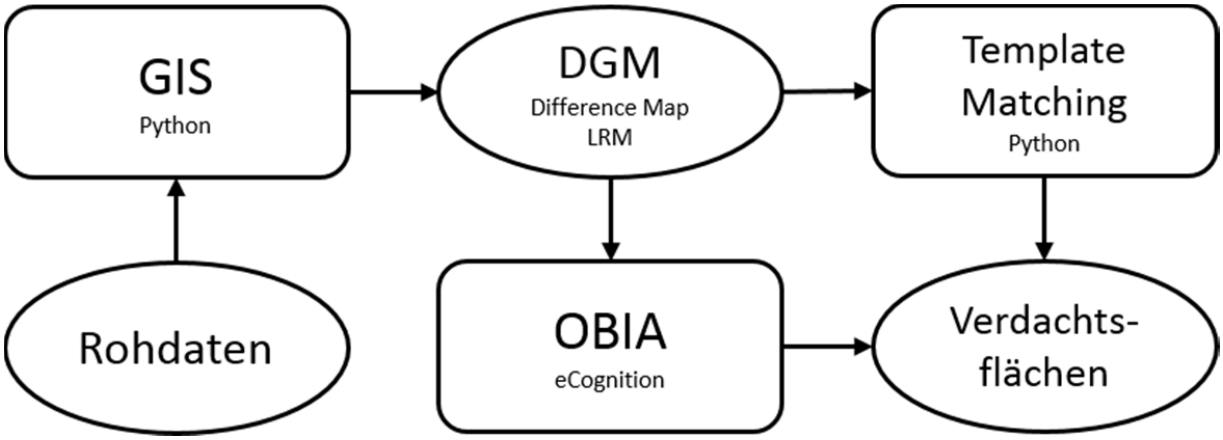


Abb. 3: Gesamtworkflow zur Erfassung verschiedener Bodendenkmäler

3.1 Vorbereitung der Geländemodelle

Die Vorbereitung der Daten geschieht vollautomatisch in einem mit Python programmierten Tool für ArcGIS. Es bietet dem Nutzer alle relevanten Einstellmöglichkeiten und nimmt gleichzeitig die Einstellungen vor, die von den Daten abhängig sind (Abb. 4).
Zunächst benötigt das Tool einen Input-Ordner mit den Datensätzen des gewünschten Untersuchungsgebietes. Dort wird dann eine Geodatabase mit demselben Namen erstellt, die im weiteren

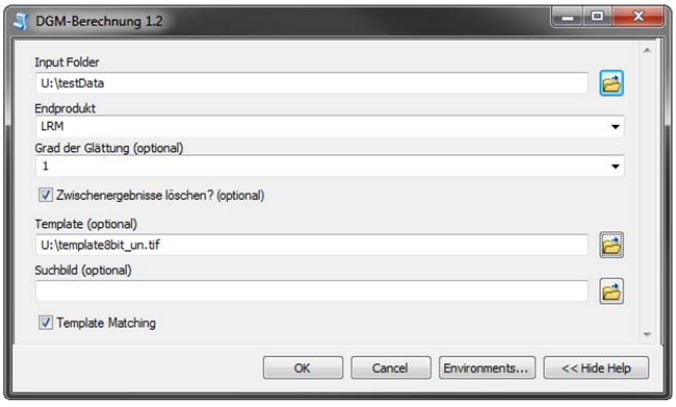


Abb. 4: Benutzeroberfläche des GIS-Tools

Verlauf als Speicherort dient, sodass die Ergebnisse nach den Untersuchungsgebieten sortiert vorliegen. Außerdem wird der Dateipfad auf Leerzeichen überprüft und der Nutzer ggf. darauf hingewiesen, da dies Probleme verursachen kann. Der Name des Ordners wird ebenfalls überprüft – hier werden Leerzeichen jedoch automatisch durch Unterstriche ersetzt.
Als erstes werden die Dateien mit den Koordinatentrios eingelesen und in eine einzige Point Feature Class importiert – das Sammeln der Punkte verhindert bei den späteren DGM-Berechnungen die Entstehung von Kacheln. Dabei wird ebenfalls festgestellt, um welches der beiden möglichen DGM des Landes es sich handelt und den Point Features das passende Koordinatensystem zugewiesen (UTM 6- oder 8-stellig).
Die Aktionen zur Berechnung der DGM richten sich danach, was der Nutzer als Endprodukt wünscht. Der Ablauf bleibt stets der gleiche, doch stoppt das Skript ggf. früher (Abb. 5). Nach dem Import wird ein herkömmliches DGM berechnet, wobei die Berechnungs- bzw. Interpolationsmethode vom Datensatz abhängt. Handelt es sich um den gleichmäßig verteilten Datensatz werden die Point Features zu Rasterzellen konvertiert, was deutlich weniger Rechenaufwand bedeutet als eine IDW-Interpolation der ungleichmäßig verteilten Messpunkte.

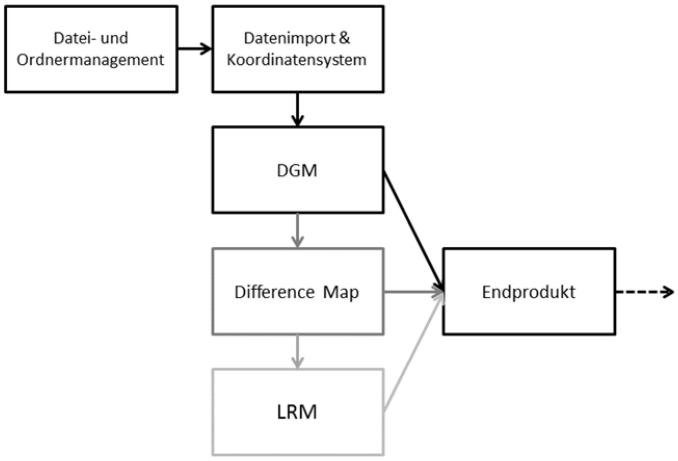


Abb. 5: Teilworkflow des GIS-Tools zur Berechnung der Geländemodelle

Bodendenkmäler lassen sich nur begrenzt in einem einfachen Geländemodell oder einer Schummerungsdarstellung erfassen. Neben dem bereits genannten Nachteil des möglicherweise nicht vorhandenen Schattenwurfs bereiten insbesondere Strukturen in Hanglagen für objektbasierte Klassifikationen Probleme, da die Grenzen der Bodendenkmäler von den Höhenlinien überprägt werden. Daher wird auf die *Difference Map* und das *Local Relief Model* (HESSE 2010) zurückgegriffen, da hier alle Strukturen in der Ebene liegen. Auch das Template Matching profitiert von diesen Visualisierungen, da das Mikrorelief deutlich besser sichtbar ist.

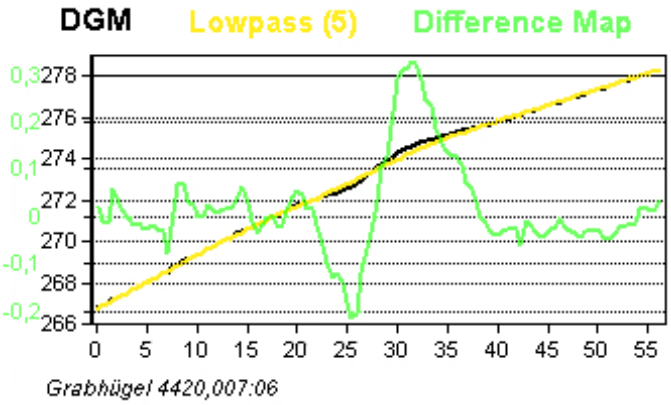


Abb. 6: Profil eines Grabhügels in Hanglage in DGM, 5x geglättetem DGM und Difference Map

Sie entstehen durch ein *Trend Removal*-Verfahren, das kleinräumige Erhebungen und Vertiefungen hervorhebt, indem es das Makrorelief aus dem DGM eliminiert und nur das Mikrorelief übriglässt. Zunächst wird das vorhandene DGM mit einem Tiefpassfilter geglättet, um dieses dann

wieder vom originalen DGM zu subtrahieren, was eine *Difference Map* entstehen lässt. Das LRM geht darüber hinaus, indem es auf Basis der *Difference Map* zusätzlich die Reliefunebenheiten ausstanzt, die Zwischenräume interpoliert und mit den so entstehenden DGM dann erneut subtrahiert. Dies führt zu einem etwas besseren Ergebnis, ist allerdings auch mit deutlich mehr Rechenaufwand verbunden. Unabhängig von der Berechnung wird die Differenz der zwei DGM schließlich visualisiert, wobei die Werte um null mäandrieren und die Amplitude von der Stärke des Tiefpassfilters abhängt (Abb. 6). Zu beachten ist, dass durch das Glätten v.a. an Geländekanten möglicherweise Pseudostrukturen entstehen, die die Klassifikation beeinflussen können (HESSE 2010).

Da es Flächen gibt, in denen keine Bodendenkmäler mehr erhalten sein können, die aber trotzdem gewisse Ähnlichkeiten zu solchen haben, z.B. die Innenflächen erhöhter Kreisverkehre, wird die berechnete Visualisierung mit Daten des digitalen Landschaftsmodells von NRW verschnitten. Aus diesem „Basis-DLM“ wurden einmalig die Flächen extrahiert, in denen Bodendenkmal überhaupt noch vorkommen können und dem Tool beigelegt. Dieser Datensatz entspricht im Wesentlichen den Flächen, die dort als Vegetation gekennzeichnet sind. Auf diese Weise werden Fehlklassifikationen vermieden, die auf der Verwechselung mit anthropogenen Strukturen beruhen (Abb. 7).

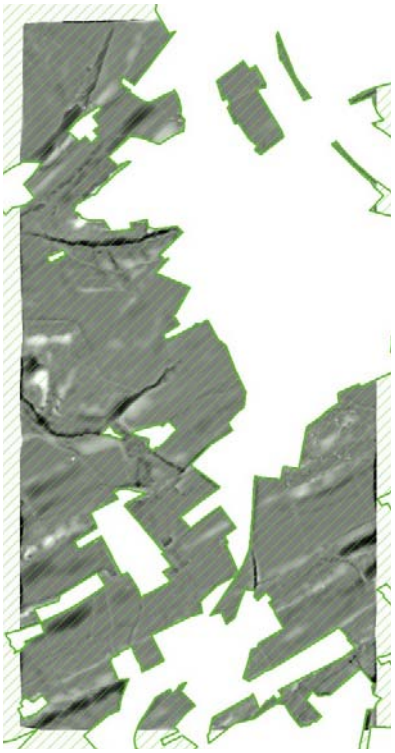


Abb. 7: Local Relief Model, das mittels des Digitalen Landschaftsmodells auf nicht versiegelte Bereiche beschränkt wurde

3.2 Erfassung von Bodendenkmälern mit OBIA

2016 wurde im Rahmen einer Masterarbeit ein Workflow entwickelt, dessen Kern die Software eCognition bildet. Diese sucht mit dem Konzept der objektbasierten Bildanalyse (OBIA, s.u.) in den speziellen Geländemodellen automatisch nach Motten, Grabhügeln und Wölbäckern. Motten, auch Turmhügelburgen genannt, sind mittelalterliche Burganlagen, die i.d.R. aus einem Hügel und einem umlaufenden Graben bestehen, Grabhügel sind Zeugnisse vorgeschichtlicher Bestattungen und beinhalten eines oder mehrere Gräber und Wölbäcker sind historische Ackerfluren, die aus parallel verlaufenden Wällen bestehen (Abb. 8).

Abb. 8: Wölbäcker im Dülmener Wildpark

Bei der objektorientierten Klassifikation werden nicht einzelne Pixel klassifiziert, sondern homogene Bildbereiche (Objekte), die im initialen Schritt der Segmentation erzeugt werden und in diesem Fall den zu suchenden Bodendenkmälern entsprechen. Für die Objekte lassen sich verschiedenste Eigenschaften wie Form, Größe oder der durchschnittliche Pixelwert berechnen. Außerdem stehen die Objekte zueinander in Beziehung, sodass auch Nachbarschaften berücksichtigt werden. Mit denselben Eigenschaften und Nachbarschaftsbeziehungen werden Klassen definiert, die die verschiedenen Bodendenkmäler beschreiben. „Kleine, runde Objekte, die ausschließlich von niedrigeren Objekten umgeben sind“ repräsentieren beispielsweise Grabhügel. „Große, runde Objekte, die von einem Graben umgeben sind“ beschreiben hingegen die Turmhügel von Mottenanlagen und „lange, schmale, runde Objekte, die von höheren Bereichen umgeben sind und neben einem Turmhügelobjekt liegen“ die dazugehörigen Gräben. Passen die Eigenschaften eines Objektes zu denen einer Klasse, wird es ihr zugewiesen.
eCognition bedient sich der berechneten *Difference Map* und greift in Einzelfällen auf das herkömmliche DGM zurück. Der Workflow ist in Form eines *rulesets* gespeichert, das sich in neuen Gebieten mit wenig Aufwand anwenden lässt. Exemplarisch wird im Folgenden der Workflow zur Erfassung von Wölbäckern dargestellt. Es geht dabei nicht darum, jede einzelne Ackerfurche und jede einzelne Kammlinie als solche zu klassifizieren – es reicht vielmehr, einzelne solcher Elemente als Bestandteil einer Flur zu erkennen. Bei dem im Folgenden kurz dargestellten Ansatz werden einzelne Furchen und Kämme klassifiziert, die sich über ihre Nachbarschaft zu weiteren, gleichen Strukturen als Teile von Wölbackerfluren zu erkennen geben. Das Ergebnis besteht daher aus repräsentativen Elementen, die auf Wölbackerfluren schließen lassen.

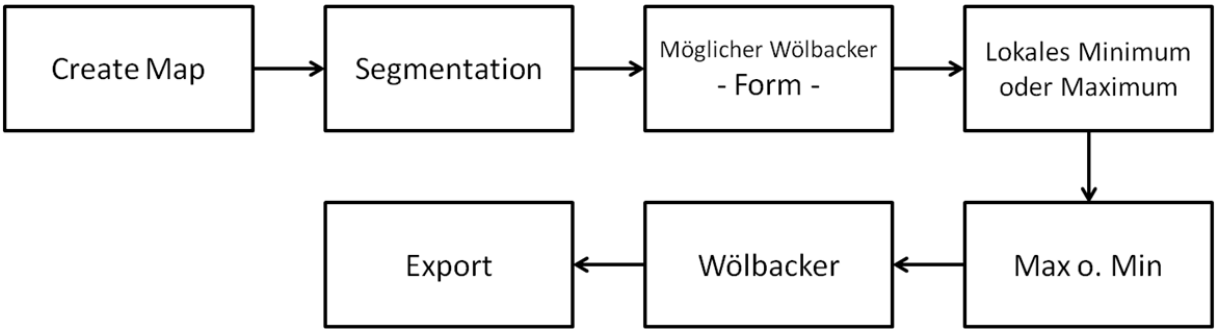


Abb. 9: Workflow in eCognition für die Erfassung von Wölbäckern

Zunächst werden alle größeren Objekte klassifiziert, die eine längliche und schmale Form haben und daher möglicherweise eine Furche oder eine Kammlinie darstellen. Zweitens werden davon alle Objekte klassifiziert, die zu mindestens 80% ein lokales Maximum oder Minimum darstellen. Da dies aus Kontrollgründen gestaffelt geschieht (80, 90, 100), werden die Objekte drittens in einer Klasse gesammelt. Viertens werden dann die Objekte, die von sechs weiteren Objekten derselben Klasse umgeben sind und damit wahrscheinlich Teil einer Wölbackerflur sind, als solche klassifiziert und exportiert (Abb. 9).

3.3 Erfassung von Celtic Fields mit Template Matching

Im Jahr 2016 wurden im Rheinland und in Westfalen sogenannte *celtic fields* entdeckt, die bisher v.a. aus dem nordwesteuropäischen Raum bekannt waren. Diese Strukturen sind teils bereits so stark erodiert, dass sie im Gelände nicht mehr zu erkennen sind. Daher ist deren schnelle flächendeckende Erfassung in Westfalen von großem Interesse.

Abb. 10: Celtic Fields unter einem Waldgebiet in Hopsten

Celtic Fields sind nach ARNOLD (2011: 439) „Ackerfluren, die mittels breiter und flacher Wälle in mehr oder weniger rechteckige oder quadratische Felder aufgeteilt sind“, „deren Kantenlängen stark schwanken, aber bei 40m kulminieren“. Anfänglich wurden sie in England gefunden und in die vorrömische, dort keltische Zeit datiert, doch stellte sich später mit weiteren Funden in den Niederlanden, Dänemark und Nordwestdeutschland heraus, dass sie von der späten Bronzezeit bis in die römische Kaiserzeit datieren. Daher ist der Name zwar irreführend, hat sich inzwischen aber durchgesetzt. Arnold spricht dennoch konsequenterweise von Irregulären Bank-Senken-Feldsystemen (IBSF, Abb. 10).
Grundsätzlich beschreibt Template Matching ein Verfahren zur Suche von Gemeinsamkeiten in Bildern. In archäologischem Kontext wird ein Template, das einer Art ‚mini-DGM‘ eines Bodendenkmals entspricht, in einem DGM gesucht. Dazu wird das Template über das DGM bewegt und an jeder Position (für jedes Pixel) die Übereinstimmung berechnet. Das Ergebnis ist eine *Correlation Map*, deren Pixelwerte den berechneten Übereinstimmungen an der jeweiligen Position entsprechen. Je höher ein Wert, desto höher die Übereinstimmung und damit die Wahrscheinlichkeit, dass es sich tatsächlich um ein Bodendenkmal handelt.
Im *Computer Vision*-Bereich ist die Technik bereits seit langem etabliert und dient der Objekterkennung. Einsatzgebiete sind bspw. die Qualitätskontrolle in Fertigungsprozessen, die Gesichtserkennung oder das Erkennen von Verkehrsschildern für das autonome Fahren.
Bereits 2005 verwendete DE BOER Template Matching zur Suche nach Grabhügeln. Diese Methodik hat sich bis heute weiter verbreitet, wurde nach derzeitigem Kenntnisstand in Verbindung mit LiDAR allerdings nur im Kontext von einfachen, z.B. runden und damit in jeder Richtung symmetrischen Objekten eingesetzt (z.B. SCHNEIDER et al. 2015 oder TRIER et al. 2015). Eine erfreuliche Ausnahme stellt VERDONCK (2016) dar, der Template Matching zur Detektion von rechteckigen römischen Hauswandüberresten in GPR-Daten einsetzte. Nicht zuletzt aufgrund der gewissen Ähnlichkeit hinsichtlich der Form ist es umso interessanter, den Transfer auf Celtic Fields in Westfalen zu wagen.
Bei der Annäherung an eine automatisierte Erfassung von Celtic Fields stellte sich das bisherige Vorgehen mit OBIA als wenig erfolgsversprechend heraus. Das Problem liegt darin, dass die nötigen Objekte in keinem der Tests den Wällen der Celtic Fields entsprachen, sobald in irgend-

einer Weise Störungen, z.B. in Form von Entwässerungsgräben, vorlagen. Gleiches gilt für Waldgebiete, in denen stets eine Tendenz zum Bildrauschen besteht.
Aus den vielen Möglichkeiten wurde Template Matching ausgewählt. Diese Technik spricht Bodendenkmäler in ihrer jeweiligen Gesamtheit an und ist damit resistenter gegenüber Störungen, die in der OBIA die Objekte zuverlässig zerschneiden. Zudem wurde sie bereits im archäologischen Kontext verwendet und ist relativ einfach über Python mit einem GIS kombinierbar, sodass kein weiteres Programm benötigt wird.
Das Template Matching für die Suche nach Celtic Fields wurde in das GIS-Tool implementiert, das bereits die Geländemodelle berechnet. Aktiviert wird es durch die entsprechende Checkbox in der Benutzeroberfläche. Das Tool benutzt dann das vom Nutzer eingegebene Template für den Vergleich mit dem berechneten Geländemodell. Da Celtic Fields in verschiedenen Orientierungen und Größen vorliegen, wird das Template schrittweise gedreht und skaliert. Für jede Kombination wird je eine *Correlation Map* berechnet. Gleiches gilt für eine Größenskalierung. Somit entstehen viele verschiedene *Correlation Maps* aus denen schließlich die Verdachtsflächen herausgefiltert werden.

4 Ergebnisse

Die Ergebnisse zeigen, dass alle vier Bodendenkmaltypen grundsätzlich erfassbar und Trefferquoten von 90-100%, je nach Testgebiet und Klasse, möglich, wenn auch nicht selbstverständlich sind. Im Folgenden werden die Ergebnisse der einzelnen Erfassungen vorgestellt.

4.1 Motten

Bei den Motten werden mit einem Algorithmus erfolgreich verschiedene Erhaltungszustände abgedeckt (Abb. 11). Somit ist es möglich, Burgentypen nach Erhaltungszustand bzw. Form zu klassifizieren und zu bewerten. Genaue Trefferquoten wurden nicht berechnet, da die Anzahl der Motten zu gering ist. Dennoch konnte gezeigt werden, wie die OBIA zusammengesetzte Bodendenkmäler erfassen kann. In den Tests trat ein falscher Treffer auf, der nicht durch anpassen der Klassifikationsparameter vermieden werden konnte und außerdem einer Motte sehr ähnlich sieht.

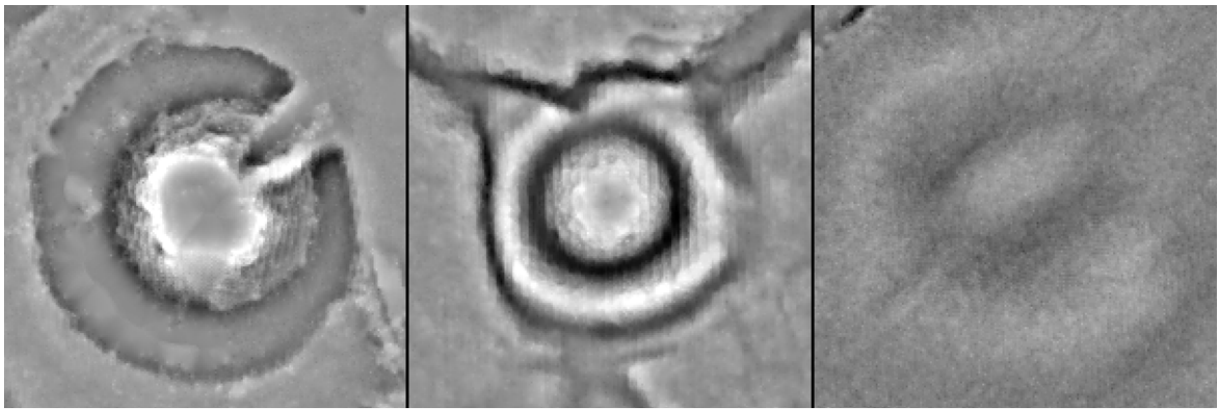


Abb. 11: Verschiedene Erhaltungszustände von Motten, die mit dem Verfahren gefunden werden können

4.2 Grabhügel

Die Grabhügel geben ein sehr diffuses Bild ab. Es zeigte sich, dass hier die Erfassungsgenauigkeit stark von mehreren Faktoren abhängt. Erstens ist dies der Erhaltungszustand – deformierte oder stark erodierte Grabhügel fallen trotz guter Segmentation in weniger idealtypische Klassen. Zweitens sind auch die Segmentation und die Klassifikation mögliche Fehlerquellen, da die Einstellungen zu nicht erkannten und unpassend klassifizierten Verdachtsflächen führen können. Drittens sind Grabhügel sehr einfache Strukturen, die in der Natur und unter anthropogenen Oberflächenformen viele Entsprechungen haben. Verwechselungen sind somit kaum zu vermeiden. Nicht zuletzt ist auch auf die Nachteile der Visualisierungen zu verweisen. Trotz aller Schwierigkeiten wurden bisher unbekannte Grabhügel gefunden und in manchen Gebieten sehr gute Trefferquoten erzielt (Abb. 12). Tabelle 1 zeigt, wie stark sich die angeführten Faktoren auf die Trefferquoten auswirken. Dabei ist zu erwähnen, dass die Zahlen noch ohne das angesprochene Ausstanzen der anthropogen überprägten Flächen zustande kamen, was die Tendenz noch besser verdeutlicht. Im Ibbenbürener Testgebiet ist einerseits ein Steinbruch vorhanden, der in den Visualisierungen viele vermeintliche Kuppen entstehen lässt, andererseits sind die Grabhügel stark erodiert. Im Testgebiet Finnenberg finden sich dagegen weitaus idealere Bedingungen.

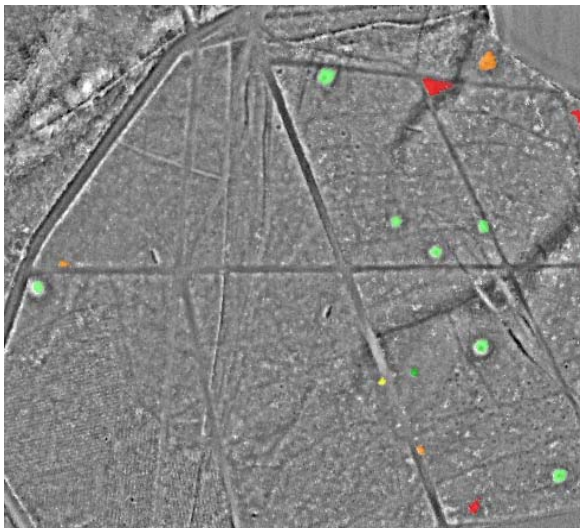


Abb. 12: Ausschnitt eines Testgebietes in Flaesheim nach abgeschlossener Klassifikation. In Grün die idealtypischen Verdachtsflächen, die für eine Interpretation am interessantesten sind. Die Grenzen für die roten, wenig idealtypischen Klassen wurden von realen Grabhügeln abgeleitet. Somit ist nicht ausgeschlossen, dass selbst diese Flächen tatsächlich Grabhügel darstellen

Tab. 1: Klassifikationsergebnisse der Testgebiete Finnenberg und Ibbenbüren im Vergleich

	Finnenberg				Ibbenbüren			
	richtige Treffer	unbestätigte Treffer	gesamt	Trefferquote	richtige Treffer	unbestätigte Treffer	gesamt	Trefferquote
1) ideal	2	0	2	100 %	0	1	1	0 %
2) fast ideal	2	1	3	66,6 %	2	14	16	12,5 %
3) mäßig ideal	2	8	10	20 %	3	21	24	12,5 %
4) wenig ideal	3	24	27	11,1 %	8	97	105	7,6 %
5) nicht ideal	2	59	61	3,2 %	5	201	206	2,4 %
gesamt	11	92	103	10,6 %	18	334	352	5,1 %
nicht erkannt	3 von 14				53 von 71			

4.3 Wölbäcker

Bei den Wölbäckern ist das Ergebnis besser als erwartet. Gerade in flachem Gelände zeigt sich, dass die Klassifikation sehr gut funktioniert (Abb. 13). Es ist dabei fraglich, inwieweit falsche Treffer durch Anpassungen des *rulesets* vermieden werden können, da sie Wölbäckern sehr ähnlich sehen. An Hängen mit hangparallelen Wegen zeigen sich Probleme, da beim Entfernen des

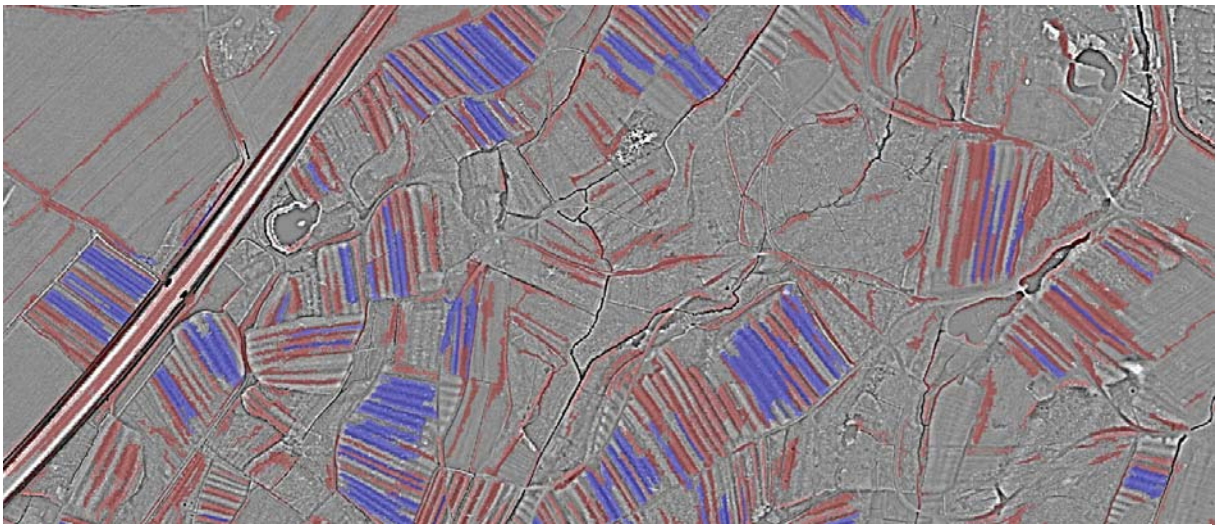


Abb. 13: Klassifikationsergebnis im Wildpark Dülmen. In Blau die erfassten Wölbackerstrukturen. In Rot weitere Strukturen, die zur Klassifikation der blauen Flächen benötigt werden, für die Interpretation aber nicht von großer Bedeutung sind

Makroreliefs Strukturen entstehen, die in *Difference Map* und LRM Wölbäckern sehr ähnlich sehen – die Verwendung des herkömmlichen DGM kommt im Gegensatz zu den Motten allerdings nicht in Frage.

Die Erfassung der Celtic Fields ist methodisch noch nicht abgeschlossen. Die bereits vorliegenden Testergebnisse sehen aber vielversprechend aus (Abb. 14). Der Vergleich mit Abb. 10 zeigt bereits, dass das Template Matching mit Celtic Fields grundsätzlich funktioniert, da in den Bereichen mit Celtic Fields die höchsten Korrelationen auftreten. Auf diesen Ergebnissen gilt es nun aufzubauen.

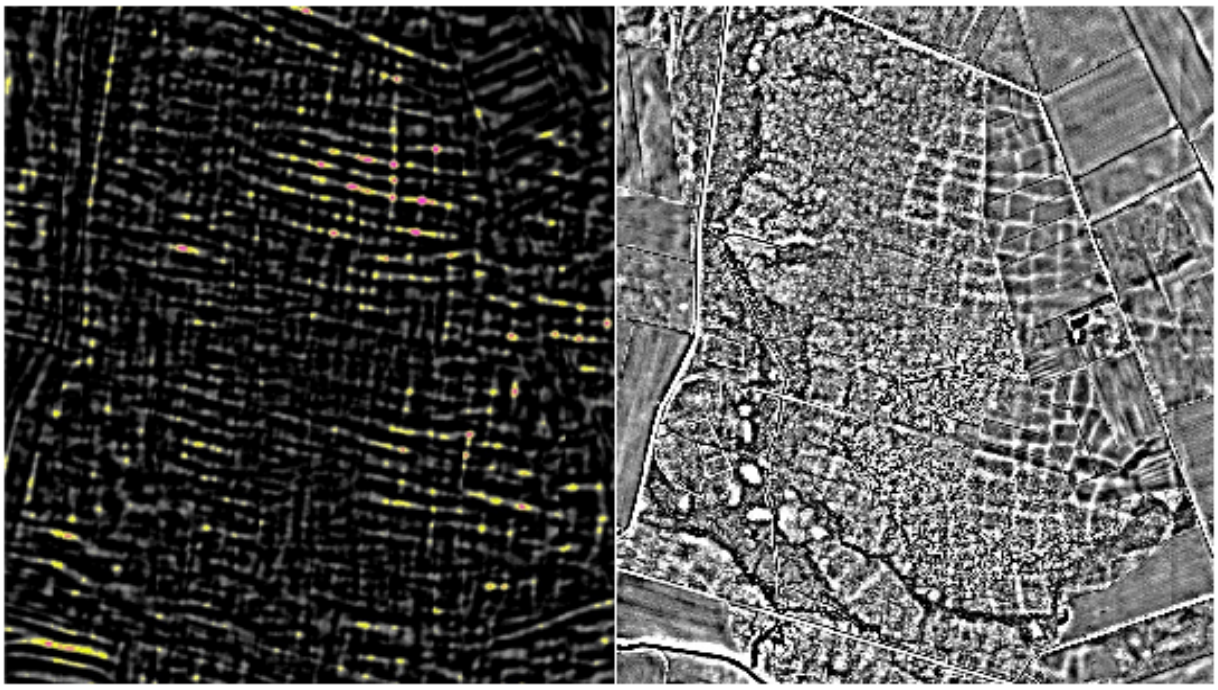


Abb. 14: *Correlation Map* mit eingefärbten Verdachtsflächen (rosa) und weiteren Flächen mit hoher Übereinstimmung zwischen Template und DGM (gelb)

5 Ausblick

Bisher konnte lediglich ein erster Eindruck der Möglichkeiten erlangt werden. Zukünftig könnte daher einerseits die Qualität der Klassifikationen und andererseits der Grad der Automation erhöht werden, was die Methoden für größere Untersuchungsgebiete interessant machte.
Auf Seiten von eCognition könnten die Trefferquoten mit strengeren, feineren und neuen Klassenbeschreibungen verbessert und verschiedene Erhaltungszustände durch mehrere unabhängige Segmentationen abgedeckt werden. Außerdem stellt sich die Frage, ob die Suche nach Motten, Grabhügeln und Wölbäckern nicht auch mittels Template Matching durchgeführt werden kann. Dies würde den Workflow vereinfachen, da alle Funktionen im GIS-Tool gesammelt würden.
Des Weiteren ist die Frage interessant, wie gut sich weitere Daten in die automatisierte Suche integrieren lassen. Interessant sind hier z.B. Satelliten- und Luftbilder im sichtbaren und nicht sichtbaren Bereich, Thermalbilder, Bodenfeuchte-Daten sowie NDVI-Klassifikationen. Möglichweise lassen sich so Bodendenkmäler von natürlichen Strukturen unterscheiden, die im LiDAR-Scan identisch aussehen.
Nicht zuletzt durch oftmals begrenzte finanzielle Mittel ist auch die Verwendung von Open Source Software-GIS wie QGIS interessant. Dies sollte jedoch nicht auf Kosten der Qualität geschehen, weswegen ein Vergleich kommerzieller und kostenloser System interessant ist. Hinsichtlich des Tools müssten lediglich die ArcGIS-Funktionen ersetzt werden.

6 Literaturverzeichnis

ARNOLD, V., 2011: Celtic Fields und andere urgeschichtliche Ackersysteme in historisch alten Waldstandorten Schleswig-Holsteins aus Laserscan-Daten. Archäologisches Korrespondenzblatt, **41**, 439-455.

BOER, A. DE, 2007: Using pattern recognition to search LIDAR data for archeological sites. The world is in your eyes. CAA 2005. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 33rd Conference, Tomar, March 2005, Figueriredo, A. & Velho, G. (Hrsg.), 245-254.

HESSE, R., 2010: LiDAR-derived Local Relief Models – a new tool for archaeological prospection. Archaeological Prospection, **17**, 62-72.

HESSE, R., 2016: Visualisierung hochauflösender Digitaler Geländemodelle mit LiVT. 3D-Anwendungen in Archäologie. Computeranwendungen und quantitative Methoden in der Archöologie – Workshop der AG CAA und des Exzellenzclusters Topoi 2013, Lieberwith, U. & Herzog, I. (Hrsg.), Berlin, 109-128.

SCHNEIDER, A., TAKLA, M., NICOLAY, A., RAAB, A. & RAAB, T., 2015: A Template-matching Approach Combining Morphometric Variables for Automated Mapping of Charcoal Kiln Sites. Archaeological Prospection, **22**, 45-62.

SEVARA, C., PREGESBAUER, M., DONEUS, M., VERHOEVEN, G. & TRINKS, I., 2016: Pixel versus object – A comparison of strategies for the semi-automated mapping of archaeological features using airborne laser scanning data. Journal of Archaeological Science: Reports, **5**, 485-498.

TRIER, Ø, ZORTEA, M. & TONNING, C., 2015: Automatic detection of mound structures in airborne laser scanning data. Journal of Archaeological Science: Reports, **2**, 69-79.

VERDONCK, L., 2016: Detection of Buried Roman Wall Remains in Ground-penetrating Radar Data using Template Matching. Archaeological Prospection, **23**, 257-272.

Klassifikation von Vegetationstypen auf Moorstandorten unter Verwendung von multisensoralen Drohnendaten

FLORIAN BEYER[1] & GÖRRES GRENZDÖRFFER[1]

Zusammenfassung: UAS liefern sehr hochaufgelöste multisensorale Daten, mit deren Hilfe und unter Verwendung moderner Klassifikatoren eine präzise Vegetationsklassifikation eines Moores möglich wird. Der im Rahmen einer UAS-Befliegung aufgenommene multisensorale Datensatz (14 Bänder), einer ca. 8 ha großen Moorfläche, besteht aus RGB-Daten, multispektralen Informationen, Thermaldaten, einem digitalen Oberflächenmodell (DOM) und mehreren Vegetationsindizes. Die Klassifikation mit einem Random Forest Ansatz lieferte bei 11 Klassen eine Gesamtgenauigkeit von ca. 89 %. Die vier relevantesten Variablen für die Klassifikation waren das DOM, die Thermaldaten, der normalisierte Vegetationsindex aus Rot und Red-Edge und der NDVI. Somit ermöglichen UAS auf der Skalenebene zwischen Blatt- und Bestandsebene neue erklärende Daten, die herkömmlichen Satelliten- und Flugzeug-basierten Untersuchungen überlegen sind.

1 Einleitung

Seit Jahrzehnten wird intensiv an der Auswertung von fernerkundlichen Satelliten- und Flugzeug-Daten geforscht. Als Referenzmessungen werden oft punkthafte Messungen am Boden als Kalibration und Validierung der jeweiligen Methode genutzt (WU & LI 2009; MARCEAU & HAY 1999). Dabei entsteht das Problem eines Skalensprungs von einem sehr großen Maßstab auf eine sehr kleine Maßstabsebene, von einer punkthaften Messung auf eine Fläche, also ein Pixel eines Rasterdatensatzes. Man muss in seinen modellhaften Annahmen also davon ausgehen, dass die punkthafte Messung im Gelände die Fläche des Pixels einwandfrei repräsentiert. Dabei entstehen Unsicherheiten und Fehler, die nicht größer werden dürfen, als die eigentliche Zielgröße (SIEGMANN et al. 2015).

Seit einigen Jahren werden Drohnen, unmanned aerial systems (UAS), vermehrt in der fernerkundlichen Forschung eingesetzt. Diese Systeme können als entscheidende Technologie betrachtet werden, um die Skalenübergänge besser zu verstehen und etablierte Ansätze der Satelliten- und Flugzeug-basierten Fernerkundung zu überprüfen.

In der vorliegenden Studie werden verschiedene Arten an Sensoren auf ihre Synergie getestet. Dafür wurden Drohnenbefliegungen mit einer hochaufgelösten RGB-Kamera, mit einem Multispektral-Sensor und einer Thermal-Kamera durchgeführt.

Das Untersuchungsgebiet ist ein wiedervernässtes Durchströmungsmoor zwischen Rostock und Greifswald in Mecklenburg-Vorpommern. Die Moorfläche bei Triebsees befindet sich seit Mitte der 1990er Jahre unter Restaurierungsmaßnahmen und weist mittlerweile eine hohe Biodiversität auf.

[1] Universität Rostock, Agrar- und Umweltwissenschaftliche Fakultät, Professur für Geodäsie und Geoinformatik, Justus-von-Liebig-Weg 6, D-18059 Rostock,
E-Mail: [florian.beyer, goerres.grenzdoeffer]@uni-rostock.de

Ziel der Studie war es zum einen die diverse Moorvegetation möglichst genau zu klassifizieren. Dafür wurde ein multisensoraler Datensatz aus den genannten Sensoren sowie daraus generierten Produkten erstellt und mit einem modernen Algorithmus des maschinellen Lernens (BISHOP 2006) klassifiziert und unabhängig validiert.
Zum anderen sollten die einzelnen Bänder des multisensoralen Datensatzes auf ihren Einfluss auf das Klassifikationsergebnis überprüft werden. Dabei wird untersucht inwiefern die klassischen Fernerkundungsinformationen (Spektralkanäle) aus dem optischen Bereich (400 – 2‘500 nm) den gleichen Stellenwert bei einer Klassifikation von Vegetation zeigen oder ob auf dieser Maßstabsebene eventuell weitere bzw. andere Einflussgrößen eine wesentliche Rolle spielen.

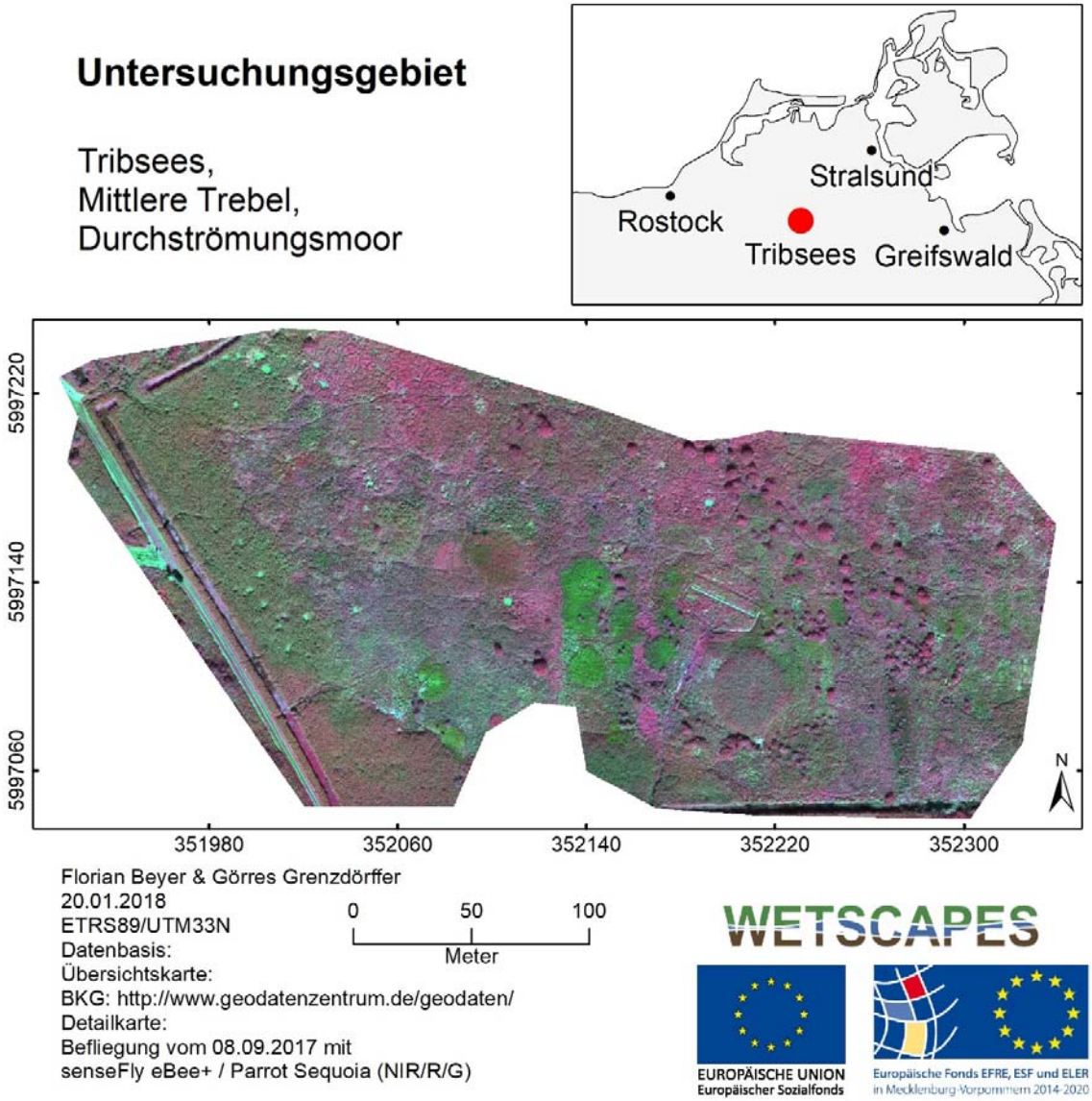


Abb. 1: Untersuchungsgebiet, wiedervernässtes Durchströmungsmoor bei Triebsees.

2 Untersuchungsgebiet

Das Untersuchungsgebiet (Abb. 1) stellt ein Flusstalmoor (WGS84: N 54°06', O 12°45') dar welches sich im Flusstal „Mittlere Trebel" direkt bei Tribsees in Mecklenburg-Vorpommern (Deutschland) befindet. Die klimatischen Jahresmittel liegen bei 685 mm Niederschlag und 8 °C Temperatur. Das Untersuchungsgebiet ist somit in leicht kontinentales Klima mit maritimen Einflüssen einzuordnen. Die Fläche wurde zu Zeiten der Moormeliorationen in der DDR der 1960er & 70er Jahre tief entwässert (ZIEBARTH et al. 2009) um sie in eine agrarische Intensiv-Nutzung zu überführen. Mit Förderung aus EU-Mitteln wird versucht die Fläche wieder in seine ursprüngliche ökologische Funktion als hoch-biodiverses Durchströmungsmoor zurückzuführen. Diese Renaturierungsmaßnahmen laufen seit 1996. Abb. 1 zeigt die etwa 8 ha große Untersuchungsfläche.

Tab. 1: Übersicht über die kartierten Pflanzenarten und der verwendete Klassenname

No.	Klasse	botanischer Name *(deutscher Name)*	Pixelzahl Training	Pixelzahl Validierung
1	Salix	Salix spec. *(Weide)*	5280	3298
2	Boden		170	135
3	Calamagrostis	Calamagrostis canesens *(Sumpfreitgras)*	2962	2301
4	Carex I	Carex acutiformis, Carex riparia *(Sumpf-Segge, Ufer-Segge)*	16650	10360
5	Epilobium	Epilobium hirsutum *(Zottiges Weidenröschen)*	6194	6822
6	Glyceria	Glyceria maxima *(Wasser-Schwaden)*	1691	1552
7	Carex II	Carex rostrata, Carex acuta *(Schnabel-Segge, Schlank-Segge)*	4509	2662
8	Juncus	Juncus effesus, Juncus subnodulosus *(Flatter-Binse, Stumpfblütige Binse)*	2342	1319
9	Phragmites	Phragmites australis *(Schilfrohr)*	41504	33145
10	Typha	Typha latifolia *(Breitblättriger Rohrkolben)*	6218	6536
11	Wasser		375	352
Total			87895	68482

Eine Kartierung des Pflanzenbestandes fand am 24. Oktober 2017 statt. Die Zusammensetzung des Pflanzenbestandes ist in Tab. 1 zusammengefasst. Dabei ist zu beachten, dass es sich fernerkundungs-konform um Dominanzbestände handelt. Die kartierten Pflanzenarten wurden zu 11 verschiedenen Klassen zusammengefasst. Die verschiedenen Dominanzbestände bilden im Gelände eine charakteristische „Patch"-Struktur aus (Abb. 1). Die größten Flächenanteile nehmen

Phragmites- und Carex-Arten ein. Besonders die Klasse Carex II ist von ökologisch besonderem Interesse, weil die Art Carex rostrata unter die seltenen Seggenarten fällt. Epilobium hirsutum ist eher als Einzelpflanze über das ganze Gebiet verteilt und bildet nur vereinzelt größere Dominanzbestände aus.
Die Kartierung wurde als Grundlage verwendet um einen Trainings- sowie einen Validierungsdatensatz zu erstellen. Die Anzahl der jeweiligen Pixel pro Klasse sind in Tab. 1 aufgelistet. Dabei fällt auf, dass manche Klassen unterrepräsentiert sind. Flächen an Wasser und offenem Boden sind im Gebiet sehr begrenzt. Aber auch Pflanzenarten wie Glyceria maxima, die Juncus-Arten oder Calamagrostis canesens gehören zu den Klassen, die auf der Untersuchungsfläche relativ wenig vorkommen.

3 Methoden

3.1 UAS-Bilddaten und Vorprozessierung

Die UAS-Daten wurden mit einer eBee+ der Firma senseFly aufgenommen (ab hier „eBee“). Diese Drohne ist ein Flächenflügler, der etwa 1‘100 g wiegt und mit verschiedenen bildgebenden Sensoren bestückt werden kann. Das RTK-GNSS-Modul ermöglicht zudem eine passpunktlose Aerotriangulation, was insbesondere in schwer zugänglichen Bereichen, wie Mooren von sehr großem Vorteil ist. Der Workflow einer photogrammetrischen eBee-Mission ist weitestgehend automatisiert worden. Aufbauend auf der Flugplanung, einschließlich der Areale für Start und Landung, wird die Drohne im Gelände mit dem Feldrechner verbunden und losgeworfen. Der Bildflug und die Landung erfolgen automatisch. Die Befliegungen fanden am 08. September 2017 statt. Das ca. 8 ha große Untersuchungsgebiet wurde mit jedem Sensor unmittelbar nacheinander dreimal beflogen. Die Flugzeit betrug jeweils ca. 15 Minuten pro Flug. Zunächst erfolgte der Bildflug mit einer hochaufgelösten RGB-Kamera (senseFly S.O.D.A.). Dies ist eine 20 Megapixel-Kamera mit Global Shutter.

Tabelle 2: Spektrale Kanäle der Parrot Sequoia-Sensors

Band	Zentrums-Wellenlänge	Bandbreite
	in Nanometer	
Grün	550	40
Rot	660	40
Red-Edge	735	10
NIR	790	40

Im Anschluss erfolgte der Flug mit einer multispektralen Kamera, der Parrot Sequoia. Diese Kamera besitzt vier einzelne Optiken für die spektralen Bereiche des grünen, roten und zwei Kanäle im infraroten Licht. Einer der Infrarotkanäle ist sensitiv im sogenannten Red Edge, einem Bereich im Übergang vom roten zum infraroten Bereich, der besonders für Vegetationsuntersuchungen von großem Wert ist (MUTANGA & SKIDMORE 2007). Der zweite Kanal liegt klassisch im nahen Infrarot um beispielsweise den NDVI zu berechnen. Jeder der vier Sensoren besitzt

eine Auflösung von 1.2 Megapixel und einen Global Shutter. Die spektralen Kenndaten des Sensors sind in Tabelle 2 aufgelistet.
Eine weitere Besonderheit des Sequoia-Sensors ist ein Sonnenschein-Sensor, der während des Fluges für jede Bildaufnahme auch Einstrahlungsinformationen (Irradiance) in die Bild-Metadaten (EXIF tags) schreibt. So kann während der Vorprozessierung der Daten jedes Bild einer Beleuchtungskorrektur unterzogen und unter Einbeziehung einer kalibrierten Grau-Referenz-Platte echte Reflexionswerte errechnet werden. Im dritten Flug wurde die Thermalkamera (senseFly ThermoMAP, FLIR Tau2) verwendet. Dieser Sensor besitzt eine Auflösung von 0.3 Megapixel (640×512 Pix.). Der Sensor ist in einem Bereich zwischen 7‘000 und 14‘000 nm sensitiv, mit einem Maximum bei 9‘000 bis 11‘000 nm. Seine thermale Auflösung liegt laut Hersteller bei 0.1 °C. Die Kalibration der Temperatur erfolgt automatisch während des Fluges in Form einer Schwarzkalibrierung. Tab. 3 fasst die Kenndaten der drei Befliegungen zusammen.

Tab. 3: Überblick über die UAS-Flüge

Datentyp	Sensor	Anzahl Bilder	Fläche	GSD	Projektions-Fehler
RGB	S.O.D.A	151	8.56 ha	1.54 cm	0.163 Pixel
Multispektral	Sequoia	174	13.57 ha	7.93 cm	0.261 Pixel
Thermal	ThermoMap	1526	12.32 ha	15.14 cm	0.536 Pixel

Während der Befliegung wurden mit einer GNSS-Basisstation RINEX-Daten aufgenommen, die im Postprozessing zur hochgenauen Bestimmung der Projektionszentren genutzt werden. Das Postprozessing der GNSS-Messungen funktionierte mit den RGB- und den Multispektraldaten ohne Fehler. Die Thermaldaten konnten wegen eines Fehlers in der Software auf diese Weise nicht korrigiert werden und die Orthophotos mussten mit Hilfe der AutoSync Workstation in Erdas Imagine 2014 zu den Multispektral-Daten auto-koregistriert werden.
Im Anschluss der Befliegung wurden mit der Software Pix4D die Daten photogrammetrisch prozessiert und Orthomosaike der drei Datensätze berechnet. Außerdem wurde aus den RGB-Daten ein digitales Oberflächenmodell und Graustufenbilder des Greennees Index sowie vier Vegetationsindizes aus den Multispektraldaten berechnet (Tab. 4). Das DOM stellt im nahezu ebenen Moor gleichzeitig ein digitales Objekthöhenmodell dar und gibt somit Auskunft über die Wuchshöhe der zu untersuchenden Pflanzengesellschaften.

Tab. 4: Übersicht über die verwendeten Vegetationsindizes

Index	Akronym	Formel	Referenz
Greennees	GI	*(G)/(R+G+B)*	BROWN et al. 2017
Normalized Difference Vegetation Index	NDVI	*(NIR-RED)/(NIR+RED)*	KROSS et al. 2015
Red Edge NDVI	reNDVI	*(NIR-RE)/(NIR+RE)*	KROSS et al. 2015
Red Edge / Red VI	reNDVI2	*(RE-RED)/(RE+RED)*	SCHUSTER et al. 2012
Green NDVI	gNDVI	*(NIR-GREEN)/(NIR+GREEN)*	KROSS et al. 2015

Wie Tab. 3 zu entnehmen ist, haben die Bilddatensätze unterschiedliche räumliche Auflösungen. Für die Auswertung wurden alle Datensätze auf eine einheitliche Auflösung von GSD = 20 cm umgerechnet (Resampling: Nearest Neighbor).

Nachdem alle Datensätze angepasst waren, wurde ein repräsentativer Ausschnitt des Untersuchungsgebiets ausgesucht und alle Daten an diesen angepasst. Danach erfolgte die Aneinanderreihung aller zur Verfügung stehenden Datensätze (Stacking). In Tab. 5 ist die neue Zusammensetzung des 14-kanaligen multisensoralen Datensatzes aufgelistet.

Tab. 5: Neue Kanalzusammensetzung des multisensoralen Datensatzes

No.	Band	Sensor	No.	Band	Sensor
1	B_1	S.O.D.A.	8	DOM	generiert aus S.O.D.A.
2	G_1	S.O.D.A.	9	Thermal	ThermoMAP
3	R_1	S.O.D.A.	10	GI	generiert aus S.O.D.A.
4	G_2	Sequoia	11	NDVI	generiert aus Sequoia
5	B_2	Sequoia	12	reNDVI	generiert aus Sequoia
6	RE	Sequoia	13	reNDVI2	generiert aus Sequoia
7	NIR	Sequoia	14	gNDVI	generiert aus Sequoia

3.2 Klassifikation und Genauigkeitsbewertung

Die Klassifikation des multisensoralen Datensatzes wurde in Python 2.7.13 durchgeführt. Das Skript dafür steht auf github unter *https://github.com/florianbeyer/RandomForest-Classification* zur Verfügung. Der verwendete Algorithmus ist der Random Forest Classifier (RF, BREIMAN 2001) in einer Python-Bibliothek namens SciKit-learn 0.18.1 (PEDREGOSA 2011). Für den RF wurden 500 Bäume aufgebaut, um das Klassifikationsergebnis möglichst sicher zu bestimmen. Die Robustheit und Genauigkeit des RF gegenüber etablierter Klassifikatoren wie der Support Vector Machine oder dem Maximum Likelihood wurde unter anderem von BEYER et al. (2015) belegt.

Ein weiterer wesentlicher Punkt, warum die Wahl auf diesen Klassifikator fiel ist, dass man sich aus dem trainierten Modell die Relevanz (band importance) der eingeflossenen erklärenden Variablen (die einzelnen Bänder des multisensoralen Datensatzes) ausgeben lassen kann (BEHNAMIAN et al. 2017; YU et al 2017), was in der vorliegenden Studie als wesentliche Information galt. So kann analysiert werden, inwiefern Bänder wie der thermale Datensatz oder das DOM aus der Photogrammetrie einen Beitrag zum Klassifikationsergebnis leisten und wie wichtig sie gegenüber den klassischen Wellenlängen der optischen Fernerkundung sind.

Des Weiteren wird gezeigt, inwiefern der trainierte Algorithmus den Trainingsdatensatz vorhersagt. Da dies jedoch keine Aussagen zur Genauigkeit auf der Fläche zulässt, wurde mit einem unabhängigen Datensatz validiert. Die Gesamtgenauigkeit (OA), der Kappa-Koeffizient (kappa), sowie die Einzelklassengenauigkeiten werden mithilfe der Benutzer- (UA) und der Produzentengenauigkeit (PA) verglichen. Abb. 2 fasst den beschriebenen Workflow nochmal zusammen.

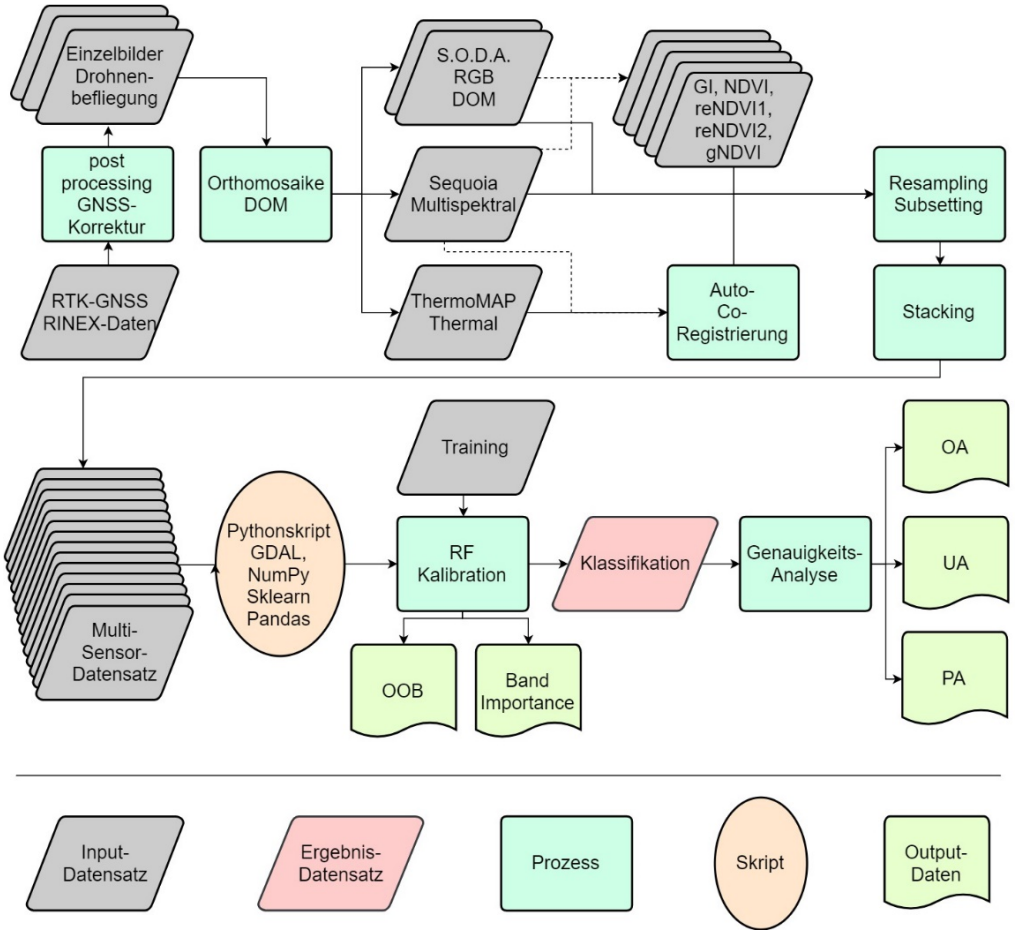


Abb. 2: Workflow der multisensoralen UAS-Datenanalyse

4 Ergebnisse & Interpretation

Nach erfolgreicher Datenvor- und Aufbereitung wurde mit Hilfe des Trainingsdatensatzes der 11 Klassen der RF trainiert. Mit dem trainierten Modell wurde zunächst der Trainingsdatensatz vorhergesagt, um die Kalibrationsgenauigkeit zu überprüfen. Der Tab. 6 kann man entnehmen, dass der Algorithmus zu 100 % (OOB = 99.85 %) an den Trainingsdatensatz angepasst wurde und jedes Pixel zur richtigen Klasse zu geordnet wurde.
Mit diesem Ergebnis wurde das RF-Modell auf das gesamte Untersuchungsgebiet angewendet. Das Ergebnis ist in Abb. 3 dargestellt und die Flächenanteile pro Klasse in Tab. 7. Phragmites australis ist mit 33 % entlang der Entwässerungsgräben im Süden und Westen flächenmäßig die dominante Klasse im Untersuchungsgebiet. Darauf folgen Carex acutiformis und riparia (Carex I: 24.85 %) innerhalb der Fläche. Diese Carex-Arten sind fast überall zu finden. Die selteneren Carex rostrata und acuta (Carex II) sind nur „patch“-artig auf einer kleinen Fläche im Zentrum des Untersuchungsgebietes zu finden (1.26 %).

Tab. 6: Fehlermatrix des vorhergesagten Trainingsdatensatzes

Vorhersage-Genauigkeit	Salix	Boden	Calamagrostis	Carex I	Epilobium	Glyceria	Carex II	Juncus	Phragmites	Typhia	Wasser	Total
Salix	5280	0	0	0	0	0	0	0	0	0	0	5280
Soil	0	170	0	0	0	0	0	0	0	0	0	170
Calamagrostis	0	0	2962	0	0	0	0	0	0	0	0	2962
Carex I	0	0	0	16650	0	0	0	0	0	0	0	16650
Epilobium	0	0	0	0	6194	0	0	0	0	0	0	6194
Glyceria	0	0	0	0	0	1691	0	0	0	0	0	1691
Carex II	0	0	0	0	0	0	4509	0	0	0	0	4509
Juncus	0	0	0	0	0	0	0	2342	0	0	0	2342
Phragmites	0	0	0	0	0	0	0	0	41504	0	0	41504
Typhia	0	0	0	0	0	0	0	0	0	6218	0	6218
Water	0	0	0	0	0	0	0	0	0	0	375	375
Total	5280	170	2962	16650	6194	1691	4509	2342	41504	6218	375	87895

Typha latifolia und Epilobium hirsutum haben beide einen Flächenanteil von etwa 14 %. Dabei kann grob eine Zonierung erkannt werden, wo im Norden eher Epilobium und südlich eher Typha zu finden ist. Dies kann mit der Nähe zum Entwässerungsgraben erklärt werden. BARRAQUETA et al. (1999) bestätigen, dass Epilobium- und verschiedenen Carex-Sorten bei Wassertiefen bis maximal 20 cm auftreten und sich Typha eher in tieferem Wasser (20 – 50 cm) entwickelt.

Offener Boden und Wasser sind die Klassen, die am wenigsten vorkommen. Jedoch muss Wasser als fehlklassifizierte Klasse interpretiert werden. Der Entwässerungsgraben im Westen des Gebietes wurde größtenteils als Typha identifiziert. Hier bedeckt Typha die Wasseroberfläche nahezu komplett und hat das dominierende Signal, was eine Klassifizierung der Wasserwege unmöglich macht. Westlich des Grabens verläuft ein Wirtschaftsweg, der irrtümlich als Glyceria identifiziert wurde. Die Gräser auf dem Wirtschaftsweg wurden in der Klassifikation nicht berücksichtigt und sind Glyceria maxima (Wasser-Schwaden) spektral ähnlicher als die restlichen Klassen.

Um die Gefahr einer Überanpassung (Overfitting) des Modells zu überprüfen, wurde der Validierungsdatensatz für eine unabhängige Überprüfung verwendet. Diese Validierung ergab, dass 60‘917 der 68‘482 Validierungs-Pixel den richtigen Klassen zugeordnet worden. Dies entspricht einer Gesamtgenauigkeit von OA = 88.95 % und einem Kappakoeffizienten von kappa = 0.8457. Dies ist unter der gegebenen Anzahl an Klassen (11) und der Klassifikation zu einem einzigen Aufnahmezeitpunkt ein sehr gutes Ergebnis.

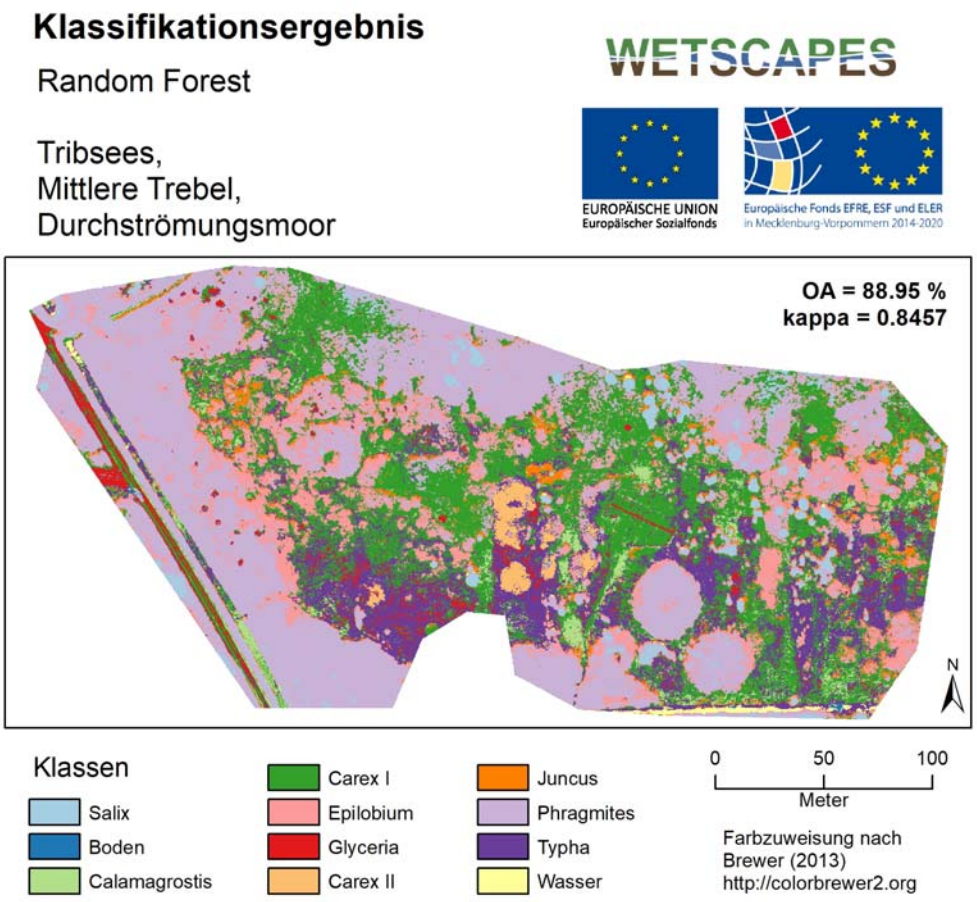


Abb. 3: Klassifikationsergebnis

Tab. 7: Flächenanteile der Klassen im Untersuchungsgebiet

No.	Klasse	Pixel	Prozent	No.	Klasse	Pixel	Prozent
1	Salix	271033	2.72	7	Carex II	125943	1.26
2	Boden	16193	0.16	8	Juncus	345577	3.47
3	Calamagrostis	305766	3.07	9	Phragmites	3295576	33.09
4	Carex I	2474499	24.85	10	Typha	1392772	13.99
5	Epilobium	1419847	14.26	11	Wasser	47049	0.47
6	Glyceria	264333	2.65	Total		9958588	100.00

In Tab. 8 sind die Fehlermatrizen in prozentualen Angaben dargestellt. Falsch klassifizierte Pixel sind in Gelb dargestellt und als Commission-Fehler in der letzten Spalte gezeigt. Das sind Pixel, die irrtümlich einer falschen Klasse zugewiesen wurden. Alle von ihrer Klasse nicht erkannte Pixel sind in der blauen Fläche gekennzeichnet und als Omission-Fehler in der letzten Zeile aufgelistet. Zusätzlich sollen PA und UA nochmal pro Klasse interpretiert werden. Dafür sind in Abb. 4 diese Einzelklassengenauigkeiten direkt gegenübergestellt und die Werte in Tab. 8 ablesbar. Die PA beschreibt wie viele Pixel einer Klasse X tatsächlich richtig klassiert wurden. Die

UA nimmt die Summe aller als Klasse X identifizierten Pixel und setzt diese in Beziehung zu dem Anteil der davon tatsächlich der Klasse X angehört.

Tab. 8: Fehlermatrix der unabhängigen Validierung
(UA = Benutzergenauigkeit, PA = Produzentengenauigkeit)

Klassen in Prozent	Salix	Boden	Calamagrostis	Carex I	Epilobium	Glyceria	Carex II	Juncus	Phragmites	Typha	Wasser	Total	UA	Commission Er.
Unclassified	0	0	0	0	0	0	0	0.2	0	0	0	0		
Salix	79.1	0	0	0	0	0	0	0	0.3	0	0	4.0	95.7	4.3
Boden	0	94.8	0	0	0	0	0	0	0	0	0	0.2	100	0
Calamagrostis	0	0	63.3	1.4	0	0	0	7.4	0	0.8	0	2.6	83.1	16.9
Carex I	0	0	32.86	93	5.8	6.4	0	27.6	0	7.6	0	17.1	82.0	18.0
Epilobium	0	0	0	0.09	80.9	0	0	0.6	4.8	1.1	0	10.5	76.8	23.2
Glyceria	0	0	0	0	0	60.2	0	0	0	3.1	0	1.7	82.4	17.6
Carex II	0	0	0	0	0	0	100	0	0	0.5	0	3.9	98.7	1.3
Juncus	0	0	0.6	4.8	0	0.4	0	39.1	0	0.3	0	1.5	48.7	51.3
Phragmites	20.9	0	0	0	12.4	0	0	0	94.9	0	0	48.2	95.3	4.7
Typha	0	0	3.3	0.7	0.8	32.9	0	25.0	0	86.5	0	9.8	84.5	15.5
Wasser	0	5.2	0	0	0	0	0	0	0	0	100	0.5	98.0	1.9
Total	100	100	100	100	100	100	100	100	100	100	100	100		
PA	79.1	94.8	63.3	93	80.9	60.2	100	39.1	94.9	86.5	100			
Omission Er.	20.9	5.2	36.7	7	19.0	39.7	0	60.9	5.1	13.5	0			

Der Anteil an richtig klassifizierten Pixel ist bei den meisten Klassen sehr hoch (Abb. 4). Dazu gehören die Klassen Boden, Carex I, Carex II, Phragmites, Typha und Wasser mit einer PA und UA über 80 %. Auch Salix ergab gute Ergebnisse, denn die PA liegt bei fast 80 % und der richtige Anteil an allen der Klasse Salix zugeordneten Pixel beträgt 95.7 %. Über 80 % der Pixel von Epilobium wurden richtig zugeordnet, jedoch wurden auch knapp 19 % aus anderen Klassen dieser Klasse zugeordnet. Die sehr kleine Klasse Glyceria kommt immerhin noch auf 60 % richtig klassifizierte Pixel und weist eine hohe UA von über 80 % auf. Die Klasse Juncus ist die problematischste Klasse. Nur 40 % konnten der Klasse Juncus richtig zugeordnet werden und auch nur knapp 50% der als Juncus klassifizierten Pixel sind tatsächlich Juncus.

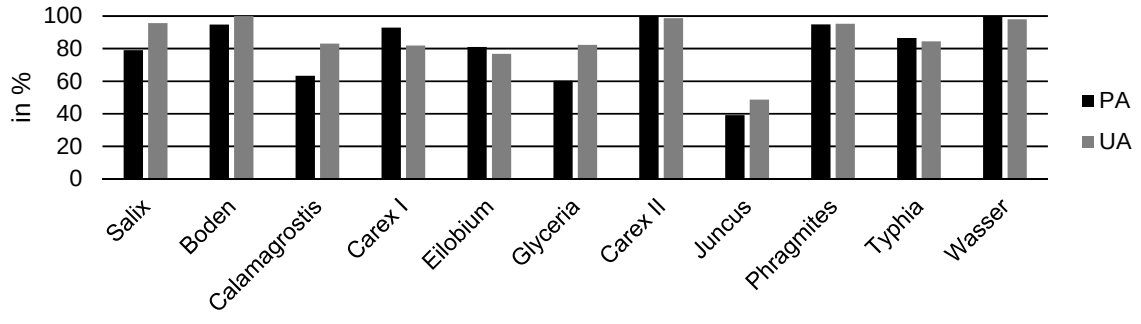


Abb. 4: Produzenten- (PA) und Benutzergenauigkeit (UA) pro Klasse

Die zentrale Frage in der vorliegenden Studie war, inwiefern zusätzliche Daten aus der Photogrammetrie, thermale Daten sowie Vegetationsindizes neben den etablierten spektralen Echtfarb-

und Infrarot-Kanälen einen Beitrag zur Klassifikation liefern können. Dafür wurde ein relativer Wert „Relevanz“ aus dem trainierten RF-Modell für jedes Band des multisensoralen Datensatzes extrahiert. Erstaunlicherweise sind die drei wichtigsten Bänder für den RF das DOM, die Thermaldaten und der reNDVI2, ein Vegetationsindex aus Rot und dem Red-Edge-Kanal (Tab. 9). Erst an vierter Stelle kommt der etablierte NDVI und der erste Farbkanal „NIR“ folgt auf Platz fünf. Dies ist daher erstaunlich, weil bisherige Forschungsarbeiten meistens die reinen Farbkanäle optischer Fernerkundungsdaten als wichtiger als daraus generierte Produkte identifizierten (u.a. LI et al. 2017; SCHUSTER et al 2012). Dies könnte ein Hinweis darauf sein, dass die neuen hochaufgelösten Skalen, die mit Fernerkundungs-Drohnen erreicht werden, im Gegensatz zu Flugzeug und Satelliten-Sensoren, neue Ergebnisse mit etablierten Methoden generieren können. Entgegen bisheriger Möglichkeiten lassen sich aus sich überlappenden Drohnen-Einzelbildaufnahmen immer Oberflächenmodelle generieren, die einen wesentlichen Informationsgehalt (Wuchshöhe) auf der Mikrorelief-Ebene zu haben scheinen. Außerdem bewegen sich die hochaufgelösten Daten im kontinuierlichen Übergangsbereich von der Bestandsebene zur Einzelpflanzenbetrachtung. Daher geben Vegetationsindizes eventuell neue Informationen preis, die auf gröberen Daten vorher rausgemittelt wurden. Unerwartet war auch, dass die bei Tage aufgenommenen thermalen Daten bei Vegetationsanalysen einen so wichtigen Einfluss haben. Dies soll zukünftig bei weiteren Studien überprüft und vertieft untersucht werden.

Tab. 9: Relevanz der einzelnen Bänder als erklärende Variable in der Random Forest Klassifikation

No.	Band	Relevanz [%]	No.	Band	Relevanz [%]
1	DOM	39.84	8	gNDVI	3.53
2	Thermal	11.84	9	R_2	3.20
3	reNDVI2	9.99	10	R_1	2.57
4	NDVI	8.61	11	reNDVI	2.19
5	NIR	5.45	12	B_1	2.03
6	RE	4.37	12	G_1	1.48
7	G_2	3.69	14	GI	1.21
Total					100.00

5 Zusammenfassung & Ausblick

Die vorliegende Studie untersucht zum einen die Genauigkeit der Klassifikation von Moorvegetation mithilfe verschiedener Sensoren eines UAS. Dabei kamen eine hochauflösende Echtfarb-Kamera, eine Multispektrale Frame-Kamera und eine Thermal-Kamera zur Anwendung. Zum anderen sollte die Studie zeigen, welcher dieser Datenarten sowie daraus generierte Produkte den größten Einfluss (Relevanz) auf die Klassifikationsgenauigkeit haben. Elf Klassen wurden mit dem Random Forest-Algorithmus klassifiziert.

Es wurde eine hohe Klassifikationsgenauigkeit von 88.95 % erreicht, welche durch eine unabhängige Validierung festgestellt wurde. Abgesehen von drei flächenmäßig kleinen Gruppen sind die Einzelbetrachtungen der Klassen mit über 80 % Genauigkeit sehr zufriedenstellend.

Die jedoch wichtigste Erkenntnis der Studie erbrachte die Analyse der Relevanz der berücksichtigten Variablen. Die vier wichtigsten Variablen für die Klassifikation waren nach Einfluss geordnet das Digitale Höhenmodell, die Thermaldaten, der normalisierte Vegetationsindex aus Rot

und Red-Edge und der NDVI. Erst an fünfter Stelle kommt der erste Farbkanal (NIR). Dies war so nicht zu erwarten und soll weiter untersucht werden. Auf der neuen Skalenebene, auf der UAS operieren, scheinen andere Parameter wichtig zu sein als bei herkömmlichen Satelliten- und Flugzeug-basierten Untersuchungen.
In Zukunft soll zum einen das Ergebnis auf anderen Moorstandorten verifiziert werden und zum anderen auch andere Aufnahmezeitpunkte während der phänologischen Entwicklung der Moorvegetation Aufschluss darüber geben wann welcher Vegetationstyp durch welche Datenart am besten beschrieben werden kann. Möglicherweise können so auch die schwachen Klassen besser identifiziert werden.
Außerdem war diese Analyse eine Vorstudie für tiefergehende Untersuchungen, welchen Beitrag UAS-Daten auf Moorstandorten bezüglich der Beschreibung des ökologischen Wertes wiedervernässter Moore leisten kann. So sollen beispielsweise die Vegetationsklassifikationen in Zukunft mit Treibhausgas-Emissionen verknüpft werden (COUWENBERG et al. 2011).

6 Danksagung

Besonderer Dank gilt Moritz KAISER (Universität Greifswald) für die am 24. Oktober 2017 durchgeführte Pflanzenkartierung und der anschließenden GIS-Aufbereitung der Daten. Des Weiteren gilt unser Dank Frau Dipl.-Biol. Birgit SCHRÖDER für die hilfreichen Tipps und Ratschläge bezüglich botanischer Fragen. Die vorliegende Studie ist eingebettet in das Verbundprojekt WETSCAPES der Universitäten Rostock und Greifswald, ein Exzellenzforschungsprogramm finanziert durch das Ministerium für Bildung, Wissenschaft und Kultur Mecklenburg-Vorpommern und dem Europäischen Sozial Fond (ESF/14-BM-A55-0034/16).

7 Literatur

BEHNAMIAN, A., MILLARD, K., BANKS, S.N., WHITE, L., RICHARDSON, M. & PASHER, J., 2017: A Systematic Approach for Variable Selection With Random Forests: Achieving Stable Variable Importance Values. IEEE Geoscience and Remote Sensing Letters, **14**(11), 1988-1992. DOI: 10.1109/LGRS.2017.2745049.

BEYER, F., JARMER, T. & SIEGMANN, B., 2015: Identification of Agricultural Crop Types in Northern Israel using Multitemporal RapidEye Data. PFG – Journal for Photogrammetry, Remote Sensing and Geoinformation, **2015** (1), 21-32. DOI: 10.1127/pfg/2015/0249.

BISHOP, C., 2006: Pattern Recognition and Machine Learning. Springer-Verlag New York.

BREIMAN, L., 2001: Random Forests. Machine Learning, **45**(1), 5-32. DOI: 10.1023/A:1010933404324.

BROWN, L.A., DASHA, J., OGUTUA, B.O. & RICHARDSON, A.D., 2017: On the relationship between continuous measures of canopy greenness derived using near-surface remote sensing and satellite-derived vegetation products. Agricultural and Forest Meteorology, **147**, 280-292. DOI: 10.1016/j.agrformet.2017.08.012.

BARRAQUETA, P., HERAS, P. & INFANTE, M., 1999: Saldropo: Vom Moor zum Feuchtgebiet. Bodenoekologie interdiziplinaer, Koehler, H., Mathes, K., Breckiling, B. (Hrsg.), Springer-Verlag Berlin Heidelberg GmbH, 189-198. DOI: 10.1007/978-3-642-58495-4.

Couwenberg, J., Thiele, A., Tanneberger, F., Augustin, J., Bärisch, S., Dubovik, D., Liashchynskaya, N., Michaelis, D., Minke, M., Skuratovich, A. & Joosten, H., 2011: Assessing greenhouse gas emissions from peatlands using vegatation as a proxy. Hydrobiologia, **674**, 67–89. DOI: 10.1007/s10750-011-0729-x.

KROSS, A., MCNAIRN, H., LAPEN, D., SUNOHARA, M. & CHAMPAGNE, C., 2015: Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. International Journal of Applied Earth Observation and Geoinformation, **34**, 235–248. DOI: 10.1016/j.jag.2014.08.002.

LI, X., CHEN, W., CHENG, X., LIAO, Y. & CHEN, G., 2017: Comparison and integration of feature reduction methods for land cover classification with RapidEye imagery. Multimedia Tools and Applications, 76(21), 23041-23057. DOI: 10.1007/s11042-016-4311-4.

MARCEAU, D.J. & HAY, G.J., 1999: Remote Sensing Contributions to the Scale Issue. Canadian Journal of Remote Sensing, **25**(4), 357-366. DOI: 10.1080/07038992.1999.10874735.

Mutanga, O. & Skidmore, A.K., 2007: Red edge shift and biochemical content in grass canopies. ISPRS Journal of Photogrammetry and Remote Sensing, **62**(1), 34-42. DOI: 10.1016/j.isprsjprs.2007.02.001.

PEDREGOSA, G., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, B., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M. & DUCHESNAY, E., 2011: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, **12**, 2825-2830.

SCHUSTER, C., FÖRSTER, M. & KLEINSCHMIT, B., 2012: Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. International Journal of Remote Sensing, **33**(17), 5583–5599. DOI: 10.1080/01431161.2012.666812.

SIEGMANN, B., JARMER, T., BEYER, F. & EHLERS, M., 2015: The Potential of Pan-Sharpened EnMAP Data for the Assessment of Wheat LAI. Remore Sensing, **7**(10), 12737-12762. DOI: 10.3390/rs71012737.

WU, H. & LI, Z.L., 2009: Scale Issues in Remote Sensing: A Review on Analysis, Processing and Modeling. Sensors, **9**(3), 1768-1793. DOI: 10.3390/s90301768.

Yu, B., Chen, F. & Chen, H., 2017: NPP estimation using random forest and impact feature variable importance analysis. Journal of Spatial Science, 1-20. DOI: 10.1080/14498596 .2017.1367331.

ZIEBARTH, M., LENSCHOW, U. & PERMIEN, T., 2009: Konzept zum Schutz und zur Nutzung der Moore - Fortschreibung des Konzeptes zur Bestandssicherung und zur Entwicklung der Moore. Ministerium für Landwirtschaft, Umwelt und Verbraucherschutz Mecklenburg-Vorpommern.

Bilanzierung der Landnutzung 2007 und 2017 – Ist ein Vergleich möglich?

DETLEF GÜNTHER-DIRINGER [1] & DAVID BRAUN [2]

Zusammenfassung: Im Auftrag des Bundesamtes für Naturschutz hatte die Hochschule Karlsruhe 2006 – 2009 eine deutschlandweite Auenbilanzierung durchgeführt. Die Daten werden derzeit in einem F+E-Projekt aktualisiert und sollen im Auenzustandsbericht der Bundesregierung 2020 Verwendung finden. Es stellt sich die Frage für das Projekt, ob ein Vergleich der bilanzierten Landnutzungsdaten der beiden Auendatensätze (2007 und 2017) auf Basis der unterschiedlichen Landnutzungsdatensätze belastbare Ergebnisse der tatsächlichen Landschaftsveränderung erzielen kann.

1 Stand des Projektes

Im Auftrag des Bundesamtes für Naturschutz hatte die Hochschule Karlsruhe 2006 – 2009 eine deutschlandweite Auenbilanzierung durchgeführt. In einem dreijährigen F+E-Projekt (2016 – 2019) werden die Daten derzeit aktualisiert und sollen im Auenzustandsbericht der Bundesregierung 2020 Verwendung finden.

Aufgrund genauerer und neu verfügbarer Datensätze (wie DGM10, LBM-DE, Ergebnisse der EU-Hochwasserrisikomanagement-Richtlinie) werden sich die vor ca. 10 Jahren erfassten Auenabgrenzungen verändern. Hinzu kommen natürlich die anthropogen beeinflussten Landschaftsveränderungen, die dokumentiert und ausgewertet werden sollen.

Neben der Erfassung der Auenflächen an Flüssen in Deutschland von ca. 10.000 km, stand vor allem die Bilanzierung der Landnutzung im Vordergrund. Basierend auf dem damals zur Verfügung stehenden BasisDLM mussten die zahlreichen, sich zum Teil überlagernden bzw. nicht flächendeckend abschließenden Objektarten zunächst aufwändig vorbereitet und dann nach sieben Hauptlandnutzungsklassen ausgewertet werden. Die Ergebnisse sind in der BfN-Schriftenreihe Naturschutz und Biologische Vielfalt (Brunotte et al. 2009) dokumentiert.

Seit 2009 steht von Seiten des BKG der für solche Bilanzierungsfragen hervorragend geeignete Datensatz des LBM-DE (Landschaftsbedeckungsmodell) zur Verfügung. Dabei handelt es sich um einen flächendeckenden Datensatz u.a. mit den CORINE-Landcover-Klassen, der auf den BasisDLM-Daten basiert, aber vor allem auf Grundlage von Satellitenbildauswertungen aktualisiert und verfeinert wurde.

Es stellt sich die entscheidende Frage für das Projekt, ob ein Vergleich der bilanzierten Landnutzungsdaten der beiden Auendatensätze (2007 und 2017) auf Basis der unterschiedlichen Landnutzungsdatensätze überhaupt belastbare Ergebnisse erzielen kann, d.h. kann dadurch die tatsächliche Änderung in der Landnutzung bilanziert und dokumentiert werden oder sind die

[1] Hochschule Karlsruhe – Technik und Wirtschaft, Fakultät Informationsmanagement und Medien, Moltkestraße 30, 76133 Karlsruhe, E-Mail: detlef.guenther-diringer@hs-karlsruhe.de
[2] Karlsruher Institut für Technologie (KIT), Am Fasanengarten, D-76131 Karlsruhe, E-Mail: braun(at)tmb.uni-karlsruhe de

methodischen Unterschiede in den Daten so groß (Detaillierungsgrad, unterschiedliche Landnutzungsklassen, etc), dass sie die tatsächlichen Änderungen in der Landnutzung wesentlich überlagern.
Im Rahmen einer Bachelor-Thesis wird eine Landnutzungsbilanzierung mit den aktuellen LBM-DE-Daten auf Basis der alten Auenabgrenzung durchgeführt. Anschließend werden die Unterschiede pro vorhandenes KM-Flussauensegment prozentual miteinander verglichen. Auf Basis von aktuellen und historischen Satelliten- bzw. Luftbildern (2006 und früher, verfügbar in GoogleEarth) werden die Landnutzungsdaten in einzelnen KM-Segmenten miteinander verglichen und dahingehend klassifiziert, ob eine tatsächliche Landnutzungsänderung stattgefunden hat oder die unveränderte Landnutzung nur detaillierter bzw. unterschiedlich interpretiert wurde.
Die Ergebnisse der Thesis sind von großer Bedeutung für die Aussagekraft über die Entwicklung der Landnutzung in den untersuchten Auengebieten in den letzten 10 Jahren.
Zum Zeitpunkt der Texteinreichung waren die auswertenden Arbeiten noch nicht abgeschlossen. Sie werden in einem Projektworkshop des Projekts Auenbilanzierung 2.0 Ende Februar 2018 diskutiert und können dann auf dem PFGK-Kongress in München aktuell vorgestellt werden.

2 Literaturverzeichnis

BRUNOTTE, E., DISTER, E., GÜNTHER-DIRINGER, D., KOENZEN, U. & MEHL, D., 2009: Flussauen in Deutschland. Erfassung und Bewertung des Auenzustands. Naturschutz und Biologische Vielfalt, Bonn, Heft 87, 243 S.

Untersuchungen zur kartographischen Qualität von touristischen Stadtplänen

NADINE PREUß[1]

Zusammenfassung: Gedruckte Stadtpläne haben auch in Zeiten des Internets nicht ihre Bedeutung verloren, jedoch ist die Erstellung dieser nicht mehr alleiniges Hoheitsgebiet von Kartographen. Gerade im Bereich der touristischen Kartographie erstellen zum Beispiel oft Marketingagenturen die Karten. Der Beitrag beschäftigt sich mit der Frage, inwiefern diese Entwicklung Auswirkungen auf die kartographische Qualität hat. Untersucht werden dazu verschiedene Stadtpläne aus der ganzen Welt, vorrangig jedoch aus Deutschland. Dabei stehen Parameter wie das Vorhandensein aller kartographischen Basiselemente, Art der verwendeten Signaturen und Umgang mit diesen, aber auch der Karteninhalt und dessen Mehrwert für Touristen im Fokus.

Abstract: Even in the digital age printed city maps are of importance. However, today these maps are no longer produced by professionals only. Especially in the field of tourist mapping marketing agencies create the maps. This article asks if this development has effects on the quality of the cartographic products. In focus are various city maps from the whole world, prior though from Germany. Investigated factors are: the question whether all cartographic base elements exit, the type of used signatures and the handling of them but also the content of the maps and its additional benefits for tourists.

1 Einleitung

Auch in Zeiten von Smartphones und Navigationsapps, interaktiven Karten im Internet und GPS Geräten haben gedruckte Karten nach wie vor nicht an Bedeutung verloren und finden in verschiedenen Bereichen Einsatz. Ein Gebiet, in dem solche Karten Einsatz finden, ist der Tourismus. Insbesondere Städtetouristen finden eine große Anzahl von Plänen zur Auswahl vor. Diese können bereits vor der Reise im örtlichen Buchhandel gekauft oder per Internet bestellt werden. Oft gibt es aber auch kostenlose Angebote vor Ort: Touristeninformationen oder Hotels bieten ebenfalls Stadtpläne an. Diese gebührenfrei verteilten Karten stehen im Fokus des Forschungsinteresses.

Im Bereich der touristischen Kartographie hat sich jedoch mittlerweile ein neues Problem ergeben: die Karten werden nicht mehr ausschließlich von ausgebildeten Kartographen erstellt, sondern vielfach auch von Marketingagenturen oder von den Touristeninformationen selbst. Somit ergibt sich die Frage, ob und inwiefern diese Entwicklung Auswirkungen auf die kartographische Qualität hat. Liegt bei Laien genug kartographisches Wissen vor, um Karten zu erstellen, die verständlich sind und gleichzeitig auch einen angemessenen Mehrwert für Touristen bieten? Es ergibt sich auch das Problem der Finanzierung eines Angebotes in Form

[1] TU Dortmund, Fakultät für Raumplanung, Lehrstuhl für Bodenpolitik, Bodenmanagement und kommunales Vermessungswesen, August-Schmidt-Str. 10, D-44221 Dortmund, E-Mail: nadine.preuss@tu-dortmund.de

einer kostenlosen „Karte“. Welche Interessenslagen beispielsweise von ortsansässigen Unternehmen stehen möglicherweise hinter einem solchen Sponsoring und haben auch Einfluss auf die kartographische Aussage?
Um eine Antwort auf diese Fragen zu geben, wurde eine Auswahl an touristischen Stadtplänen begutachtet.

2 Hintergrund

Während der Entwicklungsgeschichte der Kartographie war die Kartenerstellung immer nur eine Aufgabe von einigen wenigen Fachleuten, die die Karten in mühevoller Handarbeit anfertigten. Mittlerweile haben Computer und die entsprechende Software die Kartographie für eine größere Gruppe von Menschen geöffnet. Parallel dazu existieren so viele verschiedene Kartentypen wie niemals zuvor. Gerade im touristischen Bereich findet sich eine Vielzahl von unterschiedlichen kartographischen Produkten: Stadtpläne, Wanderkarten, Autokarten zur Anreise, Fahrradkarten, Skikarten usw.
Im Fokus dieses Beitrags stehen die touristischen Stadtpläne, die in der Regel von jeder Stadt, die nur annähernd touristisch interessant oder erschlossen ist, herausgegeben werden. Im Gegensatz zu den oftmals hochkomplexen Wanderkarten mit Höhenlinien, Schummerungen usw. werden die Stadtpläne aufgrund der bereits angesprochenen digitalen Entwicklung nicht mehr nur von Kartographen angefertigt.
Aus dieser Entwicklung ergibt sich das Forschungsinteresse, zumal touristische Karten auch in der Fachliteratur bislang wenig Beachtung finden, obwohl sie allgegenwärtig sind.

3 Die Untersuchung

Für genaue Aussagen über die kartographische Qualität von touristischen Stadtplänen wurden verschiedene Stadtpläne aus der ganzen Welt, vordergründig jedoch aus Deutschland, analysiert. Dabei wurde untersucht, wer die Karte erstellt hat und ob die Karte alle kartographischen Basiselemente enthält. Weiterhin interessant sind die Art der verwendeten Signaturen und die Lesbarkeit bzw. Nutzbarkeit dieser. Darüberhinaus zu begutachten ist die Kartenbeschriftung sowie die Frage, ob der Karteninhalt auf die gewählte Zielgruppe abgestimmt ist. Zuguterletzt wird sich noch mit der möglicherweise vorhandenen Werbung auf den Karten beschäftigt, da Werbeanzeigen oftmals zur Finanzierung beitragen.

3.1 Untersuchte Parameter

In diesem Absatz werden die verschiedenen untersuchten Parameter vorgestellt:

Kartenersteller

Wie bereits festgestellt, können Karten heutzutage von den verschiedensten Personen erstellt werden. Im Bereich der touristischen Kartographie sind dies vor allem professionelle Kartographen bzw. Kartographiefirmen, aber auch Werbeagenturen und die Tourismusbehörden und deren Mitarbeiter selbst. Es ist auch eine Kombination denkbar: Die Karte selbst wird von Kartographen angefertigt, das Layout um die Karte herum oder das Faltblatt oder die Broschüre,

in die die Karte möglicherweise eingebettet ist, wird wiederum von beispielsweise der Marketingagentur entworfen.
Es ist essentiell zu wissen, wer genau die Karte erstellt hat, um so beurteilen zu können, ob Qualitätsunterschiede zwischen den Karten, die von professionellen Kartographen und Karten, die von Laien erstellt wurden, bestehen.

Kartographische Basiselemente

Als kartographische Basiselemente werden der Kartentitel, die Legende, das Impressum, die Maßstabszahl und Maßstabsleiste sowie der Nordpfeil betrachtet.
Der *Kartentitel* ist wichtig, um auf einen Blick die geneigten Kartennutzer über das Thema der Karte zu informieren.
Die *Legende* wiederum ist eins der wichtigsten Kartenelemente (GROßER 2001). Hier werden alle Symbole erklärt, die in der Karte Verwendung finden. Manche Symbole sind leicht verständlich, da allgemein bekannt oder sehr bildhaft (WILHELMY 1981:245) während andere möglicherweise sehr abstrakt und komplex und den Kartennutzern nicht so geläufig sind. Trotzdem sollte eigentlich jedes Symbol in der Legende erklärt werden, unabhängig davon, ob diese Symbole eher bildhaft oder eher abstrakt sind (GROßER 2001). Gegenstand der Untersuchung war hier also die Frage, ob, erstens, eine Legende vorhanden ist oder nicht, und zweitens, ob diese Legende auch vollständig ist und alle verwendeten Signaturen in ihr erklärt werden.
Um Entfernungen abschätzen und Größen beurteilen zu können, benötigt eine Karte eine *Maßstabszahl* und/oder eine *Maßstabsleiste*. Der Maßstab kann auf verschiedene Art und Weise angegeben werden. Im Idealfall ist der Maßstab sowohl als Maßstabszahl als auch als Maßstabsleiste angegeben, um die Arbeit mit der Karte so leicht wie möglich zu machen (SALISTSCHEW 1967:40). Es wird bei jeder Karte darauf geachtet, ob eine Maßstabsangabe vorhanden ist und in welcher Form. Wichtig hierbei ist, dass die Maßstabsangabe nur eine Hilfestellung für die Kartennutzer ist und die Karte nicht optisch dominieren sollte.
Obwohl die meisten Karten heutzutage genordet sind, wird nach wie vor empfohlen, einen *Nordpfeil* zu nutzen, um die Richtung der Karte anzugeben (TYNER 2014:35). Teil der Untersuchung ist also auch, ob die Karte einen Nordpfeil besitzt und ob sie überhaupt genordet ist. Inbesondere bei einer nichtgenordeten Karte ist ein fehlener Nordpfeil ein schwerwiegender Mangel.
Das letzte kartographische Basiselement ist das *Impressum*. Das Impressum sollte mindestens den Namen und/oder die Institution des Kartenerstellers enthalten, die Datenquelle angeben und das Produktionsjahr der Karte. Die Angabe der Datenquelle ist entscheidend, um die Authenzität der Daten zu zeigen (TYNER 2014:39). Das Produktionsjahr verrät dem Kartennutzern wiederum die Aktualität der Karte und inwiefern er sich auf die darin gemachten Angaben verlassen kann. Während der Untersuchung wurde festgehalten, ob ein Impressum vorhanden ist und wenn ja, ob dieses auch vollständig ist. Gemäß dem Impressum wurde auch der Kartenersteller notiert. Da die Frage, wer die Karte erstellt hat, jedoch ein entscheidender Teil der Untersuchung ist, wurde dies in einem Extrapunkt abgehandelt (s.o.).

Signaturen

Die geographischen Phänomene, die grafisch in einer Karte dargestellt werden, lassen sich in die drei Kategorien Punkt, Linie und Fläche einteilen (KOHLSTOCK 2014:78, TYNER 2014:138) und können so auch dargestellt werden. Dabei wird beispielsweise zwischen Farbe, Strichstärke, Orientierung etc. varriiert (KOHLSTOCK 2014:78, TYNER 2014:137). Von Karten sind jedoch meistens Signaturen bekannt, die wesentlich komplexer als reine geometrische Darstellungen sind, beispielsweise ein Flugzeug, das für einen Flughafen steht. Hake teilt diese Signaturen in verschiedene Gruppen ein: symbolisch und bildhaft. Die bildhaften Signaturen können dabei noch weiter in Grundrissbilder, Aufrissbilder und Schrägbilder unterteilt werden (Abb.1).

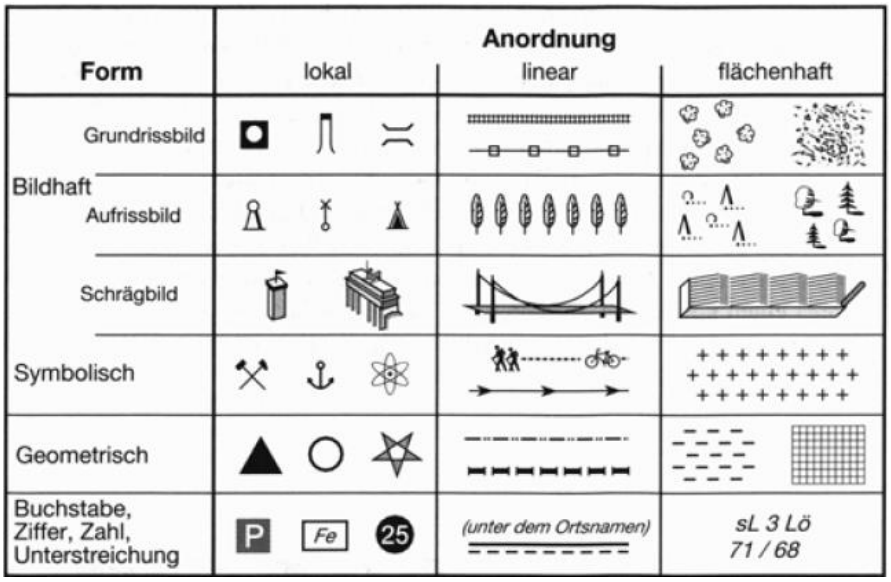


Abb. 1: Verschiedene Arten von Signaturen (HAKE 1994:101, gekürzt)

Es wird bei den vorliegenden Karten also untersucht, welche Signaturtypen in der Praxis vorherrschen.

Beschriftung

Eine vollständige Karte besteht nicht nur aus Signaturen, sondern benötigt auch eine Beschriftung. Beschriftet werden könnten im Beispiel von den Stadtplänen dabei Straßenzüge, wichtige und/oder interessante Gebäude, Stadtteile, Haltestellen von öffentlichen Verkehrsmitteln etc.. Ähnlich wie bei den klassischen Signaturen kann auch bei der Beschriftung eine Variation von zum Beispiel Farbe, Größe aber auch Schriftart oder Schriftauszeichnung erfolgen, um so verschiedene Informationen auszudrücken. Untersucht werden soll hier also, ob eine ausreichende Beschriftung vorhanden und auch lesbar ist sowie deren Variationsreichtum.

Karteninhalt

Die Untersuchung des Karteninhalts schließlich beschäftigt sich nicht mehr nur mit rein handwerklichen Aspekten der Kartenerstellung, sondern geht auf die thematische Aussage der Karte selbst ein. Eine technisch einwandfrei gestaltete Karte nützt den Touristen wenig, wenn der Inhalt dieser nicht auch an dessen Bedürfnisse angepasst ist.

Touristen benötigen zunächst natürlich einen herkömmlichen Stadtplan, mit dem er sich einfach in der Stadt orientieren kann. Hierzu sollten die Straßen aufgeführt und beschriftet sein und Gebäude und andere wichtige Orte, die der Orientierung dienen, vermerkt werden. Jedoch interessieren sich Touristen natürlich nicht nur für Angaben, die seiner Orientierung in der fremden Stadt dienen. In erster Linie wird eine Stadt bereist mit dem Gedanken, etwas bestimmtes zu sehen oder zu erleben. Einrichtungen aus verschiedenen Kategorien können für Touristen interessant sein. Dazu zählen beispielsweise: Kultur, Sport, Unterhaltung, Einkaufen… (KOSMA et al. 2003:27 ff.). Diese Kategorien lassen sich natürlich noch weiter aufteilen in konkrete Einrichtungen, die möglicherweise von Touristen besucht werden möchten, und die sich dementsprechend in der Karte wiederfinden lassen sollten, um den Touristen so darauf aufmerksam zu machen bzw. ihm zu zeigen, wo er was finden kann. In den Bereich Kultur lassen sich so beispielsweise architektonisch interessante Gebäude wie Burgen, Ruinen, Schlösser oder auch historische Plätze einordnen (KOSMA et al. 2003:28). Auch sportliche Betätigung kann im Urlaub für einige Touristen interessant sein. Hier könnten beispielsweise Minigolfplätze, Hochseilgärten, Kletter- und Abenteuerparks oder Beachvolleyballfelder (KOSMA et al. 2003:28) in der Karte vermerkt werden. Für andere Touristen wiederum spielt das Einkaufserlebnis während ihres Urlaubs eine wichtige Rolle. In diesem Bereich ist es erforderlich, Fußgängerzonen oder andere zentrale Einkausstraßen sowie Shopping Malls (KOSMA et al. 2003:29) in der Karte zu markieren. Die Kategorie der Unterhaltung ist weit gegriffen: hierunter können sowohl Pubs, Bars, Cafés und Diskotheken fallen aber auch Erlebnisbäder und Themenparks (KOSMA et al. 2003:29).
Für Touristen weiterhin wichtig ist nicht nur, wo sich welche Attraktionen befinden, sondern auch, wie diese zu erreichen sind. Hierbei könne zum Beispiel Angaben von U-Bahnhaltestellen (falls ein U-Bahnnetz vorhanden ist), Bushaltestellen oder Parkmöglichkeiten hilfreich sein.

Umgang mit Werbung
Wie bereits erwähnt, sind Stadtpläne oftmals kostenlos an Touristinformationen oder ähnlichen Orten erhältlich. Dies beinhaltet allerdings auch das „Problem“ der Finanzierung solcher Karten. Zur Finanzierung kann die Platzierung von Werbung in und um die Karte für lokale Restaurants, Gaststätten, Geschäfte etc. beitragen. Untersuchenswert hierbei ist, ob die Werbung dabei soviel Raum auf oder um die Karte einnimmt, das touristisch wichtige Informationen optisch verdrängt werden oder ob sich die Werbung harmonisch in das Kartenbild einfügt. Zusätzlich sollte nicht außer Acht gelassen werden, dass Werbung auch nicht immer negativ gesehen werden sollte. Schließlich besteht hier auch die Möglichkeit für Touristen, Informationen über Restaurants, Museen oder andere Orte zu erhalten, die seinem Geschmack oder Interesse entsprechen und die er dementsprechend aufsuchen möchte.

3.2 Erste Untersuchungsergebnisse

Mehr als 100 Karten wurden bereits hinsichtlich der eben vorgestellten Parameter analysiert, sodass bereits erste Untersuchungsergebnisse vorliegen, die hier präsentiert werden können.

Kartenersteller

In vielen Fällen ist keine Angabe darüber vorhanden, wer die Karte erstellt hat. Sofern ein Kartenersteller erwähnt wird, ist dieser entweder eine Marketing- oder eine Kartographieagentur. Selten wurden die Karten auch von Katasteramt, Stadtvermessung oder GIS-Abteilung der jeweiligen Stadt erstellt. Karten, die explizit von den Tourismusabteilungen der Städte erstellt wurden, waren bis jetzt nicht aufzufinden. Da es bei gut der Hälfte der untersuchten Karten jedoch keinen Hinweis auf den Kartenersteller gibt, kann so nicht abschließend gesagt werden, dass dies nie der Fall ist, dass Stadtpläne von Touristikfachkräften erstellt werden.
Der Herausgeber dagegen wird auf den meisten Karten genannt. Nur in einigen wenigen Fällen fehlt diese Angabe. Die Herausgeber sind in der Regel Marketingagenturen, gefolgt von Tourismusagenturen bzw. den Tourismusabteilungen der Städte. Nur in vereinzelten Fällen sind die Herausgeber Kartographieagenturen.
Aus diesen Beobachtungen lässt sich schließen, dass Kartenerstellung und Herausgabe des fertigen kartographischen Produkts nicht immer in der Hand nur eines Verantwortlichen liegen müssen. In einigen Fällen wird die Karte von professionellen Kartographen erstellt, während die Herausgabe (und Gestaltung des Gesamtwerks) von anderen erfolgt. Eine große Rolle spielen in beiden Fällen Marketingagenturen.
Das Fehlen der Angabe des Kartenerstellers ist jedoch problematisch für die Ausgangsfrage, ob Karten, die nicht von professionellen Kartographen hergestellt wurden, Qualitätsmängel aufweisen.

Kartographische Basiselemente

Zunächst ist positiv anzumerken, dass jede untersuchte Karte einen *Titel* hat, der Auskunft über das Thema der Karte geben konnte. Touristen erfahren hier also schon schnell erkennbar, was sie erwartet.
Bezüglich des Maßstabs sieht dies jedoch anders aus: eine *Maßstabszahl* ist so gut wie nie vorhanden, während sich eine *Maßstabsleiste* bei ungefähr der Hälfte der untersuchten Karten finden lässt. Wenn davon ausgegangen wird, dass die meisten Touristen Laien in der Arbeit mit Karten sind, könnte argumentiert werden, dass die Maßstabszahl nicht unbedingt notwendig ist, da sie für Menschen, die nicht so oft mit Karten zu tun haben, nicht besonders aussagekräftig ist. Die graphische Maßstabsleiste macht es für kartographisch ungeübte Menschen leichter, Entfernungen und Größen abzuschätzen. Ein Vorhandensein einer Maßstabsangabe ist gerade im Bereich der touristischen Stadtpläne unerlässlich, da es nur so für ortsfremde Touristen möglich ist, zu wissen, ob beispielsweise die favorisierte Sehenswürdigkeit noch in fußläufiger Entfernung ist oder ob der Weg mit dem Auto oder öffentlichen Verkehrsmitteln zurückgelegt werden muss. Das absolute Fehlen einer Maßstabsangabe, wie es bei vielen Stadtplänen der Fall ist, stellt also einen großen Mangel dar.
Ein *Nordpfeil* wiederum ist bei mehr als der Hälfte der gesichteten Karten nicht vorhanden. Da so gut wie alle Karten auch genordet waren, ist dieser Mangel zu vernachlässigen. In den wenigen Fällen, bei denen die Karte nicht genordet war, war auch ein Nordpfeil auf der Karte zu finden, der darauf hinweist.

Die *Legende* ist unerlässlich für das Verständnis der verwendeten Signaturen und damit auch das Gesamtverständnis der Karte. Eine Legende ist auch bei gut 80% der begutachteten Karten vorhanden, allerdings war diese oftmals nicht vollständig, das heißt, nicht alle verwendeten Signaturen wurden erklärt.
Ein *Impressum* ist in bei der Hälfte der analysierten Karten vorhanden, jedoch ist dies nicht immer vollständig, was sich auch in dem Punkt wiederspiegelt, dass nicht immer ein Kartenersteller oder ein Herausgeber ausfindig gemacht werden konnte. Problematisch ist es auch, wenn im Impressum keine Angabe zur Zeit der Veröffentlichung zu finden ist, da so die Aktualität der Karte und der darin enthaltenen Informationen nicht eingeschätzt werden kann.

Signaturen

In jeder untersuchten Karte wurde mit verschiedenen Signaturentypen gearbeitet. Die „Grundelemente“ Punkt, Strich und Fläche waren in jedem Stadtplan gegeben. Besonders von Interesse sind die Punktsignaturen, da diese meist für touristisch interessante Sehenswürdigkeiten oder Freizeitmöglichkeiten stehen und somit die Frage, wie diese besonders in der Karte dargestellt oder hervorgehoben werden.
Besonders auffällig sind Aufrisse bzw. Schrägansichten von Gebäuden, die sich in vielen Stadtplänen finden lassen (Abb.2).

Abb. 2: Schrägbild einer Kirche im Stadtplan von Halle (STADTMARKETING HALLE(SAALE) GMBH:2016)

Diese springen dem Kartenbetrachter natürlich sofort ins Auge und regen ihn möglicherweise dazu an, das entsprechende Gebäude zu besichtigen, da er es aufgrund der Darstellung für wichtig erachtet. Ansonsten finden sich hauptsächlich geometrische oder bildhafte Signaturen auf den untersuchten Stadtplänen. Wichtig bei den Signaturen ist, dass diese eine ausreichende Größe haben und sich farblich vom Hintergrund abheben, um so gut vom Betrachter erkannt zu werden. Dies ist bei den untersuchten Karten in der Regel gegeben, nur einige wenige Karten fallen dadurch auf, dass Signaturen zu klein oder ohne große farbliche Abhebung zum Hintergrund und damit schlecht erkennbar sind.

Beschriftung

Auf allen untersuchten Karten ist eine Beschriftung vorhanden. Beschriftet werden dabei Straßen, aber auch Stadtteile, wichtige Gebäude etc. Die Menge der Beschriftung ist in der Regel ausreichend, lediglich kleinere Straßenzüge werden nicht beschriftet.
Die verwendeten Schriftarten sind in so gut wie allen Fällen serifenlos; es wird wenig innerhalb einer Karte mit verschiedenen Schriftarten, -farben oder -auszeichnungen variiert.
Beim Großteil der Karten ist die Lesbarkeit der Schrift gut. Nur teilweise wird sie von anderen Signaturen verdeckt, ist zu klein oder hebt sich farblich nicht gut von der Kartengrundlage ab.
Zwar ist die Verwendung der Schrift auf den gesichteten Karten nicht unbedingt als kreativ zu bewerten und es wurde keine variierende Beschriftung für unterschiedliche Dinge genutzt, jedoch ist die Lesbarkeit der Beschriftung in der Regel gegeben, sodass dieser Punkt bei den touristischen Stadtplänen insgesamt gesehen als positiv bewertet werden kann.

Karteninhalt

Die Informationen, die sich auf den Stadtplänen befinden, sind in der Regel als durchweg positiv und nützlich zu bewerten. Die Touristen finden viele Informationen zu Sehenswürdigkeiten aus dem Bereich der Kultur wie historisch interessante Gebäude, Museen etc. Auch der Bereich für sportlich interessierte Touristen kommt nicht zu kurz: Minigolfplätze, Parks etc. sind (wenn vorhanden) ebenfalls in den Karten vermerkt. Darüber hinaus lassen sich Shopping Malls und Fußgängerzonen oder andere Einkaufsstraßen mit Hilfe der Stadtpläne auffinden. Im Bereich der Fußgängerzonen zeigt sich allerdings ein Problem, das bereits an anderer Stelle erwähnt wurde: Die entsprechenden Zonen sind in der Karte zwar in der Regel farblich markiert, jedoch wird in den oft unvollständigen Legende diese Farbe nicht erklärt, sodass dem Touristen nicht unbedingt klar wird, dass beispielsweise die hellgelbe Straße nun die gesuchte Fußgängerzone ist. Lokalitäten wie Restaurants, Cafés, Diskotheken etc. finden sich eher selten in den vorliegenden Karten. Wenn diese in den Karten markiert werden, handelt es sich hier eher um Werbeanzeigen und kein neutrales Gesamtbild der Gaststätten der Stadt (siehe auch „Umgang mit Werbung“). Die Informationen finden sich dabei direkt in der Karte vermerkt oder auch oft auf der Kartenrückseite oder rund um die Karte in Textform, oft auch mit Foto.
Insgesamt lässt sich also feststellen, dass die touristischen Stadtpläne in der Regel einen echten Mehrwert für Touristen bieten, da die Touristen hier viele nötige Informationen über interessante Sehenswürdigkeiten und weitere Freizeitmöglichkeiten erhalten.
Auch die Anreise zu den Sehenswürdigkeiten ist gewährleistet, da sich in den Karten Angaben zu Parkplätzen oder öffentlichen Verkehrsmitteln finden. Gerade in größeren Städten mit einem ausgedehnten Netz an öffentlichen Verkehrsmitteln ist der Karte oftmals, beispielsweise auf der Kartenrückseite, auch ein Extra-Plan dieser beigefügt.
Für weitere nötige Informationen sind die Standorte der Touristeninformationen in der Karte vermerkt, oftmals sind am Kartenrand auch deren Sehenswürdigkeiten und Kontaktnummern angegeben.

Einsatz von Werbung

Überraschenderweise haben einige der untersuchten Karten gar keine Werbung und dies, obwohl sie sogar gratis erhältlich sind. Die Finanzierung dieser Karten erfolgt hier also nicht über Werbepartner. Auch andere der gesichteten kartographischen Produkte weisen nur sehr wenige Werbeanzeigen auf, die nicht zu aufdringlich im Vordergrund stehen und somit nicht die kartographischen Informationen beeinträchtigen. Die Werbeanzeigen sind dabei hauptsächlich von Hotels, Restaurants und diversen Läden und Freizeiteinrichtungen, sodass sie durchaus auch einen Mehrwert für Touristen bieten können, der so darüber informiert wird, wo er essen, nächtigen oder etwas besichtigen kann.
Natürlich gibt es auch einige kartographische Produkte, die diesen positiven Gesamteindruck nicht bestätigen. Auf einigen Karten findet sich auch überproportional viel Werbung, die das Gesamtbild sehr dominiert und das Kartenbild optisch in den Hintergrund drängen. Es fallen auch vereinzelt Karten auf, die nur für ein Restaurant oder eine Einrichtung werben, also reines Werbeprodukt eines bestimmten Ladens sind. Beispielhaft soll hier eine Karte von Rhodos gezeigt werden, die zwar auch eine Sehenswürdigkeiten von Rhodos aufführt, darüber hinaus aber hauptsächlich zeigen möchte, wo sich die „Taverna Porto Antico“ befindet (Abb.3).

Abb. 3: Ausschnitt aus einer Karte von Rhodos

4 Zusammenfassung der Ergebnisse und Ausblick

Die Untersuchung der Stadtpläne hat gezeigt, dass diese von insgesamt guter Qualität sind. Mängel lassen sich vor allem im Bereich der kartographischen Basiselemente finden, die nicht auf allen Karten oder nicht immer vollständig vorhanden sind. Problematisch ist dies beispielsweise im Bereich der fehlenden Maßstabsleisten, die es Touristen so unmöglich machen, Entfernungen abzuschätzen. Die nicht vorhandene Angabe des Kartenerstellers im Impressum mag für Touristen als Kartennutzer nicht so schlimm sein; hier ergaben sich vor allem Probleme für die Ausgangsfragestellung nach Qualitätsunterschieden bei unterschiedlichen Kartenerstellern, da diese ohne diese Information natürlich nicht abschließend beantwortet werden kann. Auch die fehlende Angabe eines Produktionsdatums sowie die unvollständige Legende stellen einen Qualitätsmangel dar.
Erfreulich dagegen ist, dass die Signaturen in der Regel hochwertig sind. Auch die Beschriftung kann als positiv bewertet werden. Zwar ist sie in den meisten Fällen nicht besonders variationsreich, erfüllt aber aufgrund der guten Lesbarkeit ihren Zweck.
Weiterhin von großer Bedeutung ist der Karteninhalt. Hier kann festgestellt werden, dass dieser bei allen Karten die Bedürfnisse der Touristen erfüllt. Auch ist es positiv, dass die meisten Karten nicht von Werbung überladen sind – einige haben gar keine Werbung, bei anderen fügen sich die Werbeanzeigen harmonisch in den Karteninhalt ein und bieten eventuell sogar Zusatzinformationen für die Kartennutzer, wie beispielsweise wo er was essen gehen oder übernachten kann etc.
Abschließend lässt sich sagen, dass die Ausgangsfragestellung ob es kartographische Qualitätsunterschiede bei unterschiedlichen Kartenerstellern gibt, nicht abschließend beantwortet werden kann, da nicht immer ersichtlich ist, wer nun die Karte erstellt hat. Allerdings lässt sich bei der Gesamtbetrachtung der Untersuchungsergebnisse feststellen, dass die gesichteten Stadtpläne durchweg von guter Qualität und ohne gravierende Mängel sind. Die meisten „Fehler“ finden sich im Bereich der kartographischen Basiselemente, jedoch fallen diese einem Laien möglicherweise gar nicht erst auf bzw. machen die Karte für ihn nicht unbrauchbar. Wichtig für Touristen werden vor allem der für ihn angepasste Karteninhalt und hochwertige Signaturen sowie eine ausreichende, gut lesbare Beschriftung sein – diese Punkte werden von allen Karten in einem ausreichenden Maße erfüllt.
Weiterhin zu beachten ist, dass hier nur eine Auswahl von Parametern, die ein Anzeichen für kartographische Qualität sein können, ausgewertet und vorgestellt wurde. Weitere interessante Punkte, die noch untersucht werde, sind beispielsweise der Einsatz von Farben (Farbauswahl, Farbassoziationen usw.), die visuelle Hierarchie der Stadtpläne, das Gesamtlayout und die Platzierung der einzelnen Kartenelemente, usw. Ergebnis der noch kommenden Untersuchungen könnte ein kartographischer Leitfaden sein, der bei der Erstellung touristischer Stadtpläne zur Rate gezogen werden kann.

5 Literaturverzeichnis

GROßER, K., 2001: Definition Legende, Heidelberg. In: Bollmann, Jürgen, Koch, Wolf – Günther and Annette Lipinski (pub.) 2001a: Lexikon der Kartographie und Geomatik, Heidelberg. [Online Version]

HAKE, G., GRÜNREICH, D. & MENG, L., 2002: Kartographie, Berlin. 8. überarbeitete Auflage.

KOHLSTOCK, P., 2014: Kartographie – Eine Einführung, Paderborn. 3. überarbeitete Auflage.

KOSMA, P., GLASER, F., WÖLFL, P., SCHADEN, J., TURISSER-GALA, C. & DERFLINGER-TRAXLER, E., 2009: Betriebsorganisation, Tourismus, Marketing, Linz. 13. überarbeitete Auflage

SALISTSCHEW, K., 1967: Einführung in die Kartographie, Band 1, Gotha/Leipzig.

TYNER, J., 2014: Principles of Map Design, New York.

STADTMARKETING HALLE (SAALE) GMBH, 2016: Innenstadtplan Halle (Saale), o.O.

WILHELMY, H., 1981: Kartographie in Stichworten, Kiel. 4. überarbeitete Auflage, Gotha/Leipzig.

Three years of operations of the Sentinel-1 and Sentinel-3-OLCI PAC at German Aerospace Center

THOMAS HAHMANN[1] & ERHARD DIEDRICH[1]

Abstract: Copernicus is the European Earth Observation Programme, conducted jointly by the EC, ESA, EUMETSAT and the member states. The Sentinel satellites constitute the Copernicus space segment. Since 2014 already six Sentinel satellites have been launched successfully. The core payload data ground segment (PDGS) for the Sentinel satellites is managed by ESA and operated by national partners. In this core PDGS, DLR has been selected to install and operate Processing and Archiving Centers (PACs) for Sentinel-1 and the data from the Ocean and Land Colour Instrument (OLCI) aboard Sentinel-3. This paper describes the results of the first three years of operations of the S1- and S3-OLCI-PAC at DLR.

Daily, around 950 products are being processed based on sensor data from the OLCI instrument on Sentinel-3A, over 330 GB of associated data are stored in the Long-Term Archive of the DLR-PAC. The Sentinel-1A and Sentinel-1B satellites are already supplying many times that amount. Over 10,000 GB (10 TB) and more than 4700 products from the satellite pair are handled every day. In the first three years of operations the DLR-PAC has already processed and archived over 3.8 million data sets. They represent a data volume exceeding 6500 terabytes (6.3 petabytes).

1 Mission Context: the Copernicus Architecture

Copernicus (formerly known as Global Monitoring for Environment and Security (GMES) until December 2012) is an initiative led by the EU (ASCHBACHER & MILAGRO-PÉREZ 2012). The major Copernicus building blocks are: the Space Infrastructure, to be implemented by ESA with the Sentinel satellites, Earth Observation (EO) missions developed by ESA specifically for Copernicus; complemented by Contributing Missions, EO missions built for purposes other than Copernicus, but offering part of their capacity to Copernicus (EU/ESA Missions, EUMETSAT, commercial, international); the In-Situ Infrastructure, coordinated by the EEA; the Services Component in support of the large and diverse user community offering Copernicus Services coordinated by the EC and User (Downstream) Services, intended to be provided by commercial entities.

DLR is contributing to the development and implementation of the Copernicus system in various fields since several years (SCHREIER et al. 2013).

[1] German Aerospace Center – DLR, German Remote Sensing Data Center – DFD, Oberpfaffenhofen, D-82234 Wessling, Germany, E-Mail: [Thomas.Hahmann, Erhard.Diedrich]@dlr.de

2 Sentinel Space and Payload Data Ground Segment (PDGS) Infrastructure

The Sentinel Space Infrastructure as central part of the Copernicus Architecture is composed of the Sentinel Space Segment and the Sentinel Ground Segment consisting of the Flight Operations Segment and the Payload Data Ground Segment (PDGS).
The Sentinel Space Segment is composed of six satellite series. Sentinel-1, -2, and -3 are pairs of two identical satellites for each mission:

- Sentinel-1A/B are polar orbiting satellites, each carrying a C-Band SAR to ensure the data continuity of the ERS-1/-2, Envisat (ASAR) and Radarsat missions (TORRES et al. 2012);
- the Sentinel-2A/B polar orbiting satellites equipped each with a 13-channel Multi-spectral Imager to provide enhanced continuity of SPOT- and Landsat-type data;
- the Sentinel-3A/B polar orbiting satellites with several sensors (SLSTR, OLCI, SRAL, and MWR) for ocean and land applications building on ERS-2 (AATSR), Envisat (MERIS), and Cryosat (SRAL) heritage sensor technology (DONLON et al. 2012).

Sentinel-4 will be a Geostationary Atmosphere satellite. Sentinel-5P (P stands for Pre-Cursor) and Sentinel-5 are Low-orbit Atmosphere / Air quality satellites. The Sentinel-6 mission will be dedicated to Low-orbit Altimetry.
Since 2014 already six Sentinel satellites have been launched successfully. The first satellite Sentinel-1A was launched on 3rd April 2014. The Sentinel-1 mission was completed with the launch of the identical twin satellite Sentinel-1B on 25th April 2016. The Sentinel-3A satellite lifted off on 16th February 2016. The envisaged launch of Sentinel-3B will be in April 2018. Further Sentinel satellites will be launched in the following years.
In parallel to the satellite development, the associated PDGS for each Sentinel mission has been built up (POTIN et al. 2017; SANTELLA et al. 2013). The Sentinel PDGS consists of the following major constituents (see Fig. 1): Four X-Band Core Ground Stations (CGS) to receive all Sentinel-1/2/3 payload data; seven Processing and Archiving Centers (PAC); three Mission Performance Centres (MPC); a Data Circulation and Dissemination Network (WAN); a Payload Data Management Centre (PDMC); and the Coordinated Data Access System (CDS).
Four Core Ground Stations receive payload data in X-Band. Each ground station is able to receive data from all Sentinels. All Core Ground Stations are offering Near-Real Time (NRT) processing services. The CGS are located in Norway (Svalbard), Italy (Matera), Spain (Maspalomas), and Canada (Inuvik).

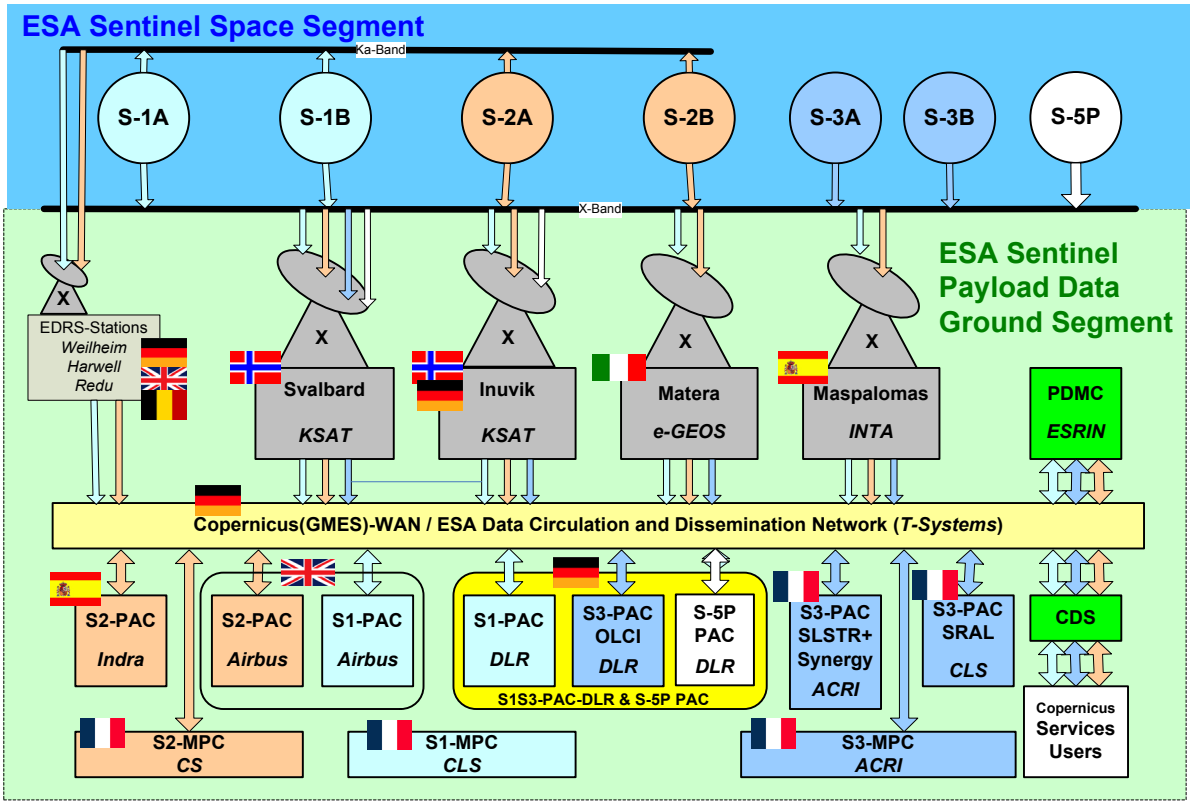


Fig. 1: Overall structure of the Copernicus Payload Data Ground Segment: Ground Stations and PACs

Seven Processing and Archiving Centers are located as follows: Sentinel-1 PAC (Newport, UK); Sentinel-1 PAC (DLR Oberpfaffenhofen, Germany); 2 Sentinel-2 PACs (Newport, UK and Madrid, Spain); Sentinel-3 OLCI-PAC (DLR Oberpfaffenhofen, Germany); Sentinel-3 SLSTR and Synergy PAC (Nice, France); Sentinel-3 SRAL-PAC (Toulouse, France). The S1-PAC and the S3-OLCI-PAC are collocated at DLR's Earth Observation Center in Oberpfaffenhofen and are integrated, operated and maintained by an experienced team of experts from the German Remote Sensing Data Center (DFD) institute.

3 Sentinel-1 and Sentinel-3-OLCI PAC at DLR

In 2011, ESA had started a series of procurement actions to select European providers to offer their facilities for the set-up and operations of the PDGS elements. Within this competitive selection process DLR was selected in 2012 by ESA for the role of a PAC operator of the Sentinel-1 and Sentinel-3 (OLCI instrument) missions (HAHMANN et al. 2013).

3.1 Tasks of the Sentinel PAC

The following high-level tasks are performed at the PAC:

- reception of Sentinel data from the network of Core Ground Stations (CGS) via electronic link (Copernicus WAN);

- ingestion of these data into the Short-Term Archive (STA) and Mid-Term Archive (MTA) of the Sentinel PGDS;
- in addition ingestion of these data in a Long-Term Archive (LTA) for addressing future science challenges;
- perform consolidation and re-assembly of level-0 data received from CGS;
- perform systematic and request-driven processing of Sentinel data to higher-level products;
- host Sentinel data products within a layered architecture of on-line dissemination elements that facilitate the data access of end-users via public networks;
- share and exchange any locally processed data with a second partner PAC for the purpose of redundancy;
- capability to reprocess large amounts of data.

For Sentinel-1 the second PAC is located in Newport (UK). For Sentinel-3 there exists a work share between the DLR-PAC (responsible for data over land areas) and EUMETSAT (datasets that cover marine areas).
The Sentinel PDGS is operationally embedded in the Copernicus Space Component Data Access (CSCDA, https://spacedata.copernicus.eu/) System that has been implemented by ESA in support of Copernicus Service Projects and their users.
Complete, free and open access to Sentinel-1, Sentinel-2 and Sentinel-3 user products for the general public is provided by the Copernicus Open Access Hub (previously known as Sentinels Scientific Data Hub, https://scihub.copernicus.eu/).
The Sentinel Online web portal provides comprehensive technical information about all Sentinel missions (https://sentinel.esa.int/).

3.2 Sentinel PAC Project Phases

The project is structured in a phased approach that consists of the Site Preparation Phase (SPP), the Deployment and Test Phase (DTP), the Commissioning Operations Phase (COP), and Routine Operations Phase (ROP). Details about the different aspects of these phases can be found in HAHMANN et al. (2013).
After the development and commissioning phase during the years 2012 – 2014, the Sentinel-1A PAC entered into the routine operations phase in March 2015. In July 2016 the routine operational phase began for Sentinel-1B and for Sentinel-3A OLCI data. The Sentinel-PAC is expected to be complete in April 2018 when the Sentinel-3B satellite will be launched into orbit. Currently the DLR-PAC is contracted by ESA until end of 2019 to fulfil its role as PAC for Sentinel-1 data and Sentinel-3 OLCI data.
In the Routine Operations Phase the following services are provided for ESA: Network Services, LTA (Long-Term Archive) Service, Production and Dissemination Service, and Support Operations such as System Administration, Database Management and Logistics.
The LAN infrastructure is operated in terms of a managed service and is continuously monitored. The LTA Service covers: Systematic ingestion of Sentinel data from the Sentinel PDGS; retrieval of archived data products for the dissemination to end-users on request of ESA; bulk retrieval of data to be used as input for reprocessing activities; systematic ingestion of data

generated in the context of reprocessing campaigns; LTA maintenance activities such as archive monitoring and archive refreshment and data re-transcription.
The Production and Dissemination Service covers systematic processing and generation of Sentinel data products of all levels; transfer of received and processed data into STA, MTA, and LTA; circulation of data products to partner centers via the multi-Sentinel WAN; request-driven dissemination of data products to Sentinel end-users.
Fig. 2 gives an overview about the high-level structure of DLR's Sentinel PAC showing the components of the LAN Service (operated by T-Systems, Germany), the LTA Service, and the Sentinel-1 and Sentinel-3 OLCI Production and Dissemination Service.

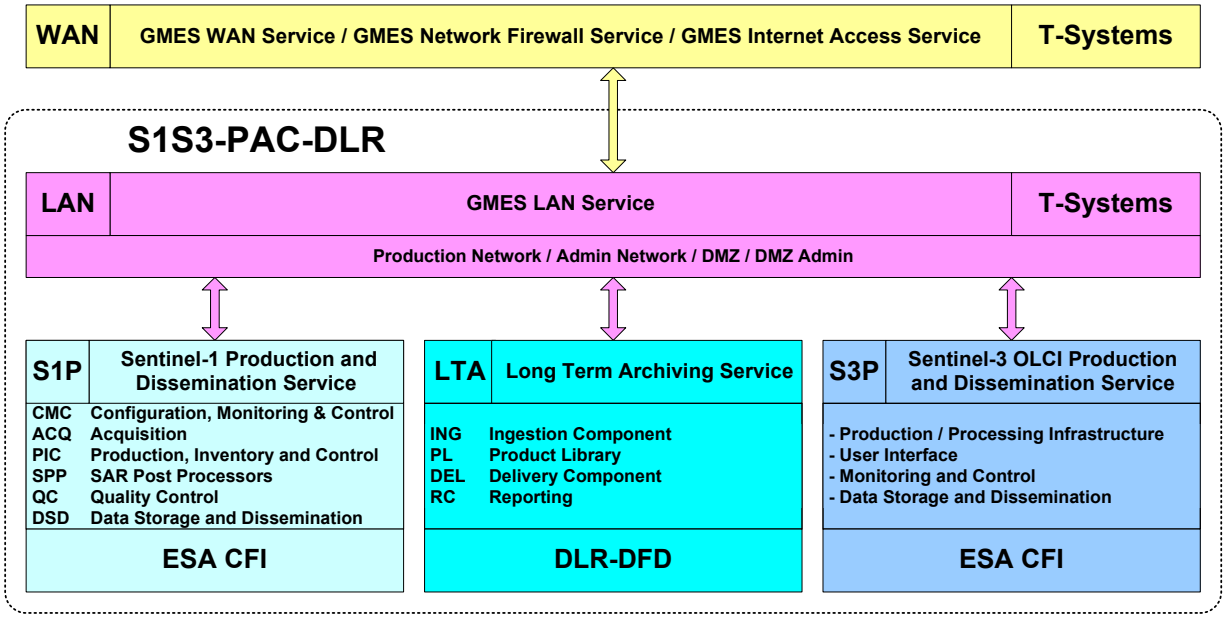


Fig. 2: Overall structure of the Sentinel-1 and Sentinel-3-OLCI PACs at DLR-DFD

4 Long-Term Archive service of DLR-PAC

In the scope of the Sentinel-PAC project, DLR's existing data archive (the so-called German Satellite Data Archive) was significantly enhanced in the year 2013. DLR has a robotic storage library available that will is to manage about 10,000 storage slots (Fig. 3). The storage tape technology which is utilized is the T10000C tape drive standard of Oracle through which it is possible to store up to 5 TB of data natively on one single cartridge. In its final configuration the DLR tape archive is able to host 50 PB (50 Million GB) of data.

Fig. 3: DLR's StorageTek SL 8500 Robotic Tape Library

4.1 Data Ingestion into the LTA

In the planning phase of the project (in the year 2012) the following Sentinel data capacity requirements were expected:

- approx. 2150 TB per year for Sentinel-1 data products (6 TB per day),
- approx. 600 TB per year for Sentinel-3 OLCI data products (1.6 TB per day).

In the course of the project it became apparent that these numbers had to be adjusted. The Sentinel-1A and Sentinel-1B satellite pair is already supplying over 10,000 GB (10 TB) and more than 4700 products per day (numbers from end of 2017). The daily Sentinel-1 data volume that needs to be archived in the DLR-PAC LTA is expected to further increase in the coming years up to 16 TB per day.
Regarding Sentinel-3 data the data volume is below the presumed amount. Daily, around 950 products are being processed based on sensor data from the OLCI instrument on Sentinel-3A, over 330 GB of associated data are stored in the Long-Term Archive of the DLR-PAC. These numbers will increase when Sentinel-3B starts acquiring data after its launch in April 2018.
In the first three years of operations the DLR-PAC has already processed and archived over 3.8 million datasets. They represent a data volume exceeding 6500 terabytes (6.3 petabytes). Table 1 provides a summary of the archived data in the LTA, whereas figures 4 and 5 show the archived Sentinel-1 and Sentinel-3 OLCI datasets over the course of the project.
Despite this enormous amount of data, the DLR-PAC is configured to additionally reprocess existing Sentinel data at any time with the latest algorithms to generate useful new geo-temporal data sets.

Tab. 1: Summary of Data archived into the LTA until end of 2017

Satellite mission	Number of products	Data Volume in GB	Data Volume in TB	Products / day	GB / day
Sentinel-1	3,181,731	6,451,848	6,301	2841	5761
Sentinel-3 OLCI	642,009	227,326	222	940	333
Total data ingestion	**3,823,740**	**6,679,174**	**6,523**		

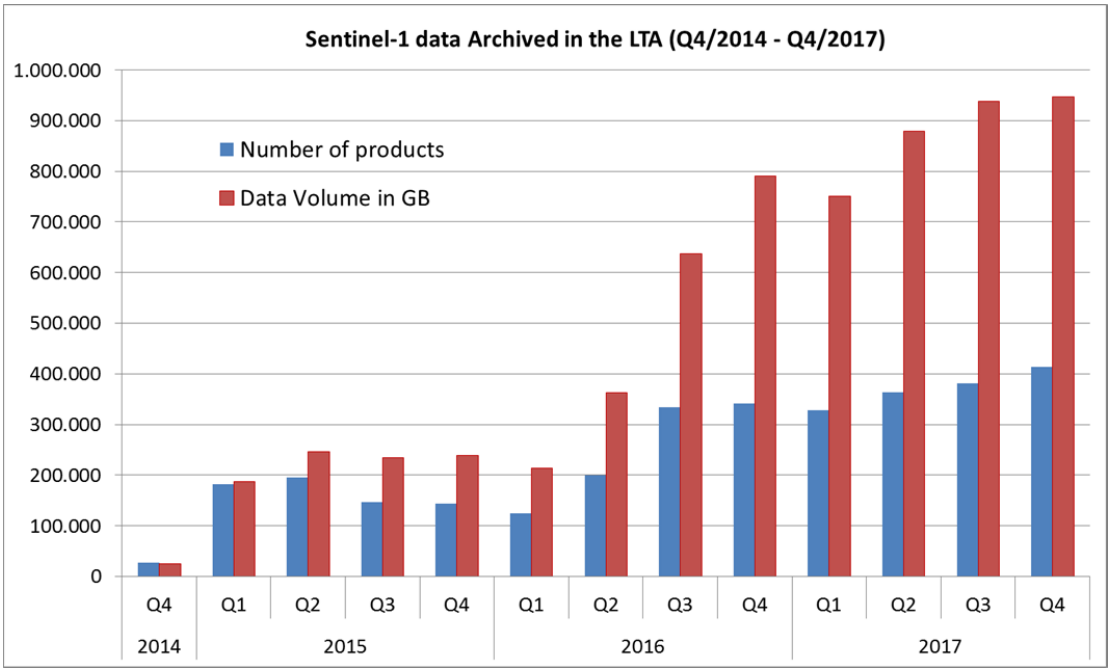


Fig. 4: Sentinel-1 data archived in the DLR-PAC LTA (per quarter, Q4/2014 – Q4/2017)

The reason for the major increase of the data volume in Q3/2016 is the start of archiving of Sentinel-1B data during this period. Further data volume increases occurred due to start of data reception/transmission via the EDRS-A system (European Data Relay Satellite) from November 2016 (Sentinel-1A) respectively March 2017 (Sentinel-1B) onwards.

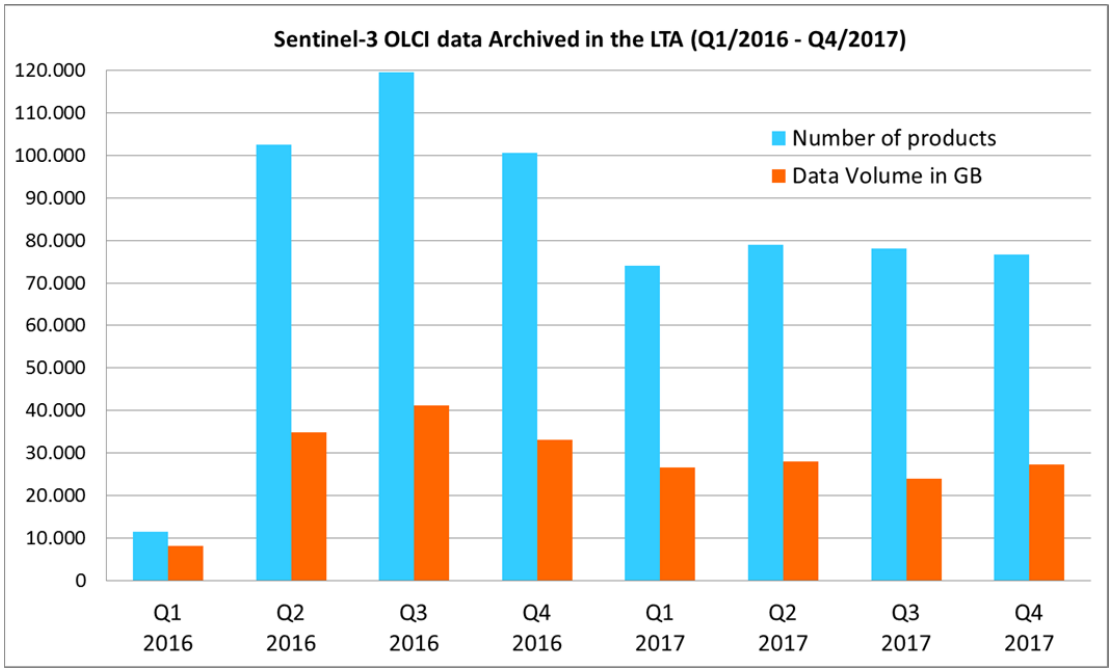


Fig. 5: Sentinel-3 OLCI data archived in the DLR-PAC LTA (per quarter, Q1/2016 – Q4/2017)

Fig. 6 shows the current and expected data volume growth in DLR's Long-Term Archive for the years 2014 until 2025.

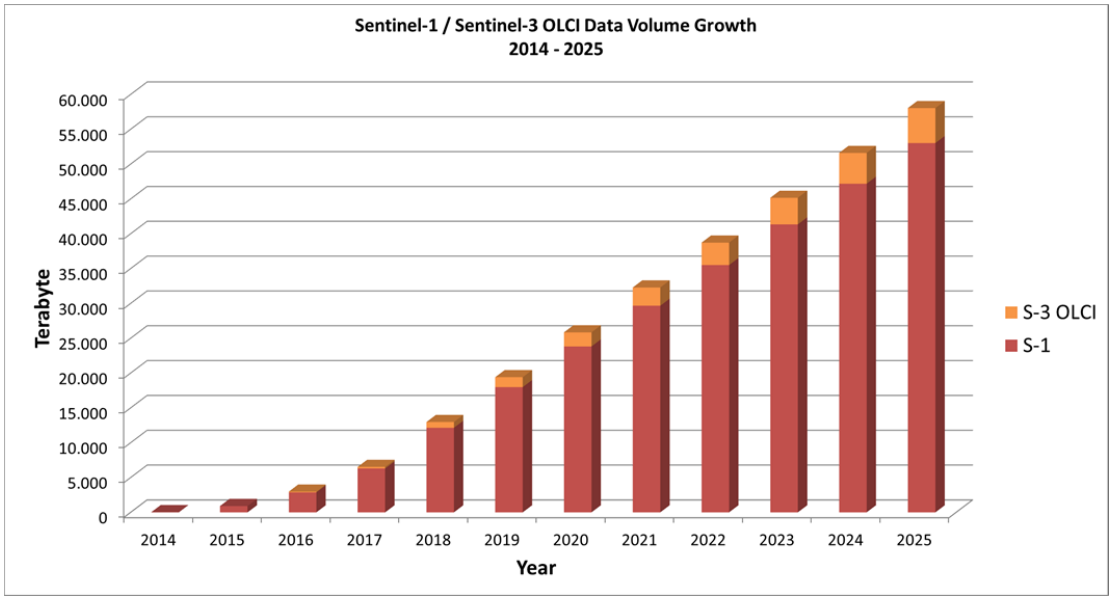


Fig. 6: Current and expected Sentinel-1 and Sentinel-3 OLCI Data Volume growth at DLR-PAC LTA in the timeframe 2014 – 2025

Fig. 7 summarizes the distribution of the archived Sentinel-1 and Sentinel-3 OLCI products with regard to the product types.

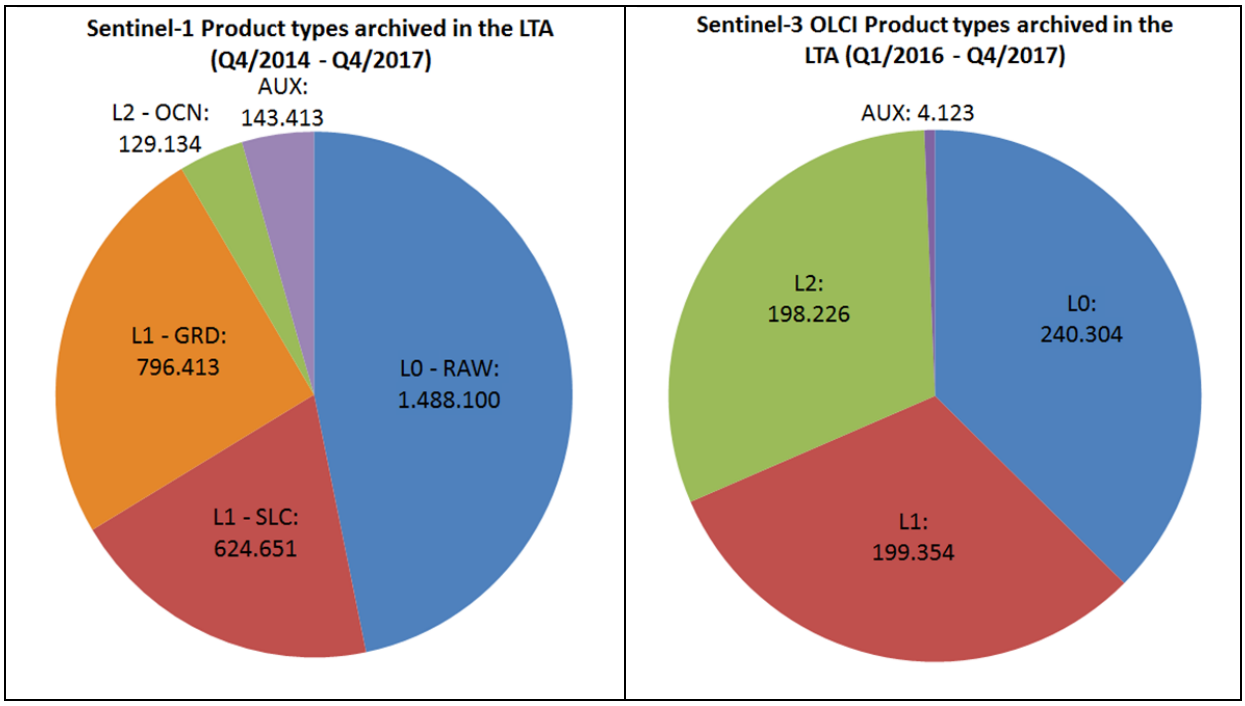


Fig. 7: Distribution of the archived Sentinel-1 and Sentinel-3 OLCI Data Product types until end of 2017

4.2 Data Retrieval from the LTA

The data retrieval from the LTA for Sentinel-1 data has started in June 2015. Operational data retrieval from the LTA for Sentinel-3 OLCI data has started in April 2016, only two months after the launch of the Sentinel-3A satellite.

The total amount of retrieved datasets is more than 70,000, whereas the data volume has reached almost 50 TB. The following table summarizes the data retrievals from the LTA until end of 2017.

Tab. 2: Summary of Data retrieved from the LTA until end of 2017

Satellite mission	Number of products	Data Volume in GB	Data Volume in TB
Sentinel-1	69,847	49,382	48.22
Sentinel-3 OLCI	2,665	895	0.87
Total data retrievals	**72,512**	**50,277**	**49.10**

The amount of data retrievals requests from the LTA, especially for Sentinel-1 data is expected to increase significantly from Q1/2018 onwards. Reason for this assumption is that ESA will no

longer be able to hold all the Sentinel-1 datasets on their Short Term Archive. The data deletion is planned to be started with Sentinel-1 L0 data older than 1 year. Sentinel-1 L1 data will follow on a later stage; Sentinel-3 data will be affected afterwards.
The DLR-PAC LTA will then ensure that all Sentinel datasets remain accessible for the data users also in a long run perspective.

5 Data access support for the CMEMS project

On request of ESA DLR-PAC is providing a dedicated data access support for the Copernicus Service Project CMEMS (Copernicus Marine Environment Monitoring Service, http://marine.copernicus.eu/). DLR-PAC has established a FTP server that is used to provide access to Sentinel-1 data in a timely manner for the members of the CMEMS project since September 2014. CMEMS is the successor project to the Copernicus Service Project MyOcean since May 2015.
CMEMS focuses on the provision of coastal and marine environment applications in the fields of marine safety, marine resources, and weather, climate and seasonal forecasting. A wide range of data is delivered by the service (e.g. temperature, salinity, sea level, currents, wind and sea ice). An important application is e.g. the derivation of sea ice drift observations and ice deformation (divergence, shear, vorticity) information (PEDERSEN & SALDO 2016).
Sentinel-1 products with two different "product Timeliness Categories" are provided via the FTP server: so-called "Fast-24h" products (share of around 70%) and "NRT-3h" products (30%). The number of delivered Sentinel-1 products and the corresponding data volume has significantly increased in the last three years. Tab. 3 and figures 8 and 9 summarize these values. More than 420,000 products and more than 310 TB of data have been provided for the CMEMS project members. On a monthly average CMEMS has access to about 12,000 Sentinel-1 products (data volume around 9 TB).

Tab. 3: Summary of Sentinel-1 data provided for the CMEMS project until end of 2017

Year	Number of products	Data Volume in TB
2015	78,269	43.33
2016	158,285	98.49
2017	185,558	174.44
Total (2015-2017)	**422,112**	**316.26**

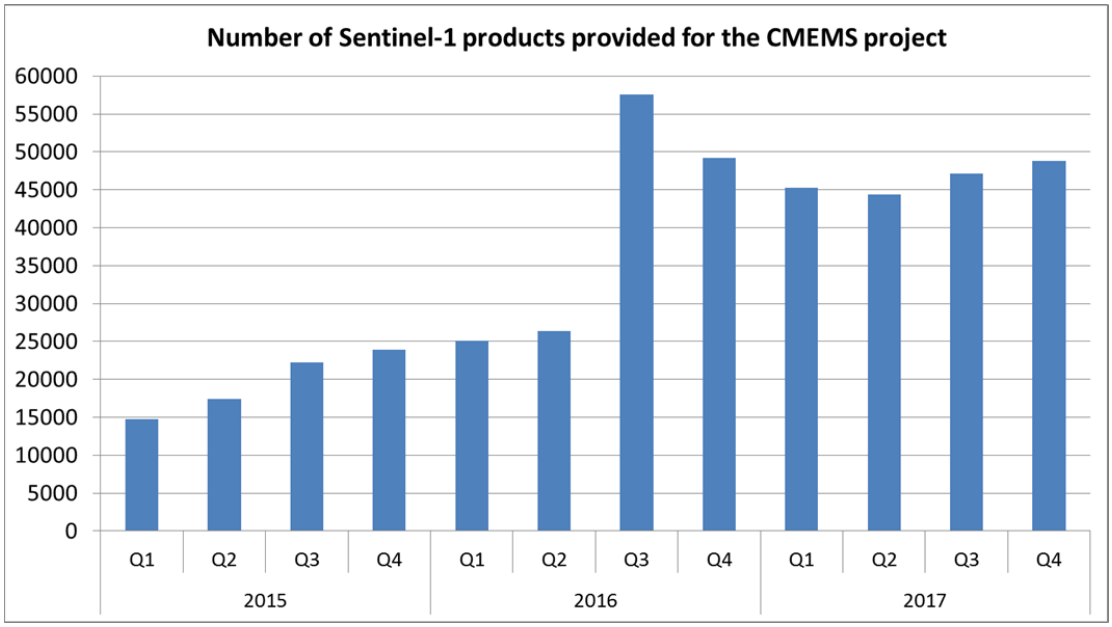


Fig. 8: Number of Sentinel-1 products provided for the CMEMS project (per quarter, 2015-2017)

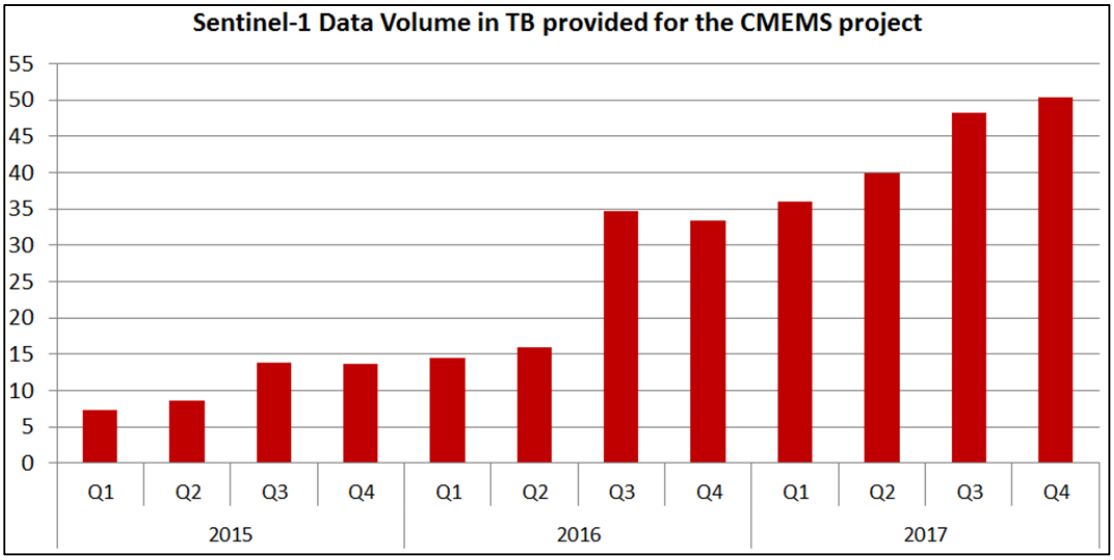


Fig. 9: Sentinel-1 Data Volume in TB provided for the CMEMS project (per quarter, 2015-2017)

6 References

ASCHBACHER, J. & MILAGRO-PÉREZ, M.P., 2012: The European Earth monitoring (GMES) programme: Status and perspectives. Remote Sensing of Environment, **120**, 3–8.

DONLON, C. J., BERRUTI, B., BUONGIORNO, A., FERREIRA, M.-H., FÉMÉNIAS, P., FRERICK, J., GORYL, P., KLEIN, U., LAUR, H., MAVROCORDATOS, C., NIEKE, J., REBHAN, H., SEITZ, B., STROEDE, J. & SCIARRA, R., 2012: The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. Remote Sensing of Environment, **120**, 37–57.

HAHMANN, T., WEBER, H., DIEDRICH, E. & TORRES, G., 2013: Sentinel-1 and Sentinel-3-OLCI PAC at DLR. Proceedings of ESA Living Planet Symposium, 9-13 September, Edinburgh, UK.

PEDERSEN, L.T. & SALDO, R., 2016: Sentinel-1 provides ice drift observations for Copernicus Marine Environment Monitoring Service (CMEMS). Proceedings of ESA Living Planet Symposium, 9-13 May, Prague, Czech Republic.

POTIN, P., ROSICH, B., MIRANDA, N., GRIMONT, P., BARGELLINI, P., MONJOUX, E., MARTIN, J., DESNOS, Y.-L., ROEDER, J., SHURMER, I., O'CONNELL, A., TORRES, R., KRASSENBURG, M. & GRATADOUR, J.-B., 2017: Sentinel-1 Mission Status. Proceedings of IGARSS 2017. 23-28 July 2017, Fort Worth, Texas, USA, 5525-5528.

SANTELLA, C., NOGUEIRA LODDO, C., FOURNIER-SICRE, V., SANTACESARIA, V., LØVSTAD, J., WILSON, H., BUONGIORNO, A., SCIARRA, R., MONJOUX, E., GORYL, P., FEMENIAS, P., SUNDA, M. & NIEZETTE, M., 2013: Sentinel-3 Payload Data Ground Segment Overview. In: Proceedings of ESA Living Planet Symposium 2013, 9-13 September, Edinburgh, UK.

SCHREIER, G., DECH, S., DIEDRICH, E., MAASS, H., MIKUSCH, E., SCHWARZ, E., WEBER, H. & HAHMANN, T., 2013: Ground Segment Facilities of the European Earth Observation Program Copernicus at the German Aerospace Center. Conference Proceedings, IAC-13, 64th International Astronautical Congress, 23-27 September, Beijing, China.

TORRES, R., SNOEIJ, P., GEUDTNER, D., BIBBY, D., DAVIDSON, M., ATTEMA, E., POTIN, P., ROMMEN, B., FLOURY, N., BROWN, M., NAVAS TRAVER, I., DEGHAYE, P., DUESMANN, B., ROSICH, B., MIRANDA, N., BRUNO, C., L'ABBATE, M., CROCI, R., PIETROPAOLO, A., HUCHLER, M. & ROSTAN, F., 2012: GMES Sentinel-1 mission. Remote Sensing of Environment, **120**, 9-24.

Automatische Extraktion von Fenstern in 3D Punktwolken mittels einer hierarchischen Methode

SVEN SCHNEIDER[1] & VOLKER COORS[1]

Zusammenfassung: Energetische Simulationen im urbanen und besonders im innerstädtischen Raum erfahren eine immer größere Nachfrage, z.B. um nachhaltige Stadtentwicklung zu betreiben. 3D Stadtmodelle sind hierfür gut geeignet. Diese können durch Simulationssoftware wie SimStadt genutzt werden, um Wärmebedarfsanalysen zu simulieren. Viele amtliche Daten sind jedoch nicht mit den nötigen Kennwerten ausgestattet. Zu diesen gehören u.a. die Stockwerksanzahl und das Fenster-Fassaden-Verhältnis von Gebäuden. Mit der hier vorgestellten Methode lassen sich diese Kennwerte über die detektierten Fenster ableiten. Zum Detektieren der Fenster werden zunächst Fassaden (Ebenen) in Punktwolken gesucht. Danach folgen das Invertieren der Fassadenpunktwolke und die Konturverfolgung der Fenster. Zuletzt werden Konturen gelabelt und die Flächen dieser anhand ihrer kleinsten umschreibenden Rechtecke berechnet. Ergebnisse für die Fensterdetektion liegen, je nach Fassade, bei einer Treffgenauigkeit von 80-90%. Mit Hilfe der Position und der Größe der Fenster lassen sich somit die entsprechenden Kennwerte für Simulationen ableiten.

1 Einleitung

Rund 40% des Energiebedarfs einer Stadt entsteht durch den Gebäudesektor. In privaten Haushalten, sind ca. 85% der Energie für den Heizbedarf notwendig (BMWI 2014). Durch die Energiewende wird es immer wichtiger, dass Neubauten energieeffizient gebaut werden und Eigentümer und Vermieter Empfehlungen zur Sanierung ihrer Gebäude erhalten, um den CO2 Ausstoß nachhaltig zu reduzieren (BMWI 2011). In diesem Zuge sind Simulationswerkzeuge wie z.B. SimStadt (NOUVEL et al. 2015) geeignete Werkzeuge um den Energie- und Wärmebedarf solcher baulicher Maßnahmen zu simulieren.

Amtliche Daten sind oft nicht mit allen relevanten Sachdaten und Kenngrößen versehen, die für solche Simulationen notwendig sind (BILJECKI et al. 2015; NOUVEL et al. 2017). In vielen Fällen besteht daher die Notwendigkeit zur Anreicherung von Datenbeständen mit Sachdaten aus anderen Quellen. Wichtige Sachdaten zur Simulation, wie Stockwerksanzahl oder dem Fenster-Fassaden-Verhältnis, sind oft nicht oder nur teilweise in Datenbanken erfasst. Durch Annahme eines quartierübergreifenden Mittelwerts (z.B. für das Fenster-Fassaden-Verhältnis) werden in Simulationen akzeptable Ergebnisse erreicht, jedoch ist die Aussagekraft für kleine Einheiten (z.B. einzelne Gebäude), mit relativ großen Unsicherheiten behaftet.

Für eine möglichst realistische energetische Gebäudebetrachtung ist das Verhältnis von Fensterfläche zur Gesamtgebäudefläche relevant, da Fenster und Fassade in der Regel erheblich unterschiedliche Wärmedurchgangskoeffizienten (U-Werte) aufweisen (MONIEN et al. 2016). Die Kenntnis der tatsächlichen Fensterflächenanteile einer Fassade liefern daher genauere Ergebnisse bei einer Wärmebedarfsanalyse (MONIEN et al. 2017).

[1] Hochschule für Technik Stuttgart, Zentrum für Geodäsie und Geoinformatik, Schellingstr. 24, D-70174 Stuttgart, E-Mail: [sven.schneider, volker.coors]@hft-stuttgart.de

Durch automatisierte Auswerteverfahren von photogrammetrisch gewonnenen Daten, beispielsweise Schrägluftbilder aber auch 3D Punktwolken, können relevante Sachdaten ergänzt werden. Daher ist es möglich, Stadtmodelle in einem niedrigen LOD (1 oder 2) zu nutzen obwohl aus diesen geometrischen Daten (i.d.R. im CityGML vorliegend) die nötigen Kennwerte i.d.R. nicht abgeleitet werden können.
Zu dem oben beschrieben Zweck wird in diesem Beitrag eine hierarchische Methode zur Extraktion von Fenstern an Gebäudefassaden präsentiert. Gebäudemodelle können mit dem vorgestellten Verfahren mit relevanten Sachdaten angereichert werden und für Simulationen genutzt werden.
Der Artikel gliedert sich folgendermaßen: Sektion 2 beschreibt die Problemstellung dieser Arbeit und gibt Kontext zu bestehenden Methoden; Sektion 3 beschreibt die zu Grunde liegenden Daten; Sektion 4 beschreibt die angewendete Methode und zeigt erste Zwischenergebnisse für jeden Verarbeitungsschritt; Sektion 5 diskutiert und fasst die Arbeit zusammen.

2 Problemstellung

Da Stadtmodelle mit detaillierten Informationen über Fenster und Stockwerksanzahl oft nur teilweise oder auch gar nicht vorliegen, besteht eine große Notwenigkeit (z.B. für Simulationen) diese Informationen auf anderen Wegen zu ergänzen. Neben Schrägluftbildern stellen auch sogenannte Mobile Mapping Verfahren mögliche Alternativen zur Verfügung. Mobile Mapping Daten werden in der Regel durch Fahrten mit einem Fahrzeug und entsprechenden Laserscannern durchgeführt (siehe Sektion 3).
Die Herausforderung, Fenster in Fassaden zu detektieren, ist aus diversen Gründen schwierig:

(i) Oft sind Fenster gar nicht oder nur teilweise in Punktwolken erfasst, da der Laserstrahl die Fensterscheibe durchdringt oder durch spiegelnde Reflexion vom Detektor weg reflektiert wird.
(ii) Fenster sind i.d.R. von der Fassade leicht nach hinten versetzt (Abbildung 1). Dieser Abstand variiert in den betrachteten Punktwolken je nach Gebäudetyp und Baujahr von ca. 15 cm, bis zu 35 cm.
(iii) Ähnlich wie der Versatz der Fenster, variiert auch die Vielfalt der Fensterformen, daher funktioniert eine Detektion über Templates nur mit relativ hohem Aufwand.
(iv) Rollläden stellen eine zusätzliche Herausforderung dar, da sie viele ungewissen Zustände haben können (ganz offen, geschlossen, halb offen und alles dazwischen).
(v) Letztlich kommen zu diesen speziellen Herausforderungen noch die üblichen messtypischen Einflussfaktoren hinzu, z.B. Verdeckungen von Fassaden und Fensterteilen durch Vegetation, parkende Fahrzeuge und / oder Laternenmasten.

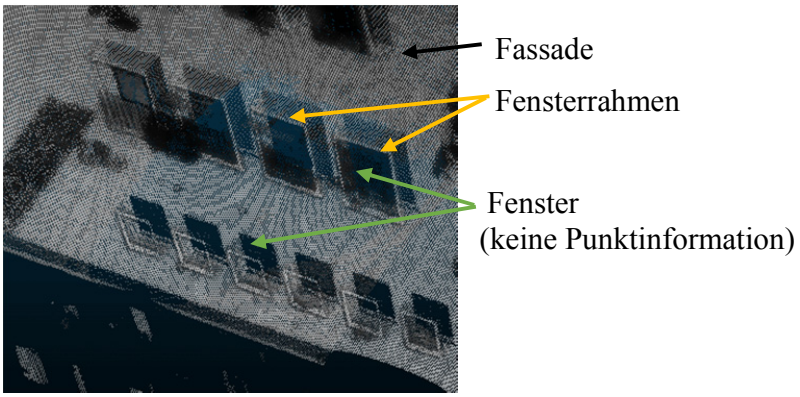


Abb. 1: Gebäudefassade betrachtet von der Innenseite. Die Fensterrahmen sind gut zu erkennen und der Versatz der Fenster zur Fassade. Hier liegen die Fenster vor der Fassade, da dieser Screenshot der Punktwolke von der Innenseite gemacht wurde

2.1 Lösungsansätze aus der Literatur

Einige der bekannten Ansätze zur Detektion von Fenster in Punktwolken nutzen bildgebende Verfahren ausschließlich oder unterstützend, d.h. dass Fassaden mit RGB-D Sensoren aufgenommen oder zusätzlich zu Laserscannern mit Kameras fotografiert wurden (z.B. ALI et al. 2007; NEUHAUSEN et al. 2018; TRINH et al. 2009; ZHANG & ZAKHOR 2014). Verfahren, die unabhängig von Kamerabildern sind und ausschließlich auf Laserscanndaten basieren sind in den letzten Jahren gestiegen. Die Methode von TUTTAS UND STILLA (2013) erzielt Resultate mit einer Klassifiziergenauigkeit von knapp 100% für Fassaden mit gleichmäßigen Strukturen. Für Häuserblocks in Wohngebieten, mit abwechselnden und kurzen Segmenten von Fenster, erzielt dieses Verfahren auf Grund der genutzten Fourier Transformation schlechte Resultate. Andere Methoden nutzen Verfahren des maschinellen Lernens um Fensterformen zu lernen und in Fassaden zu detektieren (NGUATEM et al. 2014). Dadurch ist es einerseits möglich Fenster mit unterschiedlichen Formen zu detektieren (z.B. mit Rundungen), andererseits verliert man auf diesem Wege Flexibilität, da nur gelernte Muster erkannt werden können. Das Erstellen von gelabelten Trainingsdaten ist hier ebenfalls problematisch und limitierend.

3 3D-Laserpunktwolken

3.1 Erfassung der Daten

Laserscan-Daten wurden im Rahmen des WeBest Projekts (HFT Stuttgart; MONIEN et al. 2016) von der Firma IGI im Auftrag von EUROSENSE akquiriert. Es wurden vier repräsentative Quartiere mit insgesamt etwa 500 Wohngebäuden in unterschiedlichen Baualtersklassen der Stadt Essen ausgewählt. Bei der Messfahrt wurde auf eine konstant niedrige Geschwindigkeit (< 30 km/h) geachtet. Der eingesetzte Laser Z+F Profiler (Genauigkeit ca. <1 mm bei einer Distanz von 25 m) ermöglichte die Aufnahme eines 360° großen Field-of-Views mit einer Scanrate von bis zu 1 Million Punkten pro Sekunde. Gleichzeitig kam eine Novatel GPS-Antenne L1/L2 Model 512 und eine ApplanixIMU 200Hz zum Einsatz (Aufbau siehe Abbildung 2). Da der Sensor

relativ weit oben befestigt war, führte dieses Verfahren zu sehr hochaufgelösten, georeferenzierten 3D Punktwolken mit relativ wenig Verdeckungen auf den Fassaden. Dennoch konnte nicht verhindert werden, dass auf einigen Fassaden Lücken durch geparkte Fahrzeuge, Masten von Straßenlaternen und Bäumen in den Punktwolken vorzufinden waren.

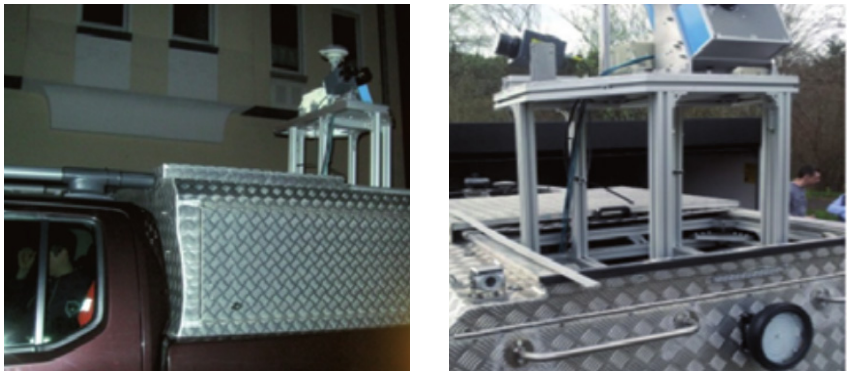


Abb. 2: Mobile Mapping Plattform mit GPS Antenne, Laserscanner, IMU und Thermographie- Kamera (in dieser Arbeit nicht verwendet). Quelle: MONIEN et al. (2016)

3.2 Merkmale der Punktwolken

Die Punktwolken sind hochauflösend und geben die Umgebung sehr detailgetreu wieder. Die Auflösung variiert mit der Distanz zum Laserscanner, d.h. die höchste Auflösung wird im Nahbereich des Lasers erreicht (z.B. der Straße, 2400-2800 Punkte pro qm) und die niedrigste Auflösung wird an den höchsten und weitest entfernten Punkten erreicht (z.B. das obere Ende der Fassade, 450-800 Punkte pro qm). Verglichen mit typischen Laserpunktwolken aus der Fernerkundung (z.B. durch amtliche Befliegungen), welche oft Werte von ca. 4 Punkten pro qm erreichen, sind die hier erreichten Auflösungen selbst am niedrigen Ende noch extrem hochauflösend. Abbildung 3 (oben) zeigt die Verteilung der Punktdichte in Abhängigkeit des Abstands und eine Fassade der Originalpunktwolke mit unterschiedlichen Intensitätswerten (unten). Punktwolken sind im LAS Format abgespeichert und die Koordinaten sind im DHDN GK Format vorhanden. Eine typische Punktwolkendatei hat zwischen 10 und 14 Millionen Punkte und eine Dateigröße von 400-500 Megabyte. Da diese große Anzahl an Punkten zu einer sehr langsamen Verarbeitung führen würde, wird die Punktwolke im Schritt der Datenvorverarbeitung zuerst in der Auflösung reduziert (Sektion 4).

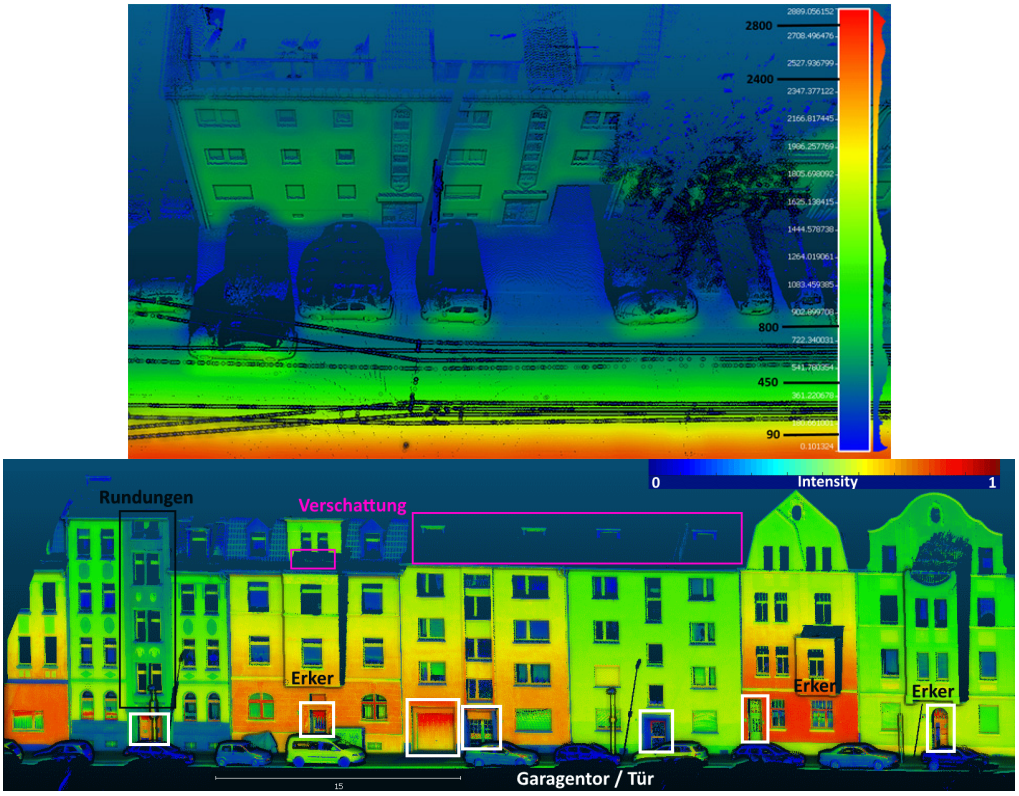


Abb. 3: Ausschnitt einer Punktwolke in Falschfarbendarstellung. Die Punktdichte pro qm ist im oberen Bild zu sehen. Intensitätswerte der reflektierten Strahlen sind unten zu sehen. Die Annotationen wie Erker, Garagentor, etc. sind für spätere Diskussionen relevant

4 Methode und Resultate

Die Methode ist in fünf große Schritte aufgeteilt welche in Abbildung 4 schematisch dargestellt sind. Die einzelnen Verarbeitungsschritte werden in den folgenden Sektionen detailliert beschrieben.

Die Methode wurde in C++ implementiert unter Verwendung gängiger Software Bibliotheken wie z.B. die Point Cloud Library (PCL; RUSU & COUSINS 2011), welche die Grundlage für viele der im Folgenden verwendeten Methoden darstellt.

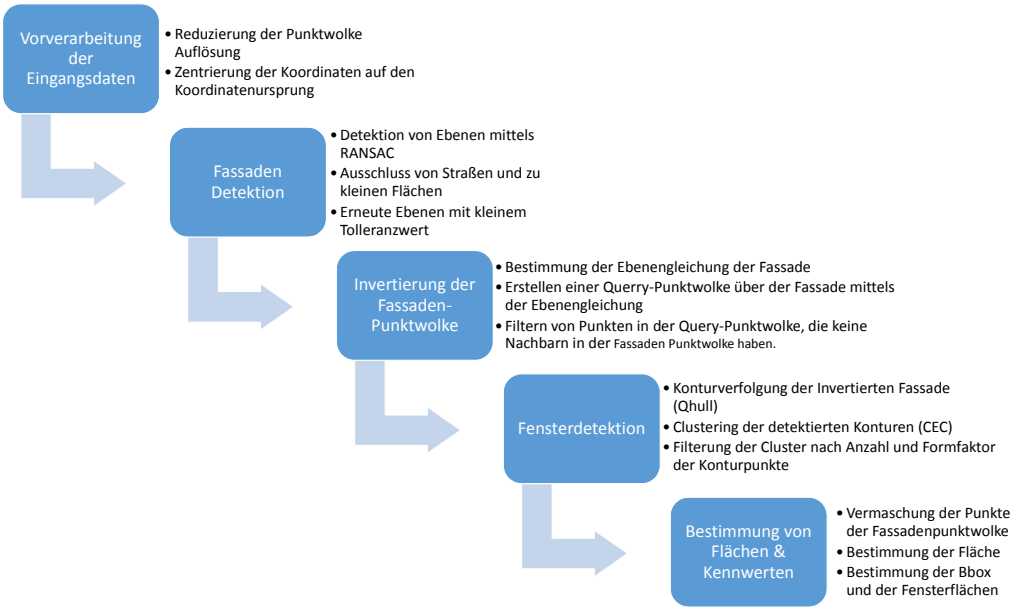


Abb. 4: Übersicht der Prozesskette zur Bestimmung der Fenster in Fassaden

4.1 Vorverarbeitung der Punktwolke

4.1.1 Zentrierung der Punktwolkenkoordinaten

Die Punktwolken im LAS Format sind im als Datentyp Double gespeichert, allerdings unterstützt die PCL nur den Float Datentyp. Würden die LAS Daten per Typumwandlung direkt in ein PCL Punktwolkenobjekt gespeichert werden, würden signifikante Stellen und Präzision verloren gehen. Dies würde zu Artefakten führen, welches für die weitere Verarbeitung problematisch wäre (Abbildung 5).

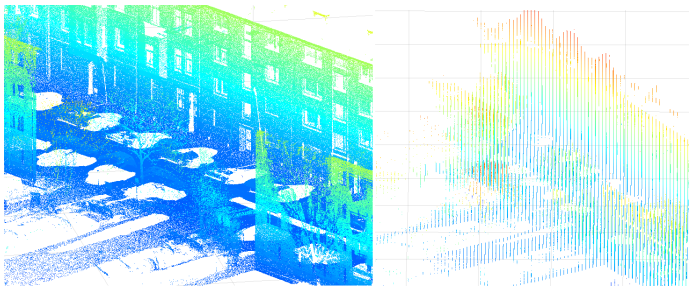


Abb. 5: Darstellung der Artefakte (nur rechts) der Punktwolke im Format Double (links, original) und Float (rechts, konvertiere Punktwolke) nach Durchführen einer einfachen Typkonvertierung.

Um diese Artefakte zu vermeiden, wird die Punktwolke zuerst auf den Koordinatenursprung zentriert, d.h. nach der Zentrierung hat die kleineste Koordinate in x,y,z (in Kartesischen Koordinaten), jeweils den Wert Null. Dadurch werden bis zu sieben Stellen vor dem Dezimalpunkt gespart, welche für die Nachkommastellen und entsprechend die Genauigkeit entscheidend sind.

4.1.2 Reduzierung der Auflösung

In diesem Schritt wird die sehr hohe Auflösung der Punktwolke reduziert. Es ist von Vorteil sehr detaillierte Daten zu haben, jedoch ist die Datenmenge für eine zügige Bearbeitung zu groß. Die Punktauflösung wird durch ein Voxel-Grid Filter der PCL reduziert, in dem die Punktwolke in viel größere Voxel aufgeteilt, und der Mittelwerte jedes Voxel als gefilterter Punkt übernommen wird. Werte von 4-6 cm³ per Voxel ergaben für die verwendeten Datensätze eine deutliche Reduzierung (von ca. 13 Mio. auf ca. 2 Mio. Punkte) bei sehr geringer Reduzierung der Qualität. Die reduzierte Punktwolke ist selbst nach der Punktreduzierung immer noch sehr hochauflösend und ausreichend für die weitere Verarbeitung.

4.2 Detektion von Fassaden in der Punktwolke

Zur Detektion der Fassade wird angenommen, dass Fassaden näherungsweise flache Ebenen sind. Diese Annahme ist für den Großteil der Fassaden legitim und führt zu guten Ergebnissen.
Ebenen in der Punktwolke werden mittels RANdom SAmple Consensus (RANSAC) Algorithmus in der Punktwolke detektiert. Ein Toleranzwert von 40 cm wird für die Detektion angenommen. D.h. alle Punkte, die innerhalb dieses Toleranzwertes sind, sind Punkte auf der gesuchten Ebene (Abbildung 6). Es werden jedoch auch andere Ebenen als die gewünschten Fassaden gefunden (z.B. Straßeneben). Diese werden durch ermitteln des Normalvektors der Ebene (besonders nützlich für die Straßenebene), als auch über die Anzahl der Punkte in der Ebene entfernt. D.h. besteht eine Eben aus weniger als 15 % der gesamten Punktwolke (hier ist die bereits reduzierten Punktwolke gemeint), wird die Ebene nicht als Fassade klassifiziert und von der weiteren Auswertung ausgeschlossen.

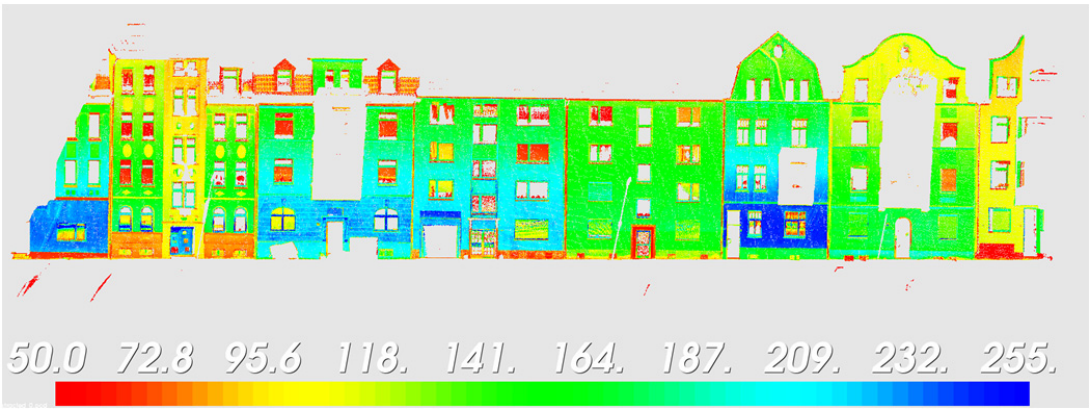


Abb. 6: Extrahierte Fassade durch Ebenen Detektion mittels RANSAC (Toleranz 40 cm). Viele Fenster, aber nicht alle, haben eine Punkinformation (meist rot). Die Farbskala entspricht der reflektierten Intensität des Laserpulses

Anschließend wird eine neue Ebenendetektion für die restlichen Punkte in der Punktwolke durchgeführt. Hier wird der Toleranzwert für die Ebenendetektion auf 15 cm reduziert. Dies führt dazu, dass nun nahezu alle Fenster ohne Punktinformation sind (Abbildung 7). In diesem Schritt sind die Fenster im Grunde bereits detektiert, dies wird durch Visualisierungen dieses Schrittes deutlich. Da die Fenster zum jetzigen Zeitpunkt allerding durch Absenz von Information gekennzeichnet sind, werden weitere Verarbeitungsschritte notwendig.

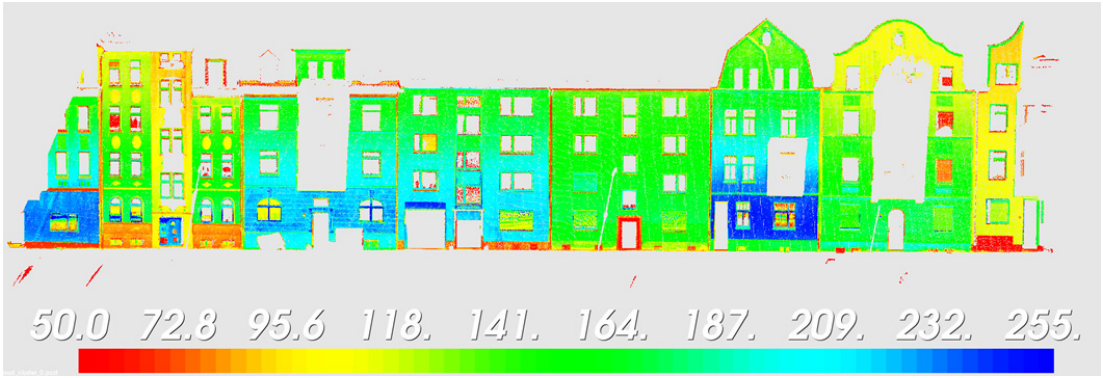


Abb. 7: Zweite Ebenen Detektion mit Toleranz von 15 cm. Fenster sind nun überwiegend ohne Punktinformation. Die Farbskala entspricht der reflektierten Intensität des Laserpulses

Es gibt jedoch Einschränkungen bei der Detektion der Fassadenebenen. Dies ist zum einen die eingangs beschriebene Annahme, dass Fassaden in einer Ebene liegen müssen, welche mit der Ebengleichung ($Ax + By + Cz + D = 0$) beschrieben werden kann. Eine weitere Einschränkung dieser Methode ist, dass hervorstehende Fassadenteile (z.B. Erker) nicht erkannt werden, da sie nicht innerhalb der Toleranzwerte des RANSAC Verfahrens für die Ebenendetektion liegen. Dies kann jedoch relativ leicht behoben werden, indem eine Ebenedetektion in einem Abstand Δd vor der Fassade durchgeführt wird (wobei $\Delta d >$ RANSAC Toleranzwert ist).

4.3 Invertierung der Fassadenpunktwolke

Im vorherigen Schritt wurden die Fenster durch fehlende Informationen der Fenster quasi detektiert. Um konkrete Punktinformationen über die Fenster zu erhalten, wird die Fassadenpunktwolke nun invertiert und folgt damit konzeptionell dem Ansatz zum Invertieren eines Binärbildes. Zum Invertieren wird eine gleichmäßige Struktur aus Punkten über die Fassade gelegt (hier im Abstand von 25 cm), nachdem die Ebenengleichung der Fassade über ein RANSAC Verfahren ermittelt wurde. Die Güte der Fensterdetektion hängt von der Auflösung der Punkte in der Gitterpunktwolke ab: je hochauflösender das Gitter ist, desto genauer werden die Fenster abgebildet. Die Punktwolke wird in die Datenstruktur eines Kd-Baums überführt, welche für die folgende Nachbarschaftssuche notwendig ist. Für jeden Punkt in der Gitterpunktwolke, wird an der gleichen Position in der Fassadenpunktwolke (und dessen unmittelbaren Umgebung), eine Suche nach Fassadenpunkten gestartet. Wird ein Fassadenpunkt in der Fassadenpunktwolke in diesem Suchbereich gefunden, so wird dieser Punkt aus der Gitterpunktwolke entfernt; andernfalls wird der Punkt in der Gitterpunktwolke gespeichert bzw. nicht eliminiert (Abbildung 8). Dies führt dazu, dass hauptsächlich die Fenster übrigbleiben und Fassadenpunkte eliminiert werden. Ein wichtiger Parameter in diesem Verarbeitungsschritt ist der Suchradius für die Nachbarschaftssuche. Für die in dieser Studie getesteten Punktwolken wurde ein Radius mit r = 15 cm genutzt.

Abb. 8: Invertierte Punktwolke durch die Gitterpunktwolke. Außerhalb der Fassade sind nun viele Punkte (oberhalb und unterhalb), die keine Fenster darstellen. Fenster haben nun Punktinformationen, die Fassade ist im Grunde eliminiert

4.4 Fensterdetektion mittels Konturverfolgung

In diesem Schritt wird die invertierte Fassade als Eingangssignal für die Detektion der konkaven Hülle der Fenster und der Fassade genutzt, welche der Kontur eines Objektes (Fenster) entspricht (Abbildung 9). Algorithmen hierfür werden ebenfalls aus der PCL genutzt, welche auf den Quickhull Algorithmus (Qhull, BARBER et al. 1996) zurückgreift. Ein wichtiger Parameter für die Hüllendetektion mit Qhull ist α. Dieser Parameter steuert wie detailliert die Kontur verfolgt wird. Kleine Werte (ca. 0.15), sorgen für relativ detaillierte Konturen, große Werte (> 0.5, mit einem Maximalwert von 1), sorgen für grobe, approximierte Konturen. Mit großen Parameterwerten bleiben mehr Fenster unentdeckt als bei kleinen Werten. Z.B. werden Fenster am Rande des Gebäudes mit zur Fassadenkontur gezählt, d.h. Details und kleinere Fenster gehen verloren. Ein zu kleiner Parameter sorgt dafür, dass zu viele kleine Details erkannt werden, z.B. vereinzelte Punkte (Rauschen, Äste von Bäumen und Laternenmaste). In der Implementierung werden jedoch im nächsten Verarbeitungsschritt kleine Objekte gefiltert. Daher ist eine leichte Übersegmentierung weniger problematisch als eine Untersegmentierung.

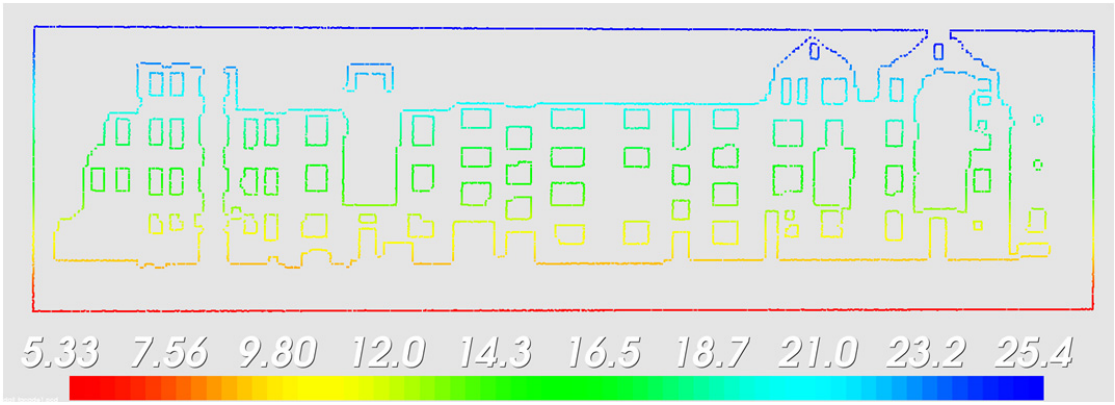


Abb. 9: Konturverfolgung der invertierten Gitterpunktwolke. Der äußere Rahmen ist in einem Nachverarbeitungsschritt leicht zu entfernen. Übrig bleiben die Umrisse der Fassade und der Fenster

4.5 Gruppierung von Punkten (Clustering)

Als vierter Schritt, wird ein „Conditional Euclidean Clustering“ Algorithmus (CEC, RUSU 2009) auf die detektierten Umrisse der Fenster angewendet. Dabei werden Objekte, die keine Fenster darstellen, weitestgehend entfernt und jedem Objekt wird ein eindeutiges Label zugeordnet (Abbildung 10).

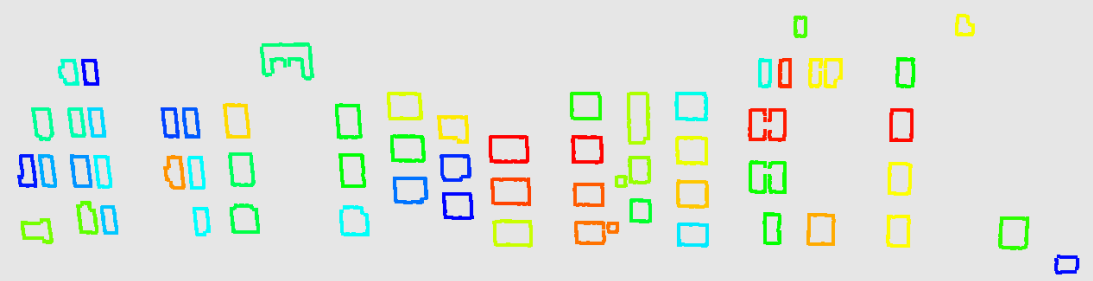


Abb. 10: Ergebnis des CEC Algorithmus mit zufällig eingefärbten Objekten / Fenstern.

Das Entfernen kleiner Objekte erfolgt über die Anzahl der Punkte per Cluster. Ein geeigneter Wert für diesen Parameter wurde empirisch ermittelt und liegt zwischen Werten von 35 und 180 Punkten, welche robuste Clustergrößen mit unterschiedlichen Punktwolken liefern. Sollten sich die Eingangsauflösungen der Punktwolken gravierend ändern, muss dieser Parameter ggf. angepasst werden. Weitere Parameter sind der Suchradius für den Normalenvektor suche (r = 7,5 cm) sowie „Cluster Toleranz“ (30 cm), also der maximale Abstand zwischen einem Punkt der bereits zu einem Cluster gehört und einem Kandidatenpunkt. Nur wenn der Abstand kleiner ist, werden die weiteren Konditionen überprüft und der Punkt entsprechend hinzugefügt oder abgewiesen. Ein abgewiesener Kandidatenpunkt kann aber zu einem späteren Zeitpunkt trotzdem zu diesem Cluster hinzugefügt werden, sollten alle Bedingungen zwischen dem Kandidatenpunkt und einem anderen Punkt des Clusters erfüllt sein. Die hier genutzten Bedingung zur Bewertung zum Aufnehmen eines Punktes waren (neben dem maximalen Abstand) die gleiche Orientierung der Flächennormalen.

4.6 Bestimmen von Fensterflächen und Kennwerten

In diesem Schritt werden die Fensterflächen und Fassadenflächen berechnet. Bei den Fassadenflächen ist es allerdings einfacher und unter Umständen genauer diese Information aus den Gebäudemodelldaten (also dem CityGML) abzuleiten.

Die Fensterflächen werden approximiert, in dem das kleinste umschreibende Rechteck (BoundingBox) für jedes Cluster aus Fensterumrisspunkten ermittelt wird. Dazu wird ebenfalls der Qhull Algorithmus aus der PCL verwendet. Die vier Eckpunkte des Fensters werden durch diesen Algorithmus bestimmt und die Fläche wird entsprechend über die Differenzen der Breiten- und Höhenkoordinaten berechnet. Durch diese Methode werden Fensterflächen mit Rundungen jedoch überbewertet, was der Einfachheit des Verfahrens geschuldet ist. Es kann jedoch angenommen werden, dass dies für die meisten Fassaden mit Fenstermischformen und für die meisten Anwendungsfälle, diese Diskrepanz keinen großen Einfluss auf die Ergebnisse hat.

Die Eckpunkte des detektierten Rechtecks können im Folgenden dann für weitere Berechnungen genutzt werden oder bei Bedarf auch als Polygonzug gespeichert werden (z.B. als CityGML), als Ergänzung zu einem bestehenden Gebäudemodell.

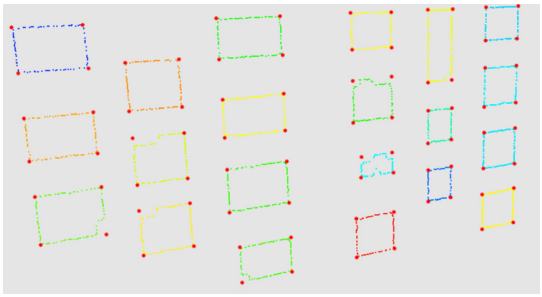


Abb. 11: Vergrößerter Teilausschnitt der Fenster mit Konturen mit überlagerten Eckpunkten (rote Punkte) der bestimmten BoundingBoxen. Unterschiedliche Farben der Fenster entsprechen unterschiedlichen Labels, die während des CEC bestimmt wurden

4.7 Auswertung

Die oben beschriebene Fassade (Abbildung 3) wurde zur ersten Evaluierung der Methode genutzt. Da eine Limitierung der Methode darin besteht, Fenster in hervorstehenden Objekten wie Erkern nicht erkennen zu können, werden Fenster, die sich an solchen Stellen befinden nicht in der Auswertung berücksichtigt. Lediglich die Fenster, die nach der Ebenendetektion noch sichtbar sind, werden berücksichtigt (in Magenta eingefärbt, Abbildung 12). Fenster, die mit einem weißen Punkt markiert sind, wurden richtig erkannt (vgl. Abbildung 10). Die quantitative Auswertung zeigt recht hohe Werte für die Treffgenauigkeit und den F1-score (2*TP / (TP+FP+TN+FN), siehe Tab. 1).

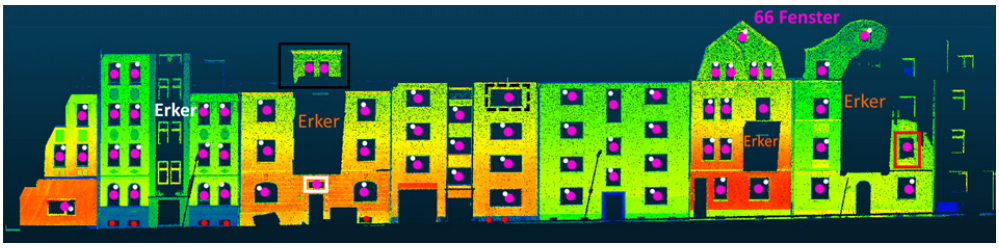


Abb. 12: Fassade mit markierten Fenstern (Magenta), welche für die Auswertung berücksichtigt werden. Rot markierte Fenster werden nicht erkannt, da sie unterhalb der minimalen Höhe zur Suche liegen

Tab. 1: Übersicht der Auswertung für die obere Fassade. Es wurde eine Treffgenauigkeit (Accuracy) von 87 % und ein F1-score von 92 % erreicht. Als „kein Fenster“ werden Türen bzw. Garagentore bezeichnet

		Tatsächlich	
		Fenster	**kein Fenster**
Prädiziert	**Fenster**	TP=60	FP=4
	kein Fenster	FN=6	TN=7

Die Auswertung zeigt auch, dass keine Türen oder Garagentore als Fenster erkannt wurden (7 TN). Dies wurde zum einen durch die Anzahl der erlaubten Punkte in einem Cluster beschränkt und zum anderen über die Höhe ausgeschlossen. Einige Fenster wurden jedoch nicht detektiert. So sind die Fenster im Verschattungsbereich (siehe schwarzes Rechteck in Abbildung 12) nicht detektiert worden, jedoch wurde der ganze obere Bereich der Fassade als Fenster detektiert was zu einer Fehldetektion führt (FP). Andere Fenster (schwarze, gestricheltes Rechteck und rotes Rechteck) wurden nicht erkannt, da durch Rauschen die Fensterkontur zur Fassadenkontur gezählt wurde und bei dem anderen Fenster die Fassade bereits lückenhaft detektiert wurde (womöglich durch Verschattung durch den Erker).
Die Methode wurde, mit den gleichen Parametern wie oben bereits beschrieben, an einer weiteren Fassade getestet. An dieser gab es keine Erker, jedoch Verschattungen durch Bäume (Abbildung 13). Punkte in Magenta markieren alle vom Algorithmus erkennbaren Fenster (unter der Annahme, dass es keine Verschattung gäbe). Schwarze Punkte in Abbildung 13 (oben) markieren die richtig detektieren Fenster, rote Punkte markieren Fenster, die nicht erkannt wurden und das schwarze Rechteck markiert eine grobe Fehldetektion (vgl. Abbildung 13 oben und unten).

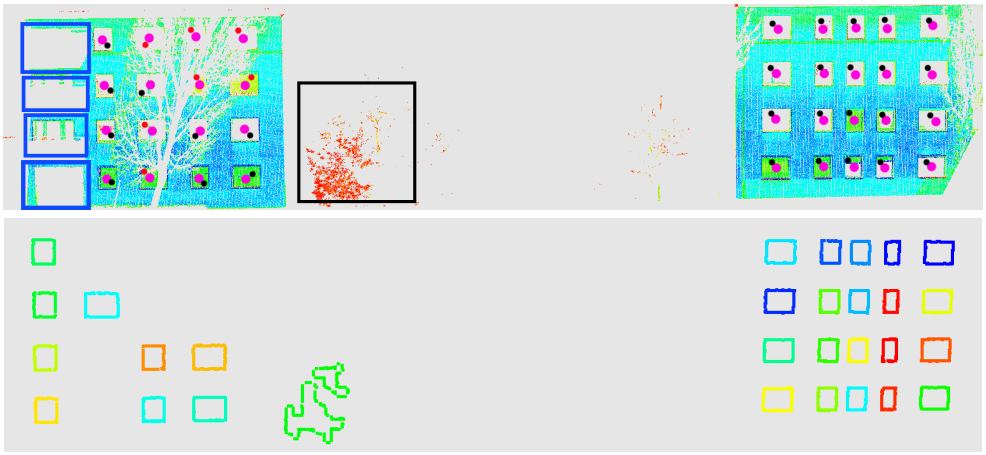


Abb. 13: Das obere Bild zeigt eine Fassade, die durch eine Straße in der Mitte getrennt wurde. Die Detektion der Ebene hat trotz der großen Lücke reibungslos funktioniert. Das untere Bild zeigt das Ergebnis der Fensterdetektion. Siehe Text für eine detaillierte Beschreibung.

Die quantitative Auswertung für diese Fassade zeigt ähnliche Ergebnisse wie für die erste (Tab. 2). Bei diesen Ergebnissen ist klar ersichtlich, dass die Methode bei Fassaden ohne Erker seht gute Ergebnisse erzielt, jedoch sind Verschattungen, wie bei vielen anderen Methoden ein Problem für die Auswertung. Die blauen Rechtecke in Abbildung 13 (oben) sind Balkons (die oberen drei) und ein Garagentor (unteres Rechteck), welche korrekt als „kein Fenster“ erkannt wurden.

Tab 2: Quantitative Auswertung für die Klassifikation der Fenster für Fassade 2. Treffgenauigkeit und F1-score liegen jeweils bei 80 % und 88 %.

		Tatsächlich	
		Fenster	**kein Fenster**
Prädiziert	**Fenster**	TP=29	FP=1
	kein Fenster	FN=7	TN=4

5 Fazit & Ausblick

Die hier präsentierte Methode nutzt ein nachvollziehbares, simples Konzept zur Herausarbeitung der Fenster in Fassaden in hochauflösenden 3D Punktwolken. Das Verfahren nutzt die hohe Punktdichte zur groben Bestimmung der Fensterkonturen. Eine Generalisierung der Fensterformen wird vorgenommen, indem für jedes Fenster ein kleinstes umschreibendes Rechteck bestimmt wird. Von diesem Rechteck kann die Fläche robust bestimmt werden und somit die generalisierte Fensterfläche. Die Fensterposition kann z.B. über den Schwerpunkt der Punkte bestimmt werden. In zukünftigen Arbeiten können diese Merkmale (Fläche, Schwerpunkt) dafür verwendet werden, die Stockwerksanzahl und das Fenster-Fassaden-Verhältnis zu bestimmen. Ergebnisse zeigen, dass eine Trefferquote zwischen 80 und 90 % erzielt werden kann. Dies steht allerdings unter der Voraussetzung, dass eine Fassade keine Erker hat oder diese bei der Auswertung nicht berücksichtigt werden. Die Detektion von Fenstern in Erkern ist eine Einschränkung dieser Methode, die aber in zukünftigen, weiterführenden Arbeiten behoben werden kann. Ein möglicher Ansatz hierfür wäre eine weitere Ebenensuche in einem gewissen Abstand vor der bereits detektieren Fassade.
Die weitere Arbeit wird sich damit befassen, das Fenster-Fassaden-Verhältnis und die Stockwerksanzahlen zu bestimmen, da dies wichtige Kenngrößen für energetische Simulationen sind. Vorläufige Ergebnisse zeigen, dass die präzise Berechnung der Fassadenfläche ein offensichtlich wichtiger Schritt für genaue Ergebnisse ist. Das getestete Verfahren (Greedy-Triangulation) funktioniert zwar für unorganisierte Punktwolken meistens gut, verursacht jedoch teils große Lücken im generierten Gitter (Mesh), was zu Fehlern bei der Flächenberechnung führt. Hierfür werden andere Ansätze getestet.

6 Dank

Die in dieser Arbeit verwendeten Punktwolkendaten sind im Rahmen des von der Bundesanstalt für Landwirtschaft und Ernährung (BLE) geförderten Projektes “WeBest – Wärmebedarfsprognose von Gebäuden und Stadtquartieren“ entstanden. Die Autoren bedanken sich bei der Stadt Essen für die gute Kooperation während des Projekts. Das Projekt „i_city: Intelligente Stadt“ wird vom Bundesministerium für Bildung und Forschung (BMBF) unter dem Förderkennzeichen 13FH9I01lA gefördert und vom Projektträger VDI Technologiezentrum GmbH für das BMBF betreut. Diese Arbeit ist im Rahmen des „i_city“ Projekts entstanden.

7 Literaturverzeichnis

ALI, H., SEIFERT, C., JINDAL, N., PALETTA, L. & PAAR, G., 2007: Window Detection in Facades. 14th International Conference on Image Analysis and Processing (ICIAP 2007), 837-842.

BARBER, C.B., DOBKIN, D.P. & HUHDANPAA, H., 1996: The Quickhull algorithm for convex hulls. ACM Transactions on Mathematical Software, **22**(4), 469-483.

BILJECKI, F., STOTER, J., LEDOUX, H., ZLATANOVA, S. & ÇÖLTEKIN, A., 2015: Applications of 3D City Models: State of the Art Review. ISPRS International Journal of Geo-Information, **4**(4), 2842-2889.

BMWI, 2011: Research for an environmentally sound, reliable and affordable energy supply. https://www.bmwi.de/Redaktion/EN/Publikationen/research-for-an-environmentally-sound-reliable-and-affordable-energy-supply.pdf?__blob=publicationFile&v=3, letzter Zugriff 11.01.2018.

BMWI, 2014: Hoher Energieverbrauch des Gebäudesektors. https://www.bmwi-energiewende.de/EWD/Redaktion/Newsletter/2014/22/Meldung/hoher-energieverbrauch-des-gebaeudesektor.html, letzter Zugriff 11.01.2018.

MONIEN, D., STRZALKA, A., KOUKOFIKIS, A., COORS, V. & EICKER, U., 2017: Comparison of building modelling assumptions and methods for urban scale heat demand forecasting. Future Cities and Environment, **3**(1), 2.

MONIEN, D., WILTING, R., CASPER, E., BRENNENSTUHL, M. & COORS, V., 2016: WeBest – Automatisierte Korrektur und Mapping von Fassadenthermographien auf 3D-Gebäudemodelle. Photogrammetrie - Fernerkundung - Geoinformation, **2016**(4), 246-257.

NEUHAUSEN, M., OBEL, M., MARTIN, A., MARK, P. & KÖNIG, M., 2018: Window detection in facade images for risk assessment in tunneling. Visualization in Engineering, **6**(1), 1.

NGUATEM, W., DRAUSCHKE, M. & MAYER, H., 2014: Localization of Windows and Doors in 3d Point Clouds of Facades. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-3, 87-94.

NOUVEL, R., BRASSEL, K.-H., BRUSE, M., DUMINIL, E., COORS, V., EICKER, U. & ROBINSON, D., 2015: SimStadt, a new workflow-driven urban energy simulation platform for CityGML city models. J.-L. Scartezzini (Ed.), Proceedings of International Conference CISBAT 2015 "Future Buildings and Districts – Sustainability from Nano to Urban Scale", EPFL Lausanne, LESO-PB, 889-894.

NOUVEL, R., ZIRAK, M., COORS, V. & EICKER, U., 2017: The influence of data quality on urban heating demand modeling using 3D city models. Computers, Environment and Urban Systems, **64**, 68-80.

RUSU, R.B., 2009: Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments. Computer Science department, Technische Universitaet Muenchen.

RUSU, R.B. & COUSINS, S., 2011: 3D is here: Point Cloud Library (PCL). IEEE International Conference on Robotics and Automation, 1-4.

TRINH, H.-H., KIM, D.-N., KANG, S.-J. & JO, K.-H., 2009: Window Extraction Using Geometrical Characteristics of Building Surface. D.-S. Huang, K.-H. Jo, H.-H. Lee, H.-J. Kang & V. Bevilacqua (Eds.), Emerging Intelligent Computing Technology and Applications, Springer Berlin Heidelberg, 585-594.

TUTTAS, S. & STILLA, U., 2013: Window Detection in Sparse Point Clouds Using Indoor Points. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(3), 131-136.

ZHANG, R. & ZAKHOR, A., 2014: Automatic identification of window regions on indoor point clouds using LiDAR and cameras. IEEE Winter Conference on Applications of Computer Vision, 107-114.

Entwicklung eines low-cost Kamerapegels zur Erfassung hydrologischer Extremereignisse

ANETTE ELTNER[1], HANNES SARDEMANN[1], MELANIE KRÖHNERT[1] & HANS-GERD MAAS[1]

Zusammenfassung: Die messtechnische Erfassung von Sturzfluten nach Starkregenereignissen ist aufgrund ihres lokalen Auftretens schwierig. Dies betrifft insbesondere Flüsse ohne ausreichende Pegelinfrastruktur. Auch fehlende Informationen über deren Einzugsgebiete erschweren Frühwarnungen sowie die Entwicklung von Vorhersagemodellen erheblich. Im Rahmen der Nachwuchsforschergruppe EXTRUSO werden kleinräumige hydrometeorologische Extremereignisse zum Aufbau eines solchen Frühwarnsystems für Sturzfluten untersucht. Ein wesentlicher Bestandteil ist die Erarbeitung neuer Methoden zur Beobachtung unüberwachter Einzugsgebiete. Der Beitrag umfasst die Entwicklung eines kamera-basierten Fließgewässer-Pegelmesssystems auf Basis von low-cost (Raspberry Pi) Kameras.

Stationäre terrestrische Kameras bieten bei entsprechenden Bildanalysemethoden und geeigneter Georeferenzierung die Möglichkeit, Gewässerstände mit einer hohen Genauigkeit zu bestimmen. Die hohe Variabilität von Fließgewässern schränkt jedoch die automatische Ableitung von Wasserständen aus terrestrischen Einzelbild-Aufnahmen stark ein. Abhilfe kann eine Ausdehnung der Analyse der Bildtextur auf die Zeitachse schaffen. Ko-registrierte Zeitraffersequenzen beschreiben den mittleren Wasserstand in Abhängigkeit des Beobachtungszeitraums und liefern zusätzlich temporale Texturmaße. Mittels Segmentierung von dynamischen (Wasser) und statischen (Land) Bereichen auf Basis der Raum-Zeit-Textur lässt sich die Wasserlinie extrahieren. Die Transformation der Bildmessung in den Objektraum erfolgt unter Nutzung eines photogrammetrisch erstellten 3D-Modells des beobachteten Gewässerabschnittes. Bei entsprechender Aufnahmekonfiguration ist eine Genauigkeit der Pegelmessung im Bereich von wenigen cm zu erwarten, welche durch erste Ergebnissen verifiziert werden konnte.

1 Einleitung

Aufgrund des Klimawandels ist eine Zunahme von hydrologischen Extremereignissen zu erwarten (COLLINS et al. 2013). U.a. ist mit einer Zunahme von Starkregenereignissen und infolgedessen Sturzfluten zu rechnen. Sturzfluten sind Hochwasserereignisse mit kurzen Reaktionszeiten von weniger als sechs Stunden (BORGA et al. 2011). Sturzfluten in kleinen und mittleren Einzugsgebieten stellen aufgrund ihres schnellen und intensiven Auftretens eine Gefahr für die zivile Bevölkerung sowie ökonomische Werte dar. Ein medial präsentes Beispiel ist das hydrologische Jahrhundertereignis vom Sommer 2016 im Einzugsgebiet des Braunsbach in Baden-Württemberg. Innerhalb von 75 Minuten erfolgte ein Niederschlag von 140 mm, was zu schnell steigenden Wasserständen und enormen Mengen mitgeführter Sedimentfracht führte (BRONSTERT et al. 2017). Ziel der Nachwuchsforschergruppe EXTRUSO (https://extruso.bu.tu-dresden.de; WIEMANN et al. 2017) ist es, ein Frühwarnsystem vor Sturzfluten zu entwickeln.

[1] Technische Universität Dresden, Institut für Photogrammetrie & Fernerkundung, Helmholtzstr. 10, D-01069 Dresden, E-Mail: [Anette.Eltner, Hannes.Sardemann]@tu-dresden.de

Dafür ist, neben Aspekten der hydro-meteorologischen Modellierung, das morphologische und hydrologische Monitoring von Flüssen in den aufgrund von hydromorphologischen sowie klimatischen Gegebenheiten gefährdeten Einzugsgebieten notwendig (für mehr Informationen bezüglich der morphologischen Erfassung siehe SARDEMANN et al. 2018). Jedoch sind bisher hauptsächlich größere Einzugsgebiete mit einem Pegelnetz abgedeckt. Kleinere Einzugsgebiete sind aus Kostengründen überwiegend unüberwacht. Um diese Datenlücke zu schließen soll ein low-cost Kamerapegelsystem entwickelt werden, welches automatisiert den Wasserstand erfasst. Im Rahmen dieser Arbeit werden ein Konzept für ein bildbasiertes Pegelsystem sowie erste Ergebnisse vorgestellt. Der Text ist wie folgt gegliedert: In Kap. 2 werden das Kamerasetup sowie die weiterhin nötige Datenakquise zur Wasserstandsmessung erläutert. In Kap. 3 wird die Methode zur bildbasierten Pegelmessung vorgestellt. Anschließend (Kap. 4) erfolgt eine Betrachtung und Diskussion der Ergebnisse für Messungen über einen mehrwöchigen Zeitraum an einem sächsischen Fluss eines mittleren Einzugsgebietes. Die Publikation schließt mit einem kurzen Ausblick auf anschließende Arbeitsschritte ab.

2 Datenakquisition

Damit aus autark akquirierten Aufnahmen des Gewässers erfolgreich Wasserstände automatisiert abgeleitet werden können, müssen verschiedene Bedingungen erfüllt sein, welche im Folgenden beschrieben werden.

2.1 Erfassung des Untersuchungsgebietes

Das gewählte Untersuchungsgebiet zur Entwicklung des bildbasierten Pegels befindet sich an der Wesenitz in Ostsachsen. Sie hat eine Länge von 55 km mit einem dazugehörigen Einzugsgebiet von 227 km². Die Wesenitz hat eine Pegelstation und ist damit gut für die Genauigkeitsverifikation im Rahmen einer Pilotstudie geeignet. Die low-cost Kamera wird in unmittelbarer Nähe zur Pegelstation Elbersdorf aufgebaut (Abb. 1). Die Station befindet sich etwa 5 km vor der Mündung in die Elbe. Wasserstand und Durchfluss werden hier im 15-minütigen Rhythmus aufgezeichnet. Der mittlere Durchfluss beträgt hier 2.2 m³/s bei einem Wasserstand von 48 cm.

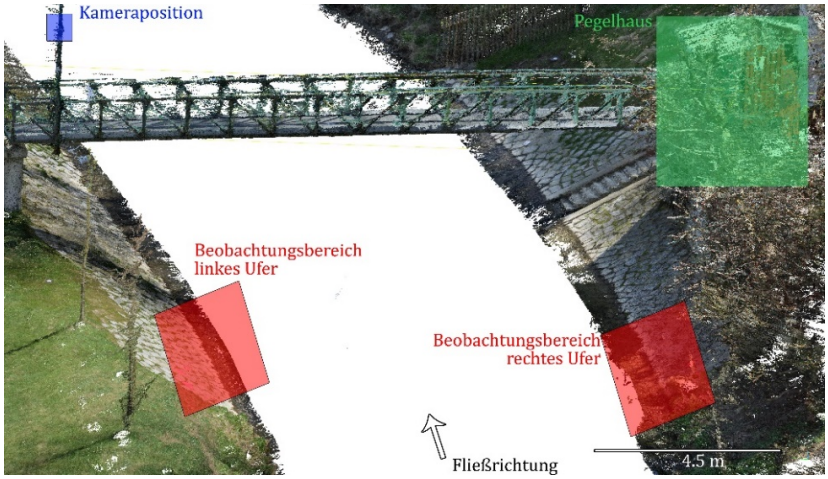


Abb. 1: 3D-Punktwolke des Beobachtungsgebietes am Pegel Elbersdorf Fluss Wesenitz. Rote Flächen kennzeichnen die Kameragesichtsfelder zur Wasserstandsmessung. Die blaue Fläche kennzeichnet die Position der Kamera

Zur geometrischen Beschreibung des Untersuchungsgebietes wurden terrestrisch und UAV-gestützt Bilddaten akquiriert. Anschließend wurden mittels Structure-from-Motion (SfM) in Agisoft PhotoScan 3D-Punktwolken aus den jeweils überlappenden Fotos berechnet (Abb. 1). Für die terrestrische Akquise von 440 Bildern sind zwei Spiegelreflexkameras (Canon EOS 500D und 1200D) genutzt worden. Alle UAV-basierten Aufnahmen sind mit der AscTec Falcon 8 (u.a. eingesetzt für Bodenerosionsüberwachung, in ELTNER (2013)) mit stabilisierender Kamerahalterung aufgenommen worden. Insgesamt dienen 20 Fotos von einer mittleren Flughöhe von 25 m zur 3D-Modellierung. Abschließend wurden beide Oberflächenmodelle (terrestrisch sowie UAV-gestützt) fusioniert, um eine Punktdichte von etwa 4 Punkten/cm² zu erhalten.

2.2 Einsatz eines low-cost Kamerasystems für Zeitrafferaufnahmen

In dieser Studie wird für die regelmäßige Aufnahme von kurzen Bildsquenzen im halbstündlichen Intervall eine kleine low-cost Kamera verwendet, deren Bildqualität vergleichbar mit der einer Smartphonekamera ist. Neben einer wetterfesten Installation müssen Referenzpunkte eingerichtet werden, und die Kamera muss kalibriert werden, damit Messungen mit höchstmöglicher Genauigkeit realisiert werden können.

2.2.1 Kamerainstallation und -kalibrierung

Der Kamerapegel besteht aus einer Raspberry Pi Kamera (V2.1). Die Kamera besitzt eine feste Brennweite von 2.9 mm. Die Auflösung beträgt 2592 × 1944 Pixel bei einer Pixelgröße von 1.4 µm. Zur Ansteuerung der Kamera wird der Einplatinencomputer Raspberry Pi Zero genutzt, um den Stromverbrauch sowie die Kosten zu minimieren. Die Kamera ist in einem wasserdichten PVC-Rohr untergebracht (Abb. 2) und an einem Laternenpfahl in einer Höhe von etwa 4 m über Grund befestigt (Abb. 1). Die Kamera ist so platziert, dass sich beide Ufergebiete innerhalb des Blickbereiches befinden (Abb. 1 und 4). Die Beobachtungsdistanz beträgt für beide Ufer wenige Meter.

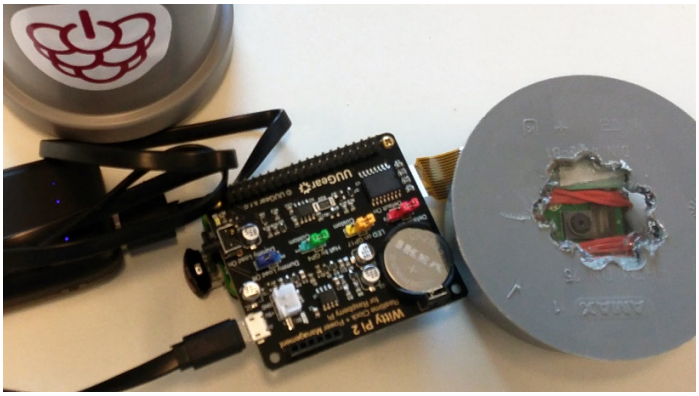


Abb. 2: Low-cost Raspberry Pi Kamera verbaut in einem wasserdichten PVC-Behälter. Ansteuerung erfolgt mittels eines Raspberry Pi Zero (hier installiert unterhalb des Witty Pi zur Speicherung der Uhrzeit)

Die Raspberry Pi Kamera wird vor ihrer Installation am Pegel zur Ermittlung der inneren Orientierung kalibriert. Die Kalibrierung erfolgt mit einer Kalibriertafel, die vor der Kamera bewegt wird, um verschiedene Perspektiven auf das Objekt zu realisieren. Anschließend erfolgt eine freie Bündelblockausgleichung in AICON 3D Studio. Die Ergebnisse der Kalibrierung zeigten eine deutliche Verzeichnung.

2.2.2 Referenzierung

Zur Übertragung der Bildmessungen in den Objektraum wurden an beiden Ufern Passpunkte (PP) installiert und mittels Tachymeter eingemessen (Abb. 3). Jedoch konnten diese PP nur zu Beginn der Messkampagne verwendet werden, da aufgrund von zunehmender Vegetation die Punkte später nicht mehr erkennbar waren. Daher wurden weitere natürliche Merkmalspunkte (z.B. an Steinecken) als PP zur Orientierung der Bilder im übergeordneten System gewählt. Die Objektkoordinaten dieser Punkte wurden mit Hilfe der terrestrischen Bildmessungen zur Rekonstruktion des Geländemodells (siehe Kap. 2.1) ermittelt. Dazu wurde der jeweilige Merkmalspunkt in mindestens sechs Bildern von deutlich verschiedenen Perspektiven detektiert.
Die Messung der PP kann in den überwiegenden Fällen der Datenakquisition nicht automatisiert innerhalb der Bilder erfolgen, da Belichtungsverhältnisse (z.B. verändernde Schatten oder Überbelichtung) sowie sich im Wind bewegende Vegetation keine eindeutige Zuordnung der PP in aufeinanderfolgenden Bildern mittels Template-Matching zulässt. Somit wurde in dieser Studie jeder PP behelfsmäßig mit Hilfe der Software ImageJ (SCHNEIDER et al. 2012) durchgeführt. Dies erlaubte die Messung in mehreren tausend Bildern in wenigen Stunden. Mit einem räumlichen Rückwärtsschnitt kann die äußere Orientierung für jede Aufnahme bestimmt werden.

Abb. 3: Position der Passpunkte (innerhalb der roten Rahmen). Während des Beobachtungszeitraumes mussten weitere natürliche Passpunkte aufgrund von Verdeckungen durch Bewuchs zur Kameraorientierung hinzugezogen werden. Bild mit Raspberry Pi Kamera aufgenommen

3 Automatisierte Wasserstandserkennung

Der Workflow zur bild-basierten, automatisierten Erkennung des Wasserstandes (Abb. 4) ist eine Erweiterung des ursprünglich für Smartphones entwickelten Tools zur Wasserlinienextraktion mit minimaler Benutzerinteraktion (KRÖHNERT & MEICHSNER 2017).

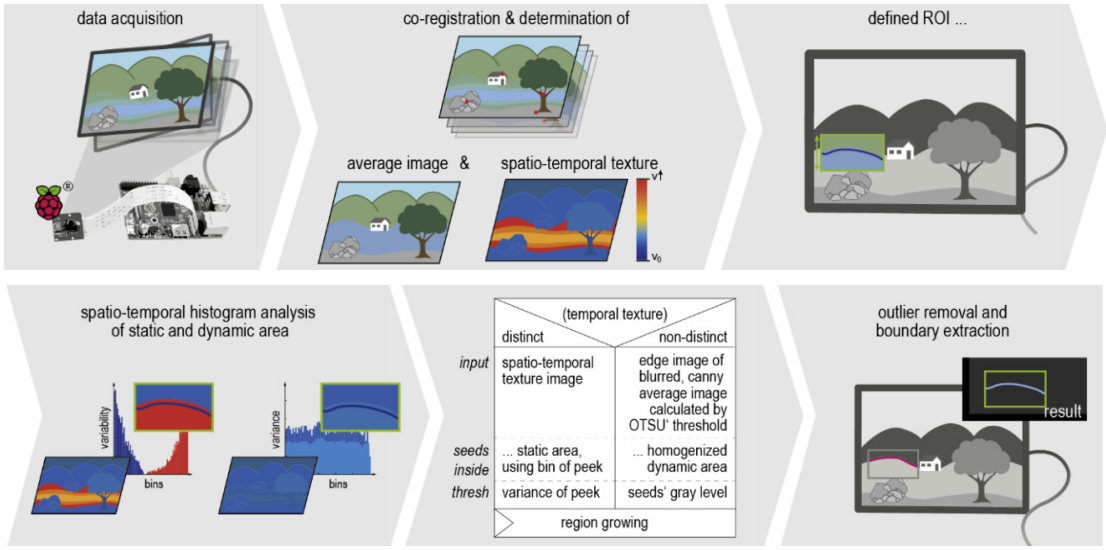


Abb. 4: Workflows zur Wasserlinienextraktion - Arbeitsschritte zur automatisierten Erkennung der Wasserlinie aus Bildsequenzen akquiriert mit einer low-cost Raspberry-Pi-Kamera (verändert nach KRÖHNERT & MEICHSNER, 2017)

3.1 Verarbeitung der Bildsequenz

Die Datenbasis zur automatischen Wasserlinienextraktion bilden kurze Bildsequenzen. Hierzu werden alle 30 Minuten bei ausreichend Tageslicht, festgestellt mittels eines Luxmeters, fünf Bilder im Abstand von 0.5 Sekunden aufgenommen. Aufgrund von möglichen Bewegungen der Kamera, z.B. infolge von Wind, werden die Bilder einer jeden Sequenz ko-registriert. Das erste

Bild einer Sequenz dient als Referenzbild (der sog. Master), zu dem alle folgenden Bilder der gleichen Sequenz registriert werden.
Nach der Ko-Registrierung der Bildsequenz wird eine temporale Textur berechnet (Abb. 5b). Dazu werden die Bilder der Sequenz (sog. Slaves) vom Master subtrahiert und anschließend die Absolutwerte der Differenzen aufsummiert. Idee dieses Vorgehens ist, dass sich bewegte Bereiche (fließendes Wasser) von statischen Regionen (bebautes Ufer) eindeutig separieren lassen. Jedoch gibt es Situationen, in denen Ufer und Wasser in der temporalen Textur nicht voneinander zu unterscheiden sind. Dies ist der Fall bei sehr langsam fließendem Wasser, Uferbereichen mit hoher Dynamik (bspw. bewegte Vegetation) oder bei bestimmten Belichtungsbedingungen (z.B. Überbelichtung). In diesen Fällen wird für die weitere Bildverarbeitung mit Kantenoperatoren ein Durchschnittsbild aus den 5 Bildern der Sequenz berechnet (Abb. 5a).

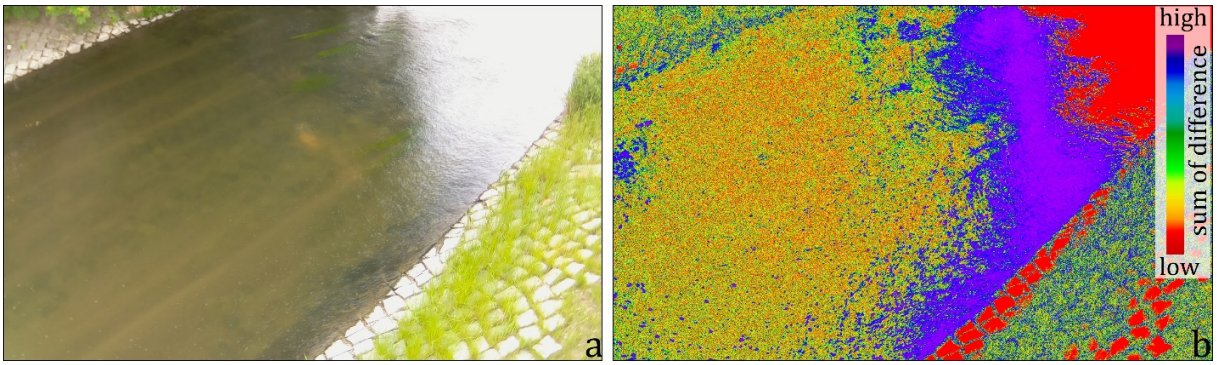


Abb. 4: a) Durchschnittsbild. b) Temporale Textur

3.2 Extraktion statistischer Bildinformationen

Ist die Bildsequenz vorbereitet, erfolgt im nächsten Schritt die Datenverarbeitung zur Entscheidung, ob das Durchschnittsbild oder die Temporaltextur zur Wasserlinienerkennung genutzt werden muss. Dazu wird ein Suchbereich festgelegt, indem mit Hilfe der Python-Bibliothek Shapely (GILLIES 2007) ein Buffer angelegt wird, welcher durch zwei Punkte definiert wird. Diese Punkte sind initial zu definieren und sollten die Wasserlinie in der Nähe des mittleren Wasserstandes nachzeichnen. Danach werden die Temporaltextur und das Durchschnittsbild mit dem Buffer ausgeschnitten, um den Suchbereich für die Wasserlinie zu minimieren. Zusätzlich müssen die Teilflächen des Buffers mit Wasser- und Landbereich attribuiert werden. Des Weiteren wird die ausgeschnittene Temporaltextur mittels eines bilateralen Filters (Python Bibliothek scikit-image; VAN DER WALT 2014) bearbeitet, damit der Einfluss des Bildrauschens der low-cost Kameras (geringere Bildqualität aufgrund kleiner Blende und sehr kleiner Pixel), welches sich beim Aufaddieren der Differenzen verstärkt, minimiert wird. Das Bild wird bei überwiegender Erhaltung von Kanten geglättet.
Nach der Filterung wird der Ausschnitt der Temporaltextur entlang der genäherten Wasserlinie in einen angenommenen statischen und dynamischen Bereich getrennt, um mit Hilfe einer Histogrammkorrelation beide Seiten miteinander zu vergleichen (OpenCV für Python; BRADSKI 2000). Die zugrundeliegende Annahme ist, dass sich der dynamische Bereich durch eine wesentlich breitere Verteilung der aufsummierten Differenzen kennzeichnet als der statische Bereich,

der eine geringe Streuung bei generell niedrigen Summen erwarten lässt. Der berechnete Korrelationskoeffizient bildet die Entscheidungsgrundlage für die folgende Bildbearbeitung.

3.3 Erkennung der Wasserlinie

Zur weiteren Bildverarbeitung wird in Abhängigkeit der Histogrammähnlichkeit das Durchschnittsbild oder die Temporaltextur genutzt. Liegt der Korrelationskoeffizient unterhalb eines definierten Schwellwertes, erfolgt ein Region Growing mittels der Temporaltextur, und im umgekehrten Fall wird ein Kantenoperator im Durchschnittsbild verwendet.

3.3.1 Region Growing

Im Fall einer geringen Ähnlichkeit im Korrelationsergebnis und somit guter Unterscheidung zwischen dynamischen und statischen Bereichen in der Temporaltextur, werden Land- und Wasserseite automatisch voneinander getrennt, indem homogene Bereiche (niedrige Summen der Differenzen) der Landseite mit dem iterativen Bildsegmentierungsverfahren Region Growing vereint werden. Dazu wird der Pixelwert des Maximums im Histogramm der aufsummierten Differenzen zur Definition von Seedpoints (Startpunkte) sowie als Schwellwert für das Region Growing verwertet. Die Seedpoints werden auf der Landseite lokalisiert, um anschließend mittels Region Growing alle Punkte innerhalb einer 4er-Nachbarschaft zur Landseite hinzuzufügen, wenn sie unterhalb des Schwellwertes liegen, d.h. bis sie auf die Wasserseite stoßen. Das Ergebnis des Region Growing ist ein binäres Bild. Die Wasserlinie wird mittels einer Konturextraktion gefunden. Bei mehreren Konturen im Fall von unterbrochenen Region Growing wird die längste Kontur behalten.

3.3.2 Kantendetektion

Ist die Unterscheidung zwischen Wasser- und Landseite in der Temporaltextur nicht ausreichend, wird das Durchschnittsbild mittels des Canny-Operators gefiltert (CANNY 1986), um Kanten zu detektieren. Hierzu wird der ausgeschnittene Durchschnittsbildbereich mit einem bilateralen Filter sowie einem Medianfilter vor der Kantendetektion gefiltert, um die Grenze zwischen Land und Wasser zu verstärken und den Einfluss von Rauschen entlang der Wasserlinie zu unterdrücken. Werden mehrere Kanten erkannt, werden diese zu einer vereint, wenn die einzelnen Kanten eine definierte Mindestlänge überschreiten.

3.4 Von der Wasserlinie zum Wasserstand

Ist die Wasserlinie im Bildraum erfasst, muss sie abschließend in den Objektraum übertragen werden, um den Wasserstand zu erhalten. Innerhalb dieser Arbeit wird das Referenzbild einer jeden Bildsequenz mittels eines Rückwärtsschnittes im gleichen System wie das Geländemodell der beobachteten Ufer orientiert (siehe Kap. 2.2.2). Danach wird die mittels SfM berechnete 3D-Punktwolke in den Bildraum der low-cost Kamera projiziert. Im so berechneten Tiefenbild werden anschließend mittels eines kd-Baumes (scikit-image) die nächsten Nachbarn der Punktwolke zu den Punkten der detektierten Wasserlinie bestimmt. Die erkannten Punkte werden abschließend wieder in den Objektraum projiziert. Für die Tiefeninformation wird dabei auf das berechnete Tiefenbild zurückgegriffen. Da für den Wasserstand lediglich ein 3D-Punkt ausreichend ist, wird der Median der Höhenwerte verwendet.

4 Ergebnisse und Diskussion

Im Rahmen einer Pilotstudie wurde das Tool zur bildbasierten Wasserstandsmessung auf der linken und rechten Uferseite für einen Zeitraum von etwa 5 Wochen im Frühsommer 2017 getestet. Neben den automatischen Messungen erfolgte des Weiteren eine manuelle bildbasierte Messung des Wasserstandes. Die Ergebnisse der manuellen Messung sind in Abb. 6 und die der automatischen Messung in Abb. 7 dargestellt. In beiden Varianten werden die Werte zu den Pegelmessungen des Pegels Elbersdorf verglichen. Dabei muss jedoch berücksichtigt werden, dass bildbasierte Werte und Pegelmessung verschiedene zeitliche Auflösungen besitzen und somit insbesondere kurzfristige Wasserstandsschwankungen nicht vergleichbar sind.

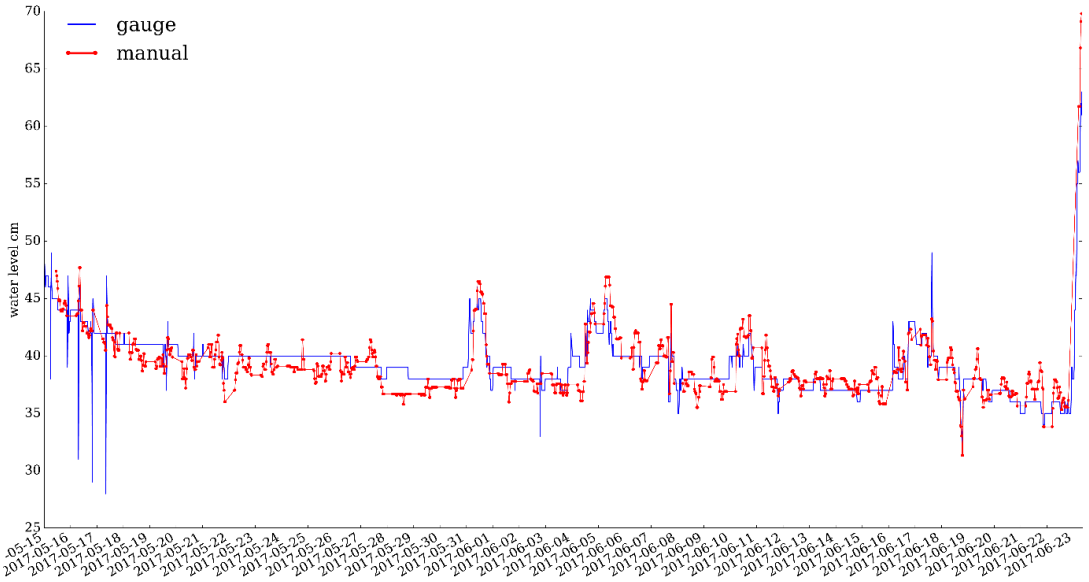


Abb. 5: Ergebnis der bildbasierten manuellen Messung des Wasserstandes (rote Linie). Blaue Linie zeigt die Pegelmessung

Mittels der bildbasierten manuellen Messung sind Schwierigkeiten bezüglich des low-cost Kamerasystems erkennbar. Aufgrund der erschwerten PP-Messung durch die notwendige Nutzung von natürlichen Merkmalspunkten gibt es Komplikationen bei der Orientierung der Kamera. Außerdem ist die Wasserlinienerkennung in den Bildern bei bestimmten Lichtverhältnissen schwierig, da z.B. durch das Wasser geschaut wird und somit eine klare Grenze nicht sichtbar ist. Daher ist in Abb. 6 ein Rauschen des bildbasierten Wasserstandes erkennbar. Dennoch ist eine prinzipielle Übereinstimmung bei der Beschreibung des zeitlichen Verlaufes der Wasserstandsschwankungen gegeben. Die Korrelation zwischen Pegelmessung und bildbasierter manueller Messung beträgt 0.81 und die mittlere Abweichung zwischen beiden Messvarianten beträgt 0.2 ± 1.4 cm (Tab. 1).

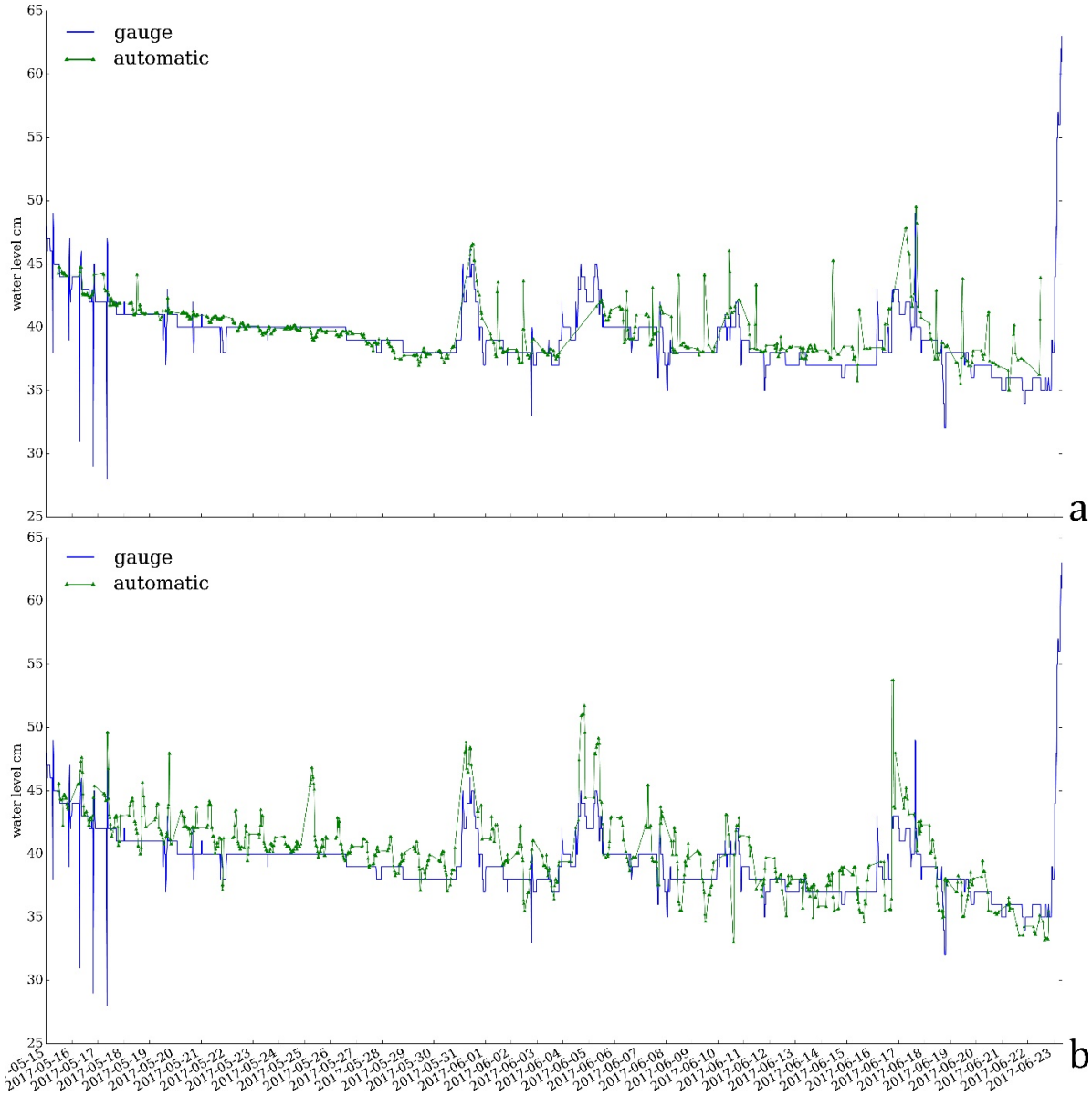


Abb. 6: Bildbasierte automatisierte Messung des Wasserstandes (grüne Linie). a) Wasserstand des rechten Flussufers. b) Wasserstand des linken Flussufers. Blaue Linie zeigt die Pegelmessung

Tab. 1: Statistik des Vergleiches der Wasserstandmessungen abgeleitet aus den Bildern (manuell und automatisiert) mit den Wasserständen gemessen am Pegel

	manuelle Messung	automatisierte Erkennung rechte Uferseite	automatisierte Erkennung linke Uferseite
Spearman-Korrelation	0.81	0.78	0.84
Anzahl N	919	802	918
Mittelwert Differenz [cm]	0.2	-0.6	-1.2
Standardabweichung Differenz [cm]	1.4	1.7	2.0

Die automatisierte Wasserstandsmessung zeigt eine unterschiedliche Qualität bezüglich der Genauigkeit und der Zuverlässigkeit für das linke und das rechte Flussufer (Abb. 7 und Tab. 1). Auf der rechten Uferseite wird hauptsächlich der Canny-Operator im Durchschnittsbild angewendet, da von der Perspektive der low-cost Kamera kaum noch Wasserbewegungen in den Bildern beobachtbar sind. Das rechte Flussufer weist ein geringeres Rauschen zur Pegelmessung auf (mittlere Abweichung -0.6 $\pm$ 1.7 cm). Dies liegt u.a. an der besseren Perspektive des Messsystems auf eine fast senkrecht zur Kamera orientierten Uferseite. Des Weiteren wird jedoch der Trend der Wasserstandsschwankung auf der rechten Uferseite mit geringerer Zuverlässigkeit gemessen (Korrelationskoeffizient von 0.78), da insbesondere bei stärkeren Wasserstands-schwankungen keine Messungen erfolgt sind. Zum Beispiel ist der Kontrast vom 5. bis zum 6. Juni, während mehrerer Regentage, zwischen nassem (dunklen) Ufer und Fluss nicht ausreichend für den Kantenoperator. Die Messungen auf der linken Uferseite zeigen ein wesentlich höheres Rauschen (-1.2 $\pm$ 2.0 cm), dafür wird jedoch der Trend besser erfasst (Korrelationskoeffizient von 0.84), da auch stärkerer Hochwasserereignisse, meist überschätzend, gemessen werden.

5 Ausblick

Der in dieser Arbeit vorgestellte Workflow zur automatisierten bildbasierten Erfassung des Wasserstandes ermöglicht es Messungen mit einer Genauigkeit von wenigen cm durchzuführen. Jedoch ist zu beachten, dass Belichtungsverhältnisse sowie zunehmende Vegetation die Qualität der Messungen erschweren können. Des Weiteren ist insbesondere bei Verwendung der Temporaltextur ein erhebliches Rauschen der Wasserstandsmessungen möglich. Zur Minimierung der falsch erkannten Wasserlinie soll zukünftig ein sich iterative anpassender Suchbereich für die Wasserlinie entwickelt werden.
In der Arbeit konnte gezeigt werden, dass der langfristige Trend der Wasserstandsänderungen mit ausreichender Genauigkeit erkannt wird. Die Methode kann als low-cost Möglichkeit insbesondere für die Überwachung unbeobachteter Einzugsgebiete genutzt werden, da neben der Installation der Kamera sowie der Passpunkte keine weiteren Materialen vor Ort gebraucht werden. Lediglich das regelmäßige aktualisieren des Gewässerprofils z.B. mit SfM ist noch nötig.

6 Dank

Die vorliegende Publikation wurde im Rahmen des Projektes *EXTRUSO* sowie *Crowdsourcing-basierte Verdichtung hydrologischer Messnetze* verfasst. Dank geht an den Europäischen Sozial Fond (ESF) für die Finanzierung beider Projekte im Rahmen der Förderperiode 2014-2020 (Nr. 100270097 und 100235479).

7 Literaturverzeichnis

BORGA, M., ANAGNOSTOU, E. N., BLÖSCHL, G. & CREUTIN, J. D., 2011: Flash flood forecasting, warning and risk management: the HYDRATE project. Environmental Science & Policy **14** (7), 834-844.

BRADSKI, G., 2000: The OpenCV Library. Dr. Dobb's Journal of Software Tools.

BRONSTERT, A., AGARWAL, A., BOESSENKOOL, B., FISCHER, M., HEISTERMANN, M., KÖHN-REICH, L., MORAN, T. & WENDI, D., 2017: Die Sturzflut von Braunsbach am 29. Mai 2016 – Entstehung, Ablauf und Schäden eines „Jahrhundertereignisses". Teil 1: Meteorologische und hydrologische Analyse. Hydrologie und Wasserbewirtschaftung, **61**, 150-162.

CANNY, J., 1986: A Computational Approach to Edge Detection. IEEE Transactions on pattern analysis and machine intelligence, **PAMI-8** (6).

COLLINS, M., KNUTTI, R., ARBLASTER, J., DUFRESNE, J.-L., FICHEFET, T., FRIEDLINGSTEIN, P., GAO, X., GUTOWSKI, W., JOHNS, T., KRINNER, G., SHONGWE, M., TEBALDI, C., WEAVER, A. & WEHNER, M., 2013: Long-term Climate Change: Projections, Commitments and Irreversibility. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

ELTNER, A., 2013: Konzeption einer Methode zur quantitativen Erfassung von Bodenerosion durch integrierte Aufnahme mit einer Mikrodrohne und einem terrestrischen Laserscanner. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **22**, 289-297.

GILLIES, S., 2007: Shapely: manipulation and analysis of geometric objects. https://github.com/Toblerity/Shapely.

KRÖHNERT, M. & MEICHSNER, R., 2017: Segmentation of Environmental Time Lapse Image Sequences for the Determination of Shore Lines Captured by Hand-Held Smartphone Cameras. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **4**(2W4), 1-8.

SARDEMANN, H., ELTNER, A. & MAAS, H.-G., 2018: Erfassung von Geometriedaten kleiner Flüsse mit einem unbemannten Wasserfahrzeug als Multisensorplattform. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., T. Kersten (Hrsg.), Band **27**.

SCHNEIDER, C. A., RASBAND, W. S. & ELICEIRI, K. W., 2012: NIH Image to ImageJ: 25 years of image analysis. Nature methods, **9**(7), 671-675.

VAN DER WALT, S., SCHÖNBERGER J., NUNEZ-IGLESIAS, J., BOULOGNE, F., WARNER, J., YAGER, N, GOUILLART, E., YU, T. & THE SCIKIT-IMAGE CONTRIBUTORS, 2014: scikit-image: image processing in Python. PeerJ **2**, e453.

WIEMANN, S., ELTNER, A., SARDEMANN, H., SPIELER, D., SINGER, T., THANH LUONG, T., AL JANABI, F., SCHÜTZE, N., BERNARD, L., BERNHOFER, C., MAAS, H.-G. (2017). On the monitoring and prediction of flash floods in small and medium-sized catchments-the EXTRUSO project. Geophysical Research Abstracts, **19**, EGU2017- 4862.

Illumination Invariant Dense Image Matching based on Sparse Features

MAX MEHLTRETTER[1] & CHRISTIAN HEIPKE[1]

Abstract: In this paper we propose an algorithm for dense depth estimation based on a cost function, which is robust against changes in illumination. For this purpose, we utilize census transformation and Phase Congruency. The process of cost computation is supported by a triangle-based depth prediction approach using a set of matched feature points. These points can reliably be detected even under drastic changes in illumination. We demonstrate the performance of our approach on the challenging KITTI stereo dataset and a set of images with simulated changes. The results show that the concept achieves state-of-the-art on images with similar illumination conditions and is robust against changed conditions.

1 Introduction

Reliable and up-to-date maps received much attention in the field of mobility in recent times (SEIF et al. 2016). Especially the sector of assisted and autonomous driving strongly depends on such maps in order to be able to accurately determine the vehicle's pose and compute a safe trajectory. With regard to the importance of the currency of the information contained in such a map, data acquisition must be done continuously.

For the purpose of data acquisition, mobile mapping using stereo cameras is well suited. In comparison to laser scanners, stereo cameras are less expensive and furthermore, they do not only provide geometric, but also multi-spectral radiometric information. Hence, additional information for later image segmentation and classification is provided. But, special attention has to be paid to the fusion of current information and data captured earlier to obtain a consistent, reliable and accurate result. Depending on the temporal distance between two recordings, changes in illumination can lead to significant differences in the acquired data. Not only the color of illumination or the geometric relationship between illumination, scene and sensor but also the number of light sources might have changed, making the fusion task a potentially difficult challenge.

In this paper, we present a novel approach for illumination invariant dense image matching. Following the quote of Aristotele “The whole is greater than the sum of its parts”, we propose a cost function combining different approaches to achieve robustness against changes in illumination. For this purpose, we merge cost functions based on census transformation (ZABIH et al. 1994) and based on images transformed with Phase Congruency (KOVESI 1999). The disparity computation is supported by a triangle-based depth prediction approach (BULATOV et al. 2011) utilizing a set of matched feature points which are reliably detectable even under drastic changes in illumination (VERDIE et al. 2015). The optimal solution for the disparity image is determined by semi-global matching (HIRSCHMÜLLER 2008).

[1] Leibniz Universität Hannover, Institute of Photogrammetry and GeoInformation, Nienburger Straße 1, D-30167 Hannover, E-Mail: [mehltretter, heipke]@ipi.uni-hannover.de

The remaining sections of this paper are organized as follows: In section 2, we provide an overview of the field of illumination invariant image matching. Section 3 contains the methodology of our approach. Then, in Section 4, the experimental results are shown and explained. Finally, the conclusion is presented in Section 5.

2 Related Work

The reconstruction of depth information from one image pair is a classical task in photogrammetry and also the minimal case of the well-known structure from motion problem. It refers to the concept that 3D structures can be recovered from the projected 2D motion field of a scene acquired with a moving sensor. Given the orientation between the images, depth can be reconstructed for each pixel using dense matching approaches like HIRSCHMÜLLER (2008), YAMAGUCHI et al. (2014) or CIGLA (2015).

Dense Matching has received much attention in recent years due to its wide range of deployment. The spectrum ranges from applications in the automotive industry to the creation of digital surface models and the digitization of cultural heritage. However, most of the approaches concentrate on matching images taken quasi-simultaneously (like in aerial photogrammetry) or in an environment with controlled illumination conditions. Consequently, there are no or only minor changes in illumination between such images. Only a few approaches address the problem that arises from the influence of different lighting conditions, despite the fact that research was already conducted on this topic back in the 1990s such as COX et al. (1995).

Early approaches tried to solve the problem by using more robust metrics for matching cost computation. There are a whole range of metrics such as Normalized Cross Correlation, Mutual Information, Rank and Census (ZABIH et al. 1994). In 2009, HIRSCHMÜLLER et al. investigated the influence of radiometric changes on the performance of these metrics. In their paper, they discuss the limitations and demonstrate that none of the examined approaches works particularly well, if there are bigger differences in illumination between the images.

To improve the robustness of dense matching against such cases, HEO at al. (2011) proposed a metric called Adaptive Normalized Cross-Correlation. Based on the observation that color consistency does not hold between stereo images in real scenes, they explicitly used the color formation model to adjust the matching cost function. Furthermore, the images were transformed using Log-Chromaticity Normalization beforehand. In contrast, ZHOU at al. (2012) and JUNG et al. (2013) rely on gradient based approaches. They assume that gradients are less sensitive to radiometric variations than the original pixel values and hence, they utilize gradient strength, orientation and distribution.

Another category of methods first transforms the images to an alternative representation before performing dense matching. KIM at al. (2014) proposes a transformation to the Mahalanobis distance transform space. It is defined as a collection of the Mahalanobis distances between the color of a pixel and the average color in this pixel's neighborhood. This is an attempt to separate the color values of the individual pixels from the influence of lighting. It is based on the assumption that the relation between the color and the average color is preserved under affine transformation. The work of MADDERN et al. (2014) proposes a transformation to an illumination invariant color space. They assume the sun as only light source and model it as a Planckian source. This approach works well for outdoor scenes illuminated by day-light. MOUATS et al.

(2015) proposes local frequency analysis which is based on Phase Congruency (KOVESI 1999). It combines information of edges and texture which is extracted from the image's frequency domain examined in a local context.

One issue that all of these approaches share is that they are based on assumptions that significantly restrict the scope of the method. Some can only deal with global, but not local lighting differences (ZABIH et al. 1994), others make specific assumptions about the type of light source (MADDERN et al. 2014) or the type of lighting change (KIM at al. 2014). As a result, these approaches are not suitable for building a reliable map, as it is based on images taken under different and quickly varying lighting conditions. Consequently, our approach is not based on any assumptions about the illumination, in order to be able to take all these images into consideration when creating the map.

3 Illumination Invariant Dense Image Matching

In order to be able to reconstruct depth information from a pair of images with highly varying illumination, we propose an approach based on the combination of multiple cost functions for dense matching. For this purpose, we assume that the orientation is known and the images have already been rectified.

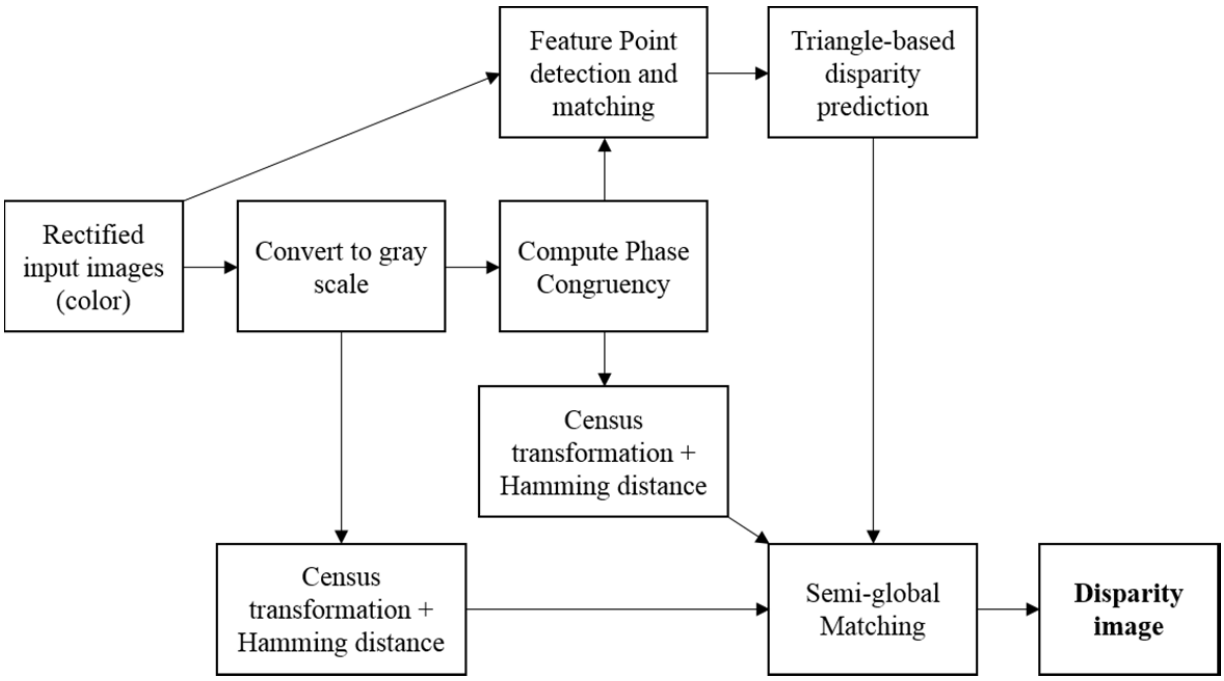


Fig. 1: Framework of the proposed dense image matching approach

As a basis we calculate the census transformation on the gray scale images as in ZABIH et al. (1994). The second element uses the Phase Congruency transformation of the images. Lastly, we use a triangle-based depth prediction, gained from a set of sparse feature matches, to support dense matching.

$$C(x,d) = \lambda_{Census} * C_{Census}(x,d) + \lambda_{PC} * C_{PC}(x,d) + \lambda_T * C_T(x,d) \quad (1)$$

The resulting cost function $C(x,d)$ is the weighted sum of the previously explained elements, with weights λ_{Census}, λ_{PC} and λ_T. It is computed for each pixel x and disparity d.

3.1 Phase Congruency

The Phase Congruency transformation is a concept developed by KOVESI (1999). It is based on the local energy model and shows the behavior of an image in frequency domain. The transformation can be used to detect and localize edges and corners and is invariant to changes in illumination and contrast. Consequently, the Phase Congruency remains constant for a scene even if the lighting conditions change.

Because of the Time-Frequency Uncertainty Principle it is not possible to exactly determine spatial position and frequency at the same time. Therefore, it is necessary to use an approximation to analyze the combined frequency-spatial space (SKARBNIK et al. 2010). Following RAGB et al. (2016) we use a bank of Log-Gabor filters for this purpose. Hence, the transformation can be computed over various scales n and orientations o with:

$$PC(x) = \frac{\sum_o \sqrt{(\sum_n (I(x) * M_{no}^e))^2 + (\sum_n (I(x) * M_{no}^o))^2}}{\sum_o \sum_n (A_{no}(x)) + \varepsilon} \quad (2)$$

where $I(x)$ is the input signal for which we use the gray scale image. M_{no}^e and M_{no}^o are the even and odd symmetric components of the Log-Gabor filters and ε is a small offset to prevent division by zero. The amplitude of the response A_{no} is defined by:

$$A_{no}(x) = \sqrt{(I(x) * M_{no}^e)^2 + (I(x) * M_{no}^o)^2} \quad (3)$$

We use 6 different orientations and 4 different scales, so we have a filter bank containing 24 different filters. We understand this as a compromise between accuracy and calculation effort. The impact on the cost function C_{PC} by increasing the number of filters was insignificantly small.

To compute the cost function C_{PC} both images are transformed with Phase Congruency, so an ordinary metric can be used to measure the similarity between image blocks:

$$C_{PC}(x,d) = F(I_{PC}^L(x), I_{PC}^R(x-d)) \quad (4)$$

where I_{PC}^L and I_{PC}^R are the left and right Phase Congruency images. As for the gray scale images, we use the Census transformation combined with Hamming distance for metric F.

3.2 Triangle-based disparity prediction

The third element of the matching cost function is based on triangle-based cost aggregation proposed by BULATOV et al. (2011). The basic idea is to build a triangle mesh based on a set of points for which the position is known in both images. The disparity for these points can then be directly computed from the difference of their positions in the two images. For all pixels within a triangle, the disparity is estimated by interpolating the disparity values of the triangle's vertices. This approach is comparable to methods based on image segmentation. However, it has the advantage that it is less prone to discretization errors and the number of disparity levels does not have to be set beforehand.

3.2.1 Sparse Feature Matches

In a first step, a set of points has to be specified which can be reliably found in both images. For this task, a feature detector and a descriptor are needed, which are robust against illumination changes. Research on feature detectors and descriptors has received much attention in recent years, but it was mainly focused on improving the reliability under viewpoint changes.

Nevertheless, there are some promising approaches like the TILDE detector proposed by VERDIE et al. (2015). Their approach is based on a piecewise linear regressor, which is learned beforehand. For training a series of images of the same scene taken under varying illumination conditions is used. On these images feature points are detected using SIFT (LOWE 2004). These points are introduced as positive samples for training. Negative samples are created by extracting patches at locations that are far away from the feature points.

In our work, we use the TILDE detector with a pre-trained linear regressor to detect feature points. In contrast to the original approach, we use a threshold on the regressor's value to filter for reliable points instead of keeping a constant number. As proposed by VERDIE et al. (2015) SIFT is used as descriptor. For the purpose of matching, a simple brute-force approach is utilized, which is improved by limiting the search space in vertical direction by a threshold ε. The set of potential matches for a feature point (x_L, y_L) detected in the first image is then defined by:

$$M_p(x_L, y_L) = \{(x_R, y_R) | y_L - \varepsilon \leq y_R \leq y_L + \varepsilon\} \tag{5}$$

where x_R, y_R are the coordinates of a potentially matching feature point in the second image. This approach is based on the assumption, that the input images are rectified and therefore, share a common image plane. Deviations in vertical directions are then caused by inaccuracies in the rectification process or movements in the scene only. Next, a distance is computed for every potential match in M_p:

$$M_d(x_L, y_L) = \{(x_R, y_R, \rho) |\ (x_R, y_R) \in M_p \cap \rho = F(x_L, y_L, x_R, y_R)\} \tag{6}$$

where F is a function that delivers the difference between two feature points based on their descriptors. From the set of potential matches a subset containing actual matches is extracted. For this purpose, we apply the ratio test (LOWE 2004) on the match with the minimal distance ρ for every feature point (x_L, y_L):

$$M(x_L, y_L) = \{(x_R, y_R, \rho_m) \in M_d(x_L, y_L) | \forall (x_R, y_R, \rho_n) \in M_d(x_L, y_L) \backslash (x_R, y_R, \rho_m) : \rho_m \leq \alpha * \rho_n\} \tag{7}$$

where α is a positive constant which regulates the strictness of the ratio test. For every feature point in the first image one match remains at most $|M(x_L, y_L)| \leq 1$. The final set of matches M is then defined as the union of all remaining feature point matches:

$$M = \bigcup_{x,y} M(x, y) \tag{8}$$

Furthermore, we detect the feature points not only on the color images, but also on the Phase Congruency images. The sets of detected feature points are quite different for both types of images. Hence, merging the sets of matched features leads to better coverage of the images and allows to create finer triangle meshes later on. Although the TILDE detector has been trained on

color images only, it works very well on the Phase Congruency images also. As descriptor SIFT is used as well.

Finally, the set of matched feature points is used to create a triangle mesh. For this purpose, the Delauney triangulation is used.

3.2.2 Triangular Interpolation

The approach of BULATOV et al. (2011) is based on the assumption that the non-occluded parts of a scene can be piecewise approximated by triangles. With a correct evaluation which triangles are nearly consistent with their covered surface, it is possible to approximate disparity values for all pixels within these triangles with high accuracy. However, this results in the difficulty to define constraints to filter triangles which include areas that cannot reasonably be approximated using a flat surface. For this purpose, we filter the set of triangles based on three policies: Size, variance of disparity values and variance of the color values of the included pixels. If the threshold of one or more policies is exceeded the triangle is excluded from the following processing steps.

In order to obtain disparity values for all pixels within the remaining triangles, the cost function introduced by BULATOV et al. (2011) is used as C_T:

$$C_T(x,d) = A(x,T) * D(d, d_{T,x}) \tag{9}$$

where $d_{T,x}$ is the interpolated disparity at pixel x within triangle T. The weight function A represents the reliability of the disparity in pixel x. It is maximal in the triangle's vertices and becomes smaller with increasing distance. Outside of the triangles boundaries A is zero.

$$A(x,T) = \hat{y} * exp\left(-\frac{g(x,T)}{\sigma}\right) \tag{10}$$

where the amplitude $\hat{y}$ and the descent parameter σ have to be greater than zero. g is the minimum distance from x to the vertices of triangle T. The function D evaluates the difference between the current disparity level d and the triangular prediction d_T in pixel x:

$$D(d, d_{T,x}) = \min\left(\frac{|d - d_{T,x}|}{d_0}, 1\right) - 1 \tag{11}$$

where $d_0 > 0$. Outside of the triangle's boundaries D is zero and accordingly, C_T is zero as well. Therefore, C_T only influences the value of the combined cost function C in areas which are covered by the triangle mesh. Based on the initial assumption this is reasonable, since it is not possible to make a reliable statement about the disparity in other regions.

3.3 Non-local optimization

In order to find the optimum disparity value from the combined cost function C for each pixel, Semi-Global Matching is used. According to HIRSCHMÜLLER (2008), the energy term E depends on the disparity image D and considers the combined matching costs C of (1) and a smoothness term:

$$E(D) = \sum_x (C(x, D_x) + \sum_{x' \in N_x} \lambda_1 T[|D_x - D_{x'}| = 1] + \lambda_2 T[|D_x - D_{x'}| > 1]) \tag{12}$$

The smoothness term adds a constant penalty λ_1 for all pixels x' in the neighborhood N_x of x, for which the difference in disparity is equal to one. For neighboring pixels with a difference in disparity bigger than 1, a constant penalty λ_2 with $\lambda_1 \leq \lambda_2$ is added. The introduction of a smoothness constraint with two different penalties is reasonable: Small changes in disparity typically indicate slanted surfaces and should be penalized less than sudden bigger changes. Also, bigger changes should be penalized uniformly in order to avoid oversmoothing.

The problem of finding the optimal disparity value for each pixel can now be solved by finding the disparity image D, which minimizes the energy $E(D)$. As mentioned in HIRSCHMÜLLER (2008) this is a NP-hard problem in 2D. Therefore, the smoothness constraint is approximated by evaluating only a specified set of paths instead of a global accumulation of all possible paths. Using (12) the path cost at pixel x and disparity d into direction r can be calculated by:

$$\begin{aligned} L_r(x,d) = \; & C(x,d) \\ & + \min\left[L_r(x-r,d), L_r(x-r,d\pm 1) + \lambda_1, \min_{d'}(L_r(x-r,d') + \lambda_2)\right] \\ & - \min_{d'}[L_r(x-r,d')] \end{aligned} \tag{13}$$

where $x - r$ is the previous pixel on path r. Since the minimum path cost of the previous pixel $\min_{d'}[L_r(x-r,d')]$ is independent from d, it can be used to bound the value of $L_r(x,d)$ by subtracting it from the whole term. Without this operation, the value of $L_r(x,d)$ would increase permanently, which could lead to very large values.

Finally, the costs for one disparity at pixel x are computed. For this purpose, the costs L_r are summed over all paths r. According to HIRSCHMÜLLER (2008), the number of paths must be at least 8 and should be 16 to ensure a good coverage of the image.

$$S(x,d) = \sum_r L_r(x,d) \tag{14}$$

The optimal disparity value at pixel x can then be determined by choosing the one with the lowest overall cost $\min_d S(x,d)$.

4 Results

In this section we first describe our experimental setup before we present both qualitative and quantitative results. For this purpose, we evaluate our approach on the KITTI stereo 2015 dataset (MENZE & GEIGER 2015) to demonstrate the performance on image pairs taken under similar lighting conditions. Furthermore, we examine the robustness of our dense matcher against different simulated changes.

4.1 Experimental Setup

For the purpose of evaluation, we use the quality measurement proposed in the KITTI stereo 2015 benchmark: The disparity of a pixel is classified as erroneous if the difference between the estimated and the ground truth value is bigger than 3px and bigger than 5% of the ground truth value. As first quality measure we use the percentage of stereo disparity outliers in the first frame, whereby only pixels for which an estimate is available in the calculated disparity image are taken into account. In order to be able to evaluate the density of the disparity images, the

coverage of the ground truth pixels for which a value is available in the estimation is also determined.

Since there is currently no benchmark for image pairs with different illumination conditions, we can only compare our approach with methods whose source code is freely available. Therefore, we compare to the following approaches: GEIGER et al. (2010), ZHANG et al. (2015), CIGLA (2015) and YAMAGUCHI et al. (2014). All of them are listed on the KITTI stereo 2015 leader board.

For all experiments we use the same set of parameter values. This set was obtained by random sampling of the parameter space with 1000 iterations and evaluation of the current configuration's performance on a subset of the data. For the weights of (1) we use $\lambda_{Census} = 0.4118$, $\lambda_{PC} = 0.3564$ and $\lambda_T = 0.2317$. The parameters in (10) and (11) are selected as $\hat{y} = 10$, $\sigma = 351$, $d_0 = 82$. And the penalties of semi-global matching in (12) are set to $\lambda_1 = 22$ and $\lambda_2 = 87$. The sizes of the Census transform masks are set to 9x11 on the gray-scale image and 5×5 on the phase congruency image.

4.2 Evaluation on image pairs with similar illumination

In order to be able to make a basic statement about the performance of our approach, we first test it on pairs of images taken under similar illumination conditions. For this purpose, we decided to use the training part of the KITTI Stereo 2015 dataset proposed by MENZE & GEIGER (2015). It contains 200 stereo images taken by a mobile mapping platform and, in contrast to the test part, the ground truth disparity images are available. The ground truth is semi-dense covering about 30% of the pixels.

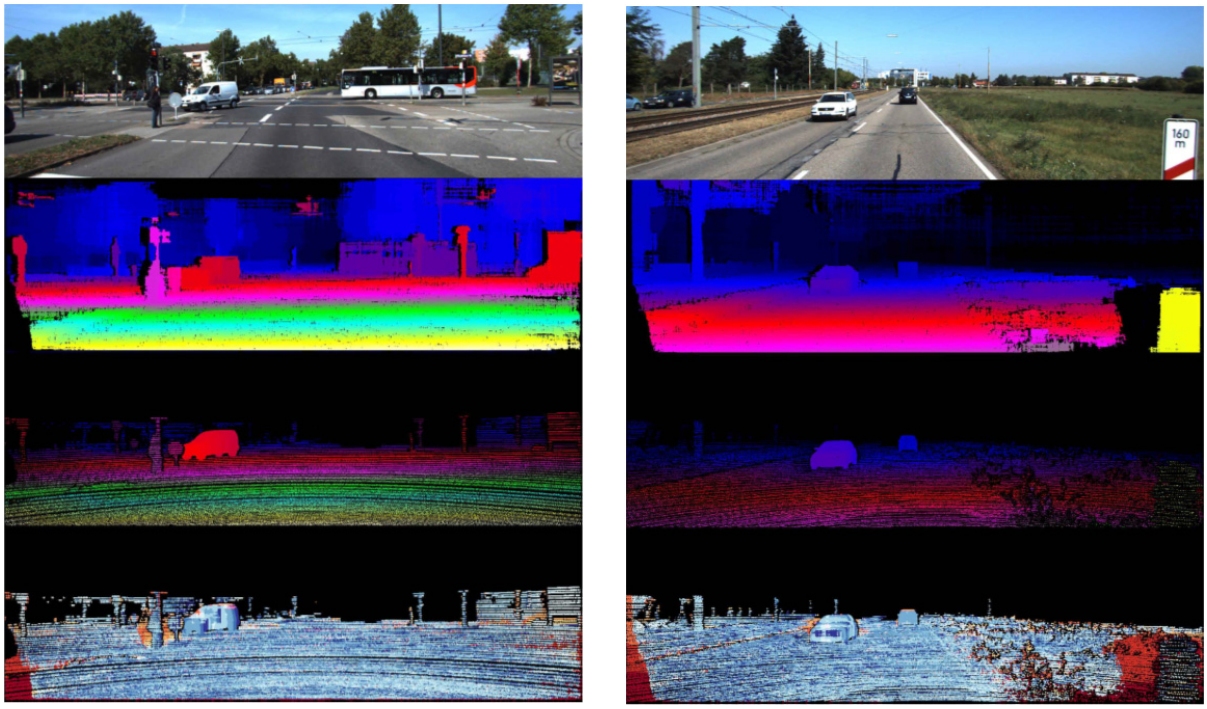


Figure 2: Qualitative results for our approach. In descending order: Left image, estimated disparity, ground truth disparity (from large disparities in yellow to small ones in blue), deviation between estimation and ground truth (from small deviations in white/light blue to large ones in orange/red).

To demonstrate the quality of the disparity images produced by our approach, Figure 2 shows two exemplary results. As one can see in the estimated disparity images (second from above) the contours are quiet sharp and only few artefacts appear. To view the quality of the results in relation to the state-of-the-art, Tab. 1 shows the averaged results for all examined approaches.

In contrast to most of the approaches on the benchmark's leader board, our approach is only semi-dense. Based on our goal to build a reliable map, it is important for us that the error contained in the estimated disparity image is as small as possible. Hence, we decided to not interpolate the disparity images for the moment to avoid the introduction of additional error. This decision results in a lower density and causes most of the differences between estimation and ground truth, which can be seen in the last images of Figure 2.

Tab. 1: Comparison with the examined approaches on the training part of the KITTI stereo 2015 dataset

Method	**Error**	**Coverage**
Elas Matcher (GEIGER et al. 2010)	5.71%	83.66%
Mesh Stereo (ZHANG et al. 2015)	7.80%	99.87%
REAF (CIGLA 2015)	13.13%	99.42%
SPS Stereo (YAMAGUCHI et al. 2014)	4.81%	99.97%
Ours	4.27%	89.00%

4.3 Evaluation on image pairs with simulated changes

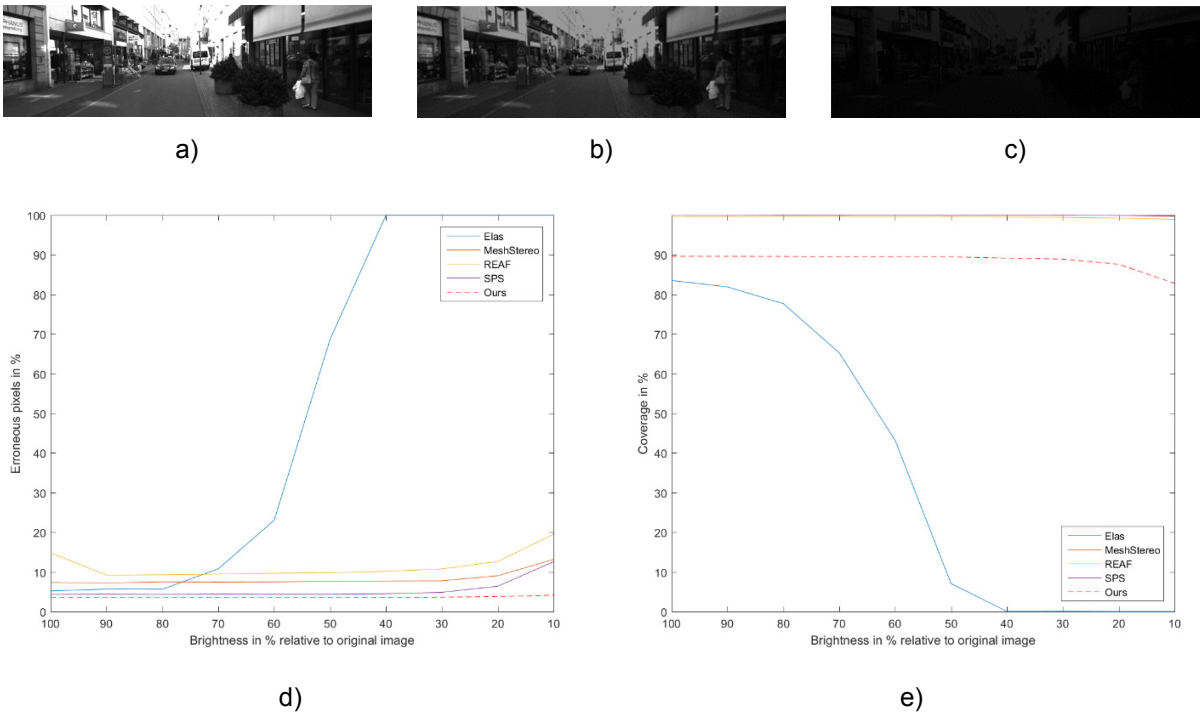


Fig. 3: Relative brightness decrease. (The brightness of the right image is reduced to the respective percentage.) a) - c) show the influence of the brightness reduction on the image at 100%, 50% and 10% of the original values. d) and e) show the effects of this change on the percentage of erroneous pixels and coverage.

In a second step, we evaluate the performance of our approach under varying conditions. For this purpose, we change the right image of a pair, while the left one is left untouched. In the following, the effects of relative brightness decrease, absolute brightness decrease and Gaussian noise are evaluated. For every type of change, the effect on the image, on the accuracy and also on the coverage is shown. The evaluation was performed on 10 different image pairs, with 10 different levels for relative brightness change, 15 different levels for absolute brightness change and 10 different levels for Gaussian noise.

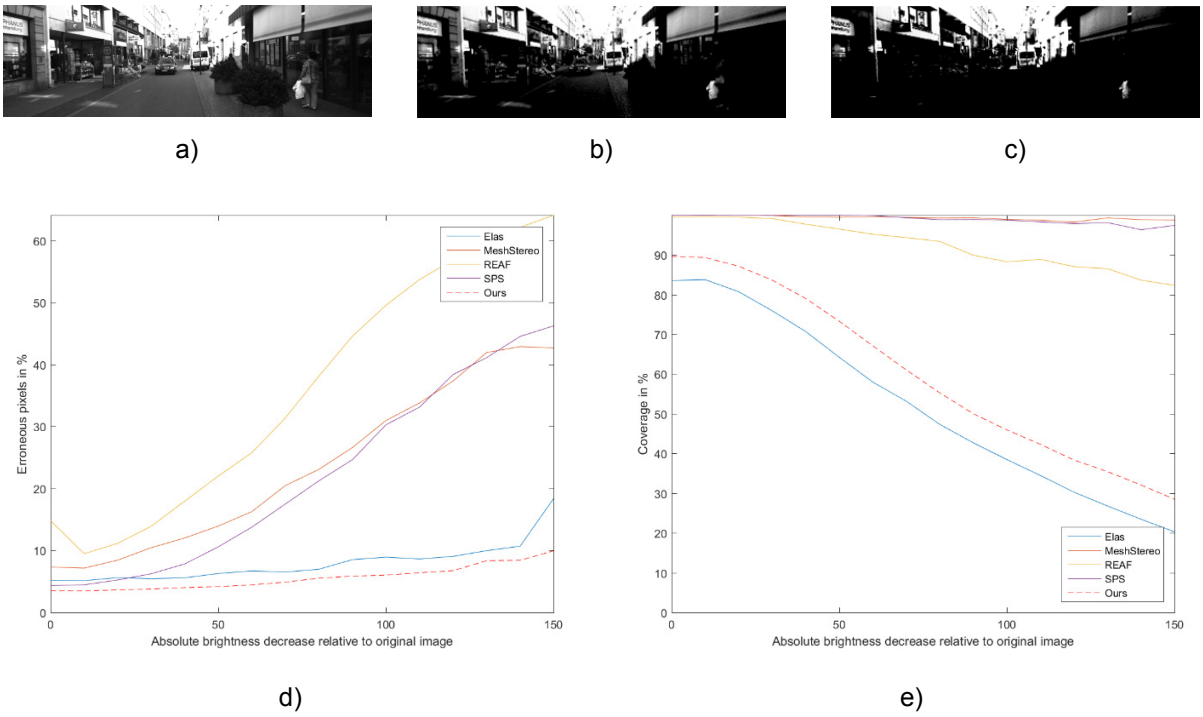


Fig. 4: Absolute brightness decrease. (The brightness of the right image is reduced by the respective value.) a) - c) show the influence on the image when the original values are reduced by 0, 80 and 150 grey levels. d) and e) show the effects of this change on the percentage of erroneous pixels and coverage.

As shown in Fig. 3 and Fig. 4 the percentage of erroneous pixels in the estimated disparity image increases much slower for our approach than for the others. In contrast, the coverage for our approach decreases comparatively quickly. This demonstrates that our approach is able to identify potentially wrong pixels up to a certain point. In ambiguous cases, the pixel is set to 0 instead of just choosing the value with the lowest energy.

A weakness of our approach can be seen in Fig. 5: The introduction of noise to one image leads to a high impact on error and coverage. Among other things, we want to address this problem in our future work.

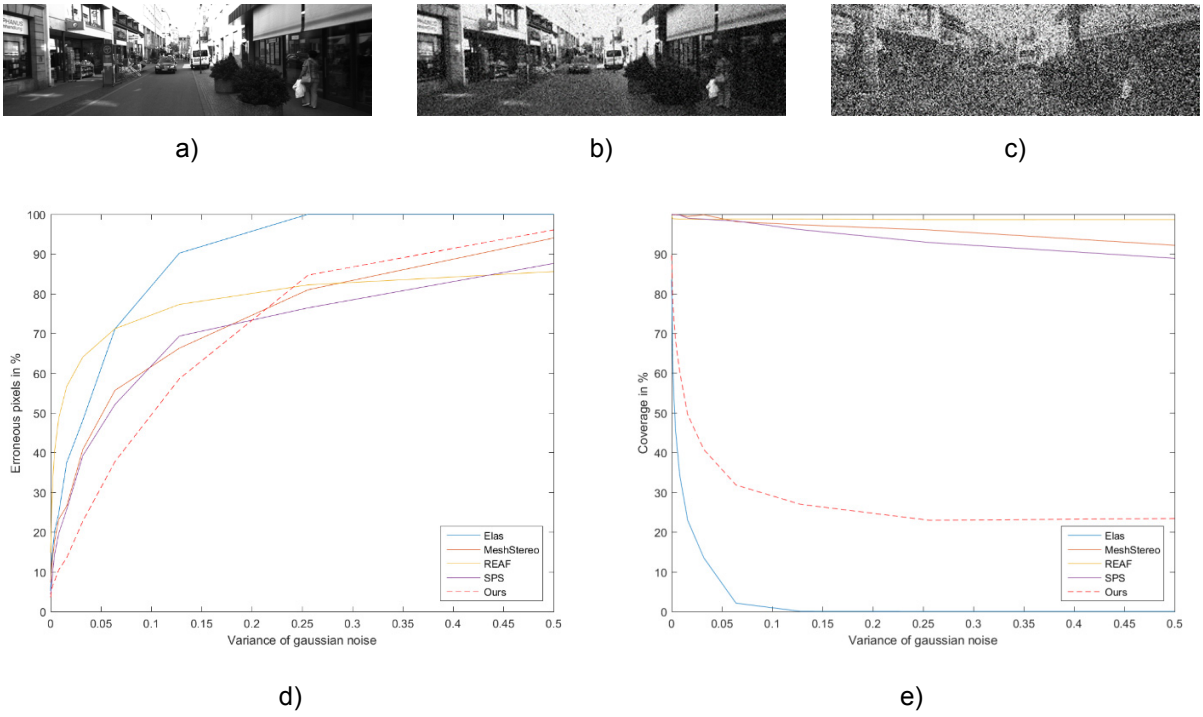


Fig. 5: Gaussian noise. (On the right picture Gaussian noise with respective variance is applied.) a) - c) show the influence on the image when applying Gaussian noise with a variance of 0, 0.25 and 0.5. d) and e) show the effects of this change on the percentage of erroneous pixels and coverage.

5 Conclusion

We have proposed an algorithm for dense image matching based on the combination of multiple cost functions to reliably reconstruct depth from an image pair. The performance of our approach was evaluated on the challenging KITTI stereo dataset, showing state-of-the-art results. Moreover, a set of images with simulated changes was used to demonstrate the robustness and the ability to detect potentially erroneous pixels. In future work, we want to investigate methods, which allow densification of the disparity image, without increasing the error. Furthermore, we believe that reliability measures for the single components within the cost function, could improve the performance under varying conditions.

6 References

BULATOV, D., WERNERUS, P. & HEIPKE, C., 2011: Multi-view dense matching supported by triangular meshes. ISPRS Journal of Photogrammetry and Remote Sensing, **66**(6), 907-918.

CIGLA, C., 2015: Recursive edge-aware filters for stereo matching. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 27-34.

COX, I.J., ROY, S. & HINGORANI, S.L., 1995: Dynamic histogram warping of image pairs for constant image brightness. Proceedings of the International Conference on Image Processing, **2**, 366-369.

GEIGER, A., ROSER, M. & URTASUN, R., 2010: Efficient large-scale stereo matching. Asian Conference on Computer Vision, 25-38.

HEO, Y.S., LEE, K.M. & LEE, S.U., 2011: Robust stereo matching using adaptive normalized cross-correlation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(4), 807-822.

HIRSCHMÜLLER, H., 2008: Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2), 328-341.

HIRSCHMÜLLER, H. & SCHARSTEIN, D., 2009: Evaluation of stereo matching costs on images with radiometric differences. IEEE Transactions on Pattern Analysis and Machine Intelligence, **31**(9), 1582-1599.

JUNG, I.L., SIM, J.Y., KIM, C.S. & LEE, S.U., 2013: Robust stereo matching under radiometric variations based on cumulative distributions of gradients, 20th IEEE International Conference on Image Processing (ICIP), 2082-2085.

KIM, S., HAM, B., KIM, B. & SOHN, K., 2014: Mahalanobis distance cross-correlation for illumination-invariant stereo matching. IEEE Transactions on Circuits and Systems for Video Technology, 24(11), 1844-1859.

KOVESI, P., 1999: Image features from phase congruency. Videre: Journal of computer vision research, **1**(3), 1-26.

LOWE, D.G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, **60**(2), 91-110.

MADDERN, W., STEWART, A., MCMANUS, C., UPCROFT, B., CHURCHILL, W. & NEWMAN, P., 2014: Illumination invariant imaging: Applications in robust vision-based localisation, mapping and classification for autonomous vehicles. Proceedings of the Visual Place Recognition in Changing Environments Workshop, IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, **2**, 3-10.

MENZE, M. & GEIGER, A., 2015: Object scene flow for autonomous vehicles. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3061-3070.

MENZE, M., HEIPKE, C. & GEIGER, A., 2015: Joint 3d estimation of vehicles and scene flow. ISPRS Workshop on Image Sequence Analysis (ISA), **8**.

MOUATS, T., AOUF, N. & RICHARDSON, M.A.,2015: A novel image representation via local frequency analysis for illumination invariant stereo matching. IEEE Transactions on Image Processing, 24(9), 2685-2700.

RAGB, H. & ASARI, V.K., 2016: Histogram of Oriented Phase (HOP): A New Descriptor Based on Phase Congruency. Electrical and Computer Engineering Faculty Publications. Paper 391.

SEIF, H. G. & HU, X., 2016: Autonomous Driving in the iCity-HD Maps as a Key Challenge of the Automotive Industry. Engineering, **2**(2), 159-162.

SKARBNIK, N., ZEEVI, Y.Y. & SAGIV, C., 2010: The Importance of phase in image processing. CCIT Report No. 773, 1-30.

VERDIE, Y., YI, K., FUA, P. & LEPETIT, V., 2015: TILDE: a temporally invariant learned detector. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 5279-5288.

YAMAGUCHI, K., MCALLESTER, D. & URTASUN, R., 2014: Efficient joint segmentation, occlusion labeling, stereo and flow estimation. European Conference on Computer Vision, 756-771.

ZABIH, R. & WOODFILL, J., 1994: Non-parametric local transforms for computing visual correspondence. European conference on computer vision, Springer, Berlin, Heidelberg, 151-158.

ZHANG, C., LI, Z., CHENG, Y., CAI, R., CHAO, H. & RUI, Y., 2015: Meshstereo: A global stereo model with mesh alignment regularization for view interpolation. Proceedings of the IEEE International Conference on Computer Vision, 2057-2065.

ZHOU, X. & BOULANGER, P., 2012: Radiometric invariant stereo matching based on relative gradients. 19th IEEE International Conference on Image Processing (ICIP), 2989-2992.

Classification of Laser Scanning Data Using Deep Learning

FLORIAN POLITZ[1], BASHIR KAZIMI[1] & MONIKA SESTER[1]

Abstract: In the last couple of years Deep Learning has gained popularity and shown potential in the field of classification. In contrast to 2D image data, Airborne Laser Scanning data is complex due to its irregular 3D structure, which turns the classification into a difficult task.
Classifying point clouds can be separated into pointwise semantic classification and object-based classification. In this paper, we investigate both classification strategies using Convolutional Neural Networks (CNNs). In the first part of this paper, we focus on the semantic classification of 3D point clouds into three classes as required for generating digital terrain models. The second part of this paper deals with classifying archaeological structures in digital terrain models.

1 Introduction

Airborne laser scanning (ALS) is a very versatile technology with many applications such as terrain modelling, hydrology or archaeology. In ALS a laser pulse measures the time a laser beam requires to travel from a remote sensing module to the earth's surface and back. 3D point coordinates can be calculated using the position and direction of the remote sensing module in a global coordinate system as well as the direction and distance of the measured laser beam. An ALS point cloud is a collection of those point measurements for a specific area and is utilised for several applications such as ground point detection, topographic mapping or object recognition.
From the raw point cloud to a final product, the point cloud undergoes several processing steps with different intermediate products. In order to generate a digital terrain model (DTM), a semantic and pointwise classification on the raw point cloud is necessary to distinguish between objects on the ground and the ground itself. Subsequently, through those classified points, a DTM can be interpolated, which is usually saved as a rasterised height map or a triangulated irregular network (TIN). Finally, height maps or TINs can be the basis for further data analysis or classification such as object recognition or object reconstruction.
In the past, traditional machine learning algorithms have been used to classify the data to the requested products. As special type of a machine learning algorithm Convolutional Neural Networks (CNNs) moved into the focus of attention due to winning at the latest image benchmark competitions such as ImageNet and PASCAL VOC for 2D image processing by a large margin (KRIZHEVSKY et al. 2012; GIRSHICK et al. 2013). In contrast to regular 2D images, 3D point clouds possess an irregular data structure, which impede using CNN for semantic or object-based point cloud classification directly.
This paper deals with pointwise and object-based laser scanning point classification using a CNN. Firstly, we briefly overview the related work in order to classify laser scanning data in chapter 2. Secondly, we introduce the CNN and its network layers in chapter 3. Thirdly, we will present and discuss our work classifying ALS data using CNNs in the experiment's section in

[1] Leibniz Universität Hannover, Institut für Kartographie und Geoinformatik, Appelstraße 9a, D-30167 Hannover, E-Mail: [Florian.Politz, Bashir.Kazimi, Monika.Sester]@ikg.uni-hannover.de

chapter 4. The experiment's section is divided in two parts. The first part deals with semantic classification of raw point cloud data. The point cloud shall be classified into the classes "ground", "non-ground" and "others" such as cars, trains and open wires in order to generate a DTM. The second part deals with object-based classification. In this approach, archaeological terrain structures and objects have to be classified from a rasterised height map. Lastly, we finish the paper with a conclusion and outlook in chapter 5 as well as our references in chapter 6.

2 Related Work

For semantic classification using CNNs, laser scanning point clouds can be classified using voxel-based, point-based or projection-based approaches. In voxel-based methods, the point cloud is discretised to equally distributed 3D voxels, which are then fed into a CNN. PROKHOROV (2010) used a 3D CNN to refine coarsely segmented point clouds of cars using a voxel-based density grid. Meanwhile MANTURANA & SCHERER (2015a, 2015b) and HUANG & YOU (2016) calculated a voxel-based occupancy grid from the point cloud as input data for a 3D CNN. While MANTURANA & SCHERER (2015a) wanted to classify if a landing zone for helicopters is safe or not, HUANG & YOU (2016) classified each point in an urban city point cloud into one of seven classes such as buildings, trees, wires and poles. In the point-based methods, the raw point coordinates and some additional information are fed into a CNN. QI et al. (2017) showed the potential of this method by feeding their network called PointNet with raw point coordinates, corresponding colour information and a normalised position.

While voxel-based and point-based methods benefit from highly dense 3D point clouds, ALS point clouds are usually in 2.5D, where the point height can be considered as attribute. This is why, in projection-based methods, an ALS point cloud is projected into 2D raster images. HU & YUAN (2016) projected airborne laser scanning point clouds into 128×128 pixels images with the normalised minimal, average and maximal height for each pixel as channel values. While HU & YUAN (2016) only distinguished between ground and non-ground classes, our network is able to classify a third class for other, artificial structures on the ground such as cars, which must be eliminated from the point cloud for the DTM generation.

In object-based classification methods, the input data can be processed with two different ways. In the first method, points are segmented into objects and then are fed into a classifier. Using this method, MALLINIS et al. (2008) aimed to delineate forest vegetation polygons in a natural forest in Northern Greece. They segmented homogeneous pixels into objects and applied a Nearest Neighbour algorithm in order to classify these merged objects into different classes. DORREN et al. (2003) also used the first method for forest mapping in steep mountainous terrains. Furthermore, they fed those objects into a decision tree algorithm to classify each object into their specific forest class.

In the second method, image patches are generated from training data by extracting digital terrain models as rasterised height maps. The most common object class within a patch decides its label. Using this method, PALAFOX et al. (2017) clipped squared grids of different sizes from the centre of a visible object to generate training data and applied those grids as an input for a classifier to detect landforms in Mars. The different sized patches are utilised to detect different sizes of the underlying phenomenon and later fed into different CNNs for training. The results of all CNNs are then aggregated to lead to a final classification. We utilise the second method in our work and feed rasterised height maps in a CNN to classify terrain structures originating from

archaeological sites. The input for the CNN are squared grids with a single channel of object heights. The model outputs a single class for the whole input grid.

3 Convolutional Neural Networks

CNN are a subgroup of general Neural Networks, which gained a lot of attention in the past couple of years due to their remarkable results in several benchmark sets. Different linear and non-linear functions build a CNN:
The linear functions can be separated into two different layer types: convolutional and fully-connected layer. A convolutional layer connects units locally using kernel matrices and it calculates a linear convolution. A convolutional layer is position-invariant, since it shares the kernel weights over the whole image. While training, the network optimizes the kernel weights as well as additional bias terms. Another linear function within a CNN is the fully connected layer (FC-layer). In FC-layers each unit of one layer is connected with all units from the previous layer leading to a very dense feature aggregation. Convolutional layers are utilised for feature extraction from the input, while FC-layers gather the most important features for classification and detection purposes at the end of CNNs.
Non-linear functions enhance the capacity of the network and allow non-linear feature aggregation. The Rectified Linear Unit (ReLU) by NAIR & HINTON (2010) is a state-of-the-art non-linear activation function for CNN. ReLU is defined as

$$f(x) = \begin{cases} max(0,x) & if\ x > 0 \\ 0 & else \end{cases}, \tag{1}$$

where x is the output of a convolutional or fully-connected layer. After ReLU only the non-negative gradients with the most valuable features remain. Compared to ReLU, LeakyReLU by MAAS et al. (2013) does not set the activation for $x \leq 0$ to zero, but scales the activation by multiplying with 0,01. As a result, the overall gradient-based optimisation algorithms improves the training's speed. Another non-linear function is max-pooling. In max-pooling two layers are connected, so that only the maximal value within a specified patch of the first layer remains in the next layer. As a result, only the most informative features persist, while spatial invariance is achieved (SCHERER et al. 2010).
Lastly, dropout regularises the network to prevent overfitting. In dropout, network units are randomly dropped while training (SRIVASTAVA et al. 2014). Consequently, in each epoch only a fraction of the network is trained, since weight and bias parameters are only updated on activated units. Therefore, the network is forced to learn more general features for each output class.
For classification purposes, the gathered features have to be transformed into discrete class numbers. In case of binary and multiclass classification, a sigmoid and a softmax function are applied respectively (BISHOP 2006).
The CNN consist of several layers: in the beginning of a CNN, convolutional layers build block structures with ReLU and max-pooling layers. The convolutional part usually finishes with a flattening of the last raster array of size n x n into a $1 \times n*n$ vector, which is necessary to connect convolutional layers with fully connected layers. For classification and detection purposes, fully connected layer with ReLU, dropout and some kind of classification function for their specific problem such as softmax are added at the end of the CNN. Figure 1 shows an exemplary CNN and its structure.

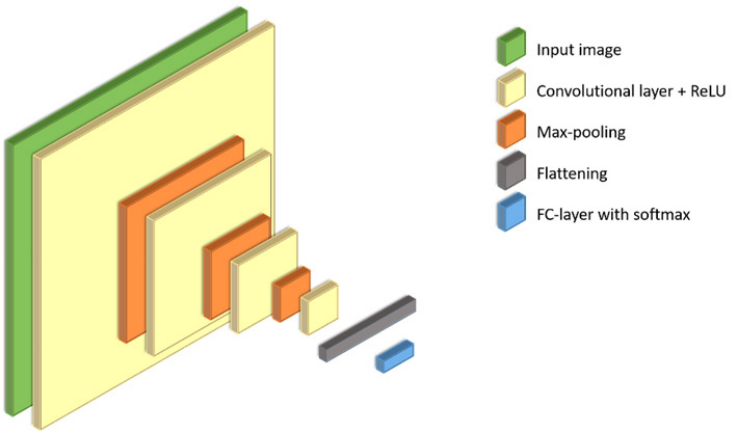


Fig. 1: Exemplary CNN structure with convolutional and ReLU layers, max-pooling, flattening and a FC-layer.

4 Experiments

4.1 Semantic Classification of airborne laser scanning point clouds

4.1.1 Input Data

For the experiments, ALS data from the survey state agency of Mecklenburg-Vorpommern (LAiV-MV) in Germany are given. The ALS data was generated in September 2012 and is located in the southeast of Rostock, Germany. The data includes a 20 km² wide area with a point density of around 4 points/m². The classes "ground", "non-ground" and "cars" are manually classified and are applied as target data for a supervised classification approach. The resulting ground points are the basis to interpolate a Digital Terrain Model (DTM). As dynamic or artificial objects such as cars are not part of a DTM, they have to be eliminated from the point cloud.

A Convolutional Neural Network (CNN) expects input data in a regular structure such as an image. For the CNN, we generate the input data similar to HU & YUAN (2016). For that reason, we project a 3D point and its environment into a 2D image structure. This is done by creating image patches around individual points, which are representing the point's environment. The patch size of 128m was chosen. Choosing such a large environment gives the network the opportunity to distinguish between large areas of flat ground and huge buildings with flat rooftops. Then, all points in that patch are sorted into 1m² raster, leading to a 128×128 pixel image. Additionally, for each raster cell within this image the minimal, average and maximal height are calculated. Finally, the height of the random point i declared as z_i is subtracted from each height value (z_{min}, z_{avg}, z_{max}) and normalized by a sigmoid function

$$Sigmoid(x) = \frac{1}{1+e^{-x}}, \tag{2}$$

where x equals $\{z_{min}, z_{avg}, z_{max}\} - z_i$. The resulting 2D structure is a 128 pixel × 128 pixel image with three channels (normalised minimal, average and maximal height values) for each point in the point cloud. The process chain from the point cloud to the input data is shown in figure 2.

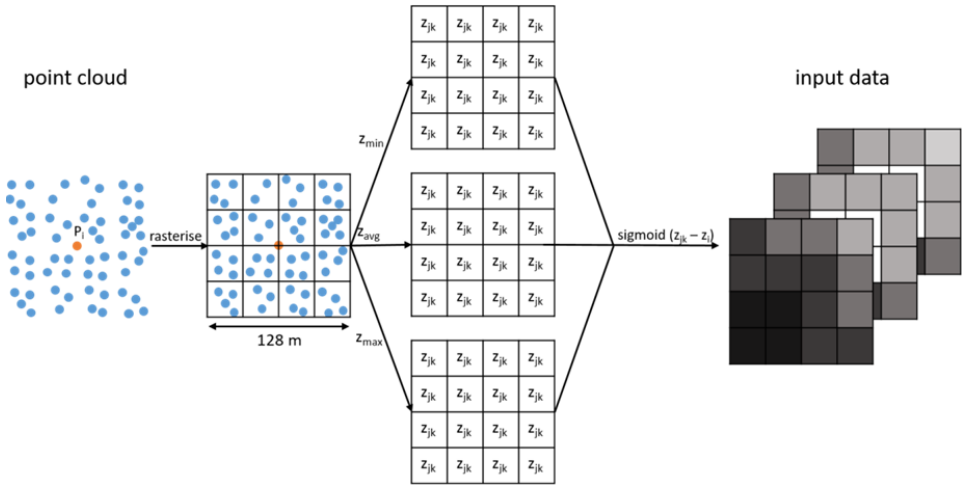


Fig. 2: Process chain to generate rasterised input data from an irregular point p_i and its environment within a point cloud.

4.1.2 Network structure

We adopt a similar CNN to the VGG16 network of SIMONYAN & ZISSERMAN (2015). In 2014, the VGG16 network achieved first and second place in the ImageNet localisation and detection challenge respectively. Since the detection challenge distinguishes 1000 classes, the capacity of the network is unnecessarily huge for our task to classify three classes. In order to prevent overfitting, we only rebuilt the first four blocks of VGG16 including several convolutional layers with ReLU and Max-Pooling at the end of each block. The kernel size for the convolutional layers is 3x3 and 2x2 for max-pooling layers following VGG16. After the last max-pooling, the aggregated features maps are flattened and two fully-connected layers with dropout are added to the network. A LeakyReLU is applied in between the FC-layers in order to fasten the training process while maintaining all gradients. At the end, a softmax function distinguishes the resulting features into the classes “ground”, “non-ground” and “cars”. Additionally, the input size is set to 128×128 pixel in comparison to the original VGG16 size of 224×224. The complete network structure is shown in table 1. Convolutional layers are abbreviated as “conv”.

4.1.3 Experiments

For the experiments, we split the point cloud data into three separate regions, leading to training, validation and testing data. The training data and their corresponding labels optimise the parameters from the CNN during training. While training, the validation data verifies the network’s capacity to generalise the data in order to prevent overfitting. Finally, the test data tests the optimised network independently from the training process. For the training region 126,000 points (42,000 points of each class) and for the validation and testing regions 18.000 points (6000 points of each class) are randomly picked from their respective regions. The ratio of training: validation: testing is 7:1:1.

Tab. 1: Overview of the network structure

Layer Name	Size (width x length x channel)	Number of Parameters
Input	128 x 128 x 3	0
Conv Block 1		
- Conv and ReLU	128 x 128 x 64	1,792
- Conv and ReLU	128 x 128 x 64	36,928
- Max-pooling	64 x 64 x 64	0
Conv Block 2		
- Conv and ReLU	64 x 64 x 128	73,856
- Conv and ReLU	64 x 64 x 128	147,584
- Max-pooling	32 x 32 x 128	0
Conv Block 3		
- Conv and ReLU	32 x 32 x 256	295,168
- Conv and ReLU	32 x 32 x 256	590,080
- Conv and ReLU	32 x 32 x 256	590,080
- Max-pooling	16 x 16 x 256	0
Conv Block 4		
- Conv and ReLU	16 x 16 x512	1,180,160
- Conv and ReLU	16 x 16 x512	2,359,808
- Conv and ReLU	16 x 16 x512	2,359,808
- Max-pooling	8 x 8 x512	0
- Flattening	1 x 1 x 32768	0
Dropout 1	1 x 1 x 32768	0
FC-layer 1	1 x 1 x 256	8,388,864
LeakyRelu	1 x 1 x 256	0
Dropout 2	1 x 1 x 256	0
FC-layer 2 (with softmax)	1 x 1 x 3	771
	Total	16,024,899

Additionally, a few hyperparameters have to be set for training the network. The specific parameters divide into network related parameters and optimisation related parameters. The network parameters include the amount of units in the first fully connected layer (FC-Layer 1, check table 1) and the drop rate parameter in the dropout layer. The amount of units in FC-Layer 1 decides about the operating network capacity. In case of fully connected layer, the parameter amount calculates as input_channel*output_channel + output_channel for weights and bias respectively. In case of 256 output_channels for FC-layer 1 as shown in table 1, the amount of necessary parameters for this layer reaches over 50% of the total training parameters. The more output units in FC-layer 1, the more the network capacity is increasing, but also the more training parameters are necessary to compute. The drop rate describes how many units are randomly dropped out during training, f.e. using a dropout rate of 0.75 leads to 75% dropped out units in this layer. The higher the drop rate, the sparser is the network while training, which forces the network to learn more general features rather than the training set itself. While testing, dropout is not applied to the network. The optimisation hyperparameter relate to all hyperparameters, which describe the optimisation process during training. The training was executed with a batch size of 50 using the cross entropy as loss function and categorical accuracy as metric. The Adam optimiser is utilized as optimising function (KINGMA & BA 2015). The optimiser consists of different parameters such as learning rate, β_1, β_2, ε and decay. For this experiment, we explore different learning rate values while keeping all other parameters for Adam as mentioned in the original paper by KINGMA & BA (2015). Additionally, we explore different parameter values and their influence towards the resulting classification accuracy.

During training, in each epoch 50 input images are used to update the parameter values. Each input image is fed into the network during training only once. Therefore, the training process lasts 2520 epochs to process all input images. The Keras API (KERAS-TEAM 2018) provides pre-trained VGG16 weights for the ImageNet data set. We initialised the weights of the convolutional layers with the provided weights and are kept untrainable for the first 20 epochs of training. Consequently, those weights do not lose their feature learning ability while the parameters in the fully connected layers are focused on classifying those extracted features from the convolutional layers to the target classes. After the first 20 epochs, all weights are included in the updating step to approach the given data.

After training, the classification ability of the network can be tested by classifying the test data set. We trained several networks with different parameter sets in order to validate the influence of each parameter and their respective classification accuracy score. Each parameter is tested separately, while the not observed parameter values within each test are set to the following values: number of units in FC-layer 1 256, dropout rate 0.75 and learning rate 0.0001.

The resulting accuracy scores for every tested parameter are shown in table 2. The first column shows the accuracy scores when the amount of units in FC-layer 1 is changed. 512 and 256 units achieve around 3% better results than networks with only 128 or even 1024 units for this layer. Although the resulting network of 512 units achieves a slightly better accuracy score, this layer also consists of 16,777,728 parameters, which have to be optimised, compared to the 8,388,864 parameters for the network with only 256 units in this layer as shown in table 2. For different drop rates in the second column in table 2 the results are not as clear as for the number of units in the fully connected layer. The accuracy score here only differs within 2% for drop rates between 0.70 and 0.85. Despite this trend, when the drop rate is set to 0.90, the accuracy score decreases for more than 5%. Such a drop indicates that the network is no longer be able to achieve correct results, because its structure is too sparse. For the learning rate, a value in the interval between 0.0005 and 0.00005 seems sufficient according to the third column in table 2 with difference of 2% in the accuracy score. Choosing a learning rate of 0.00001 results in an accuracy score of only 84,46%. In this case, the amount of training epochs is simply not enough for the low learning rate to reach the global minimum in only 2520 updates.

Tab. 2: Accuracy scores for different parameter values

Amount of units in FC-layer 1		Drop rate		Learning rate	
Parameter values	Accuracy score	Parameter values	Accuracy score	Parameter values	Accuracy score
1024	86.82	0.70	89.83	0.0005	89.03
512	90.26	0.75	90.00	0.0001	91.14
256	90.00	0.80	88.71	0.00005	89.31
128	85.23	0.85	90.42	0.00001	84.46
		0.90	84.23		

In table 3 the confusion matrix between the classes ground, non-ground and cars for the best network classification using 256 units in the FC-layer 1, a drop rate of 0.75 and a learning rate of 0.0001 are shown. This network achieves an overall accuracy score of 91,14%. The network classified around 94% of the classes ground and cars correctly, but only 84,13% non-ground points were classified correctly. 10,95% of all non-ground points were wrongly classified as ground points.

Tab. 3: Confusion matrix for 18,000 randomly picked test points [%]. Each class is equally distributed.

		Predicted label		
		Ground	Non-ground	Cars
True label	Ground	94.37	3.47	2.17
	Non-ground	10.95	84.13	4.92
	Cars	0.07	4.35	94.92

To identify the location of these misclassifications, we tested our CNN on a 250×250 m² area within the 20 km² point cloud, which have not been used for any previous training, validation or testing of the CNN. In comparison to the randomly picked testing data with an equal class distribution, we picked a test region with a highly unbalanced class distribution. The test region also consists of a challenging terrain with steep slopes as shown in figure 3. Despite the given challenges, the CNN achieved an overall accuracy of 88,41% for the test region.

Fig. 3: Orthophoto of the 250 x 250 m² test region in the southeast of Rostock, Germany in September 2014. The Orthophoto resembles the region, but the ALS data is recorded in September 2012.

Figure 4 shows the misclassifications between ground and non-ground classes. The misclassification of ground as non-ground lies mostly in the forest area in the northwest and in the mound in the southwest. The latter is caused by the lack of training data with mounds. The misclassifications within a forest are to be expected for two reasons. First, most of the ALS points within forest areas are at the first pulse on the treetops, so that points on the ground are sparse. Consequently, the ground cannot be represented correctly by z_{min} in the input images. Second, there is no guarantee, that points from last pulse are representing the ground, leaves or some undergrowth, so even for the manual classification, ground points within a forest are considered challenging. The same explanation applies for the misclassification of non-ground as ground in the forest area. Additionally, there are misclassifications on the field in the east as well as a steep slope in the southeast. The data set is recorded in September, where the fields are usually mowed down and ploughed resulting in a very rough ground structure. Consequently, the differentiation between ground and stubble field is not clear. Steep slopes are always a challenge in DTM generation. Additional training data with steep slopes will help the CNN to learn this special kind of ground better.

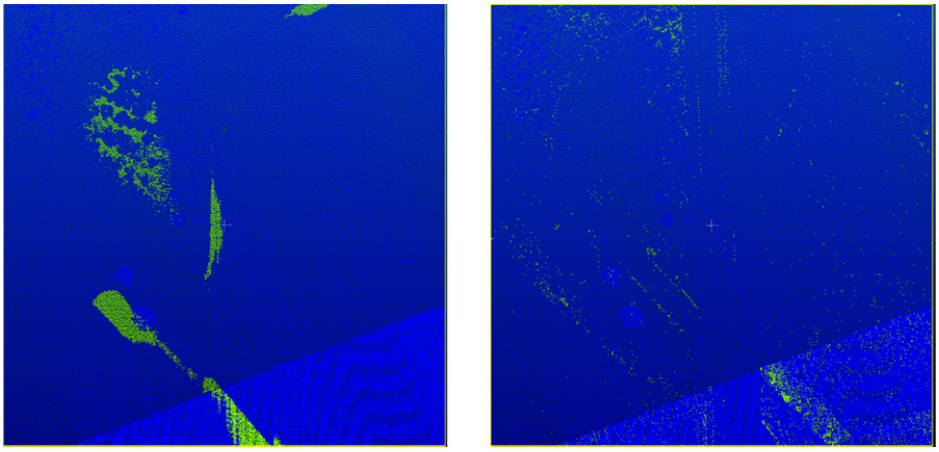


Fig. 4: Misclassifications from our CNN on the test region. Misclassifications are colour-coded as green, while the remaining classification results are blue. (left) The CNN classified ground as non-ground. (right) The CNN classified non-ground as ground.

4.2 Object-based Classification

4.2.1 Input Data

The goal of the second project is to determine archaeological terrain structures such as hollow ways, pits, heaps, shafts, sink holes, but also historical roads and water bodies in DTMs. The data set is a DTM acquired from Harz Mountains region in Lower Saxony, Germany. In our first experiments, we concentrated on two object classes, which are easily distinguishable in DTMs. To generate training data for our network, we labelled known water bodies and roads in the data set in an automatic fashion by intersecting the DTM with GIS data containing these features. Afterwards, we clipped 100×100 pixels images sampled in regular distances on the GIS data, where each pixel represents the elevation value. Figure 5 illustrates the DTM with the overlay of GIS data. The clipping process for generating training data and example inputs for the model are shown in figure 6. The images are normalised according to

$$i' = \frac{i-\mu}{\sigma}, \tag{3}$$

where $i'$is the normalized value for each pixel, i is the previous (original) pixel (elevation), μ is the mean elevation for the whole dataset and σ is the standard deviation for the whole dataset. Normalisation is necessary for the model to be able to learn from patches in one DTM and to identify features in another DTM. Without normalisation, the learnt features are specific to the training data and not transferable to any new data.

The network classifies water bodies and roads. In total, 36,248 images are used in training. 18,927 of those are road images, while the remaining 17,321 images show water bodies. Around 80% of the data is utilized for training and 20% for testing.

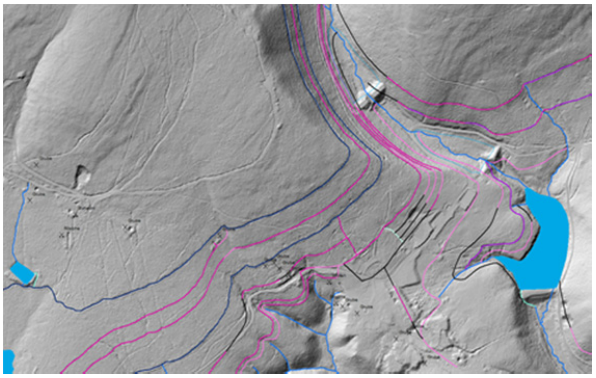


Fig. 5: Overlay of roads and water bodies on top of DTM.

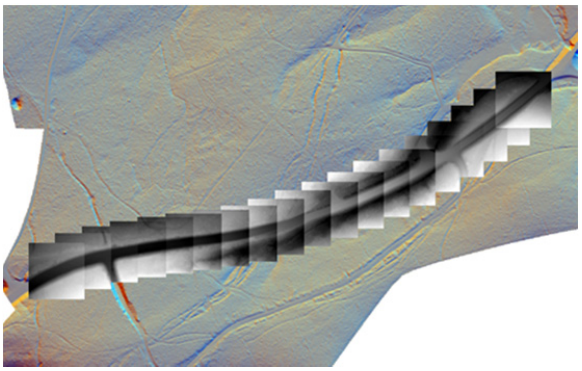


Fig. 6: Clipped roads to create labelled test data.

Table 4 summarizes the characteristics of the data set used in this experiment.

Tab. 4: Data set used for object-based classification

	Dataset 1: DTM data
# training images	28,997
# testing images	7,251
Input size	100 x 100 x 1
# classes	2 (water bodies & roads)

4.2.2 Network Structure

The CNN model for object-based classification takes a height map of size N x N as input and outputs a single label for the entire input image. The model consists of three blocks with a convolutional layer, ReLU and a max-pooling layer. Another convolutional layer and ReLU follows these blocks. The model finishes with a fully connected layer and a softmax function to classify the aggregated features into the respective classes. Table 5 summarizes the structure of the model for the data set. Convolutional layers are abbreviated as “conv”.

Tab. 5: Overview of the network structure for object-based classification

Layer Name	Size (width x length x channel)	Number of Parameters
Input	100 x100 x 1	0
Conv and ReLU	100 x 100 x 16	592
Max-pooling	50 x 50 x 16	0
Conv and ReLU	50 x 50 x 32	115,232
Max-pooling	25 x 25 x 32	0
Conv and ReLu	25 x 25 x 16	12,816
Max-pooling	12 x 12 x 16	0
Conv and ReLu	12 x 12 x 8	2,056
Flattening	1 x 1 x 1152	0
FC-layer (with softmax)	1 x 1 x 2	2,306
	Total	133,002

4.2.3 Experiments

For training the CNN, we also chose the loss function as the categorical cross entropy and Adam as optimiser (KINGMA & BA 2015). The batch size is 32 and the learning rate is 0.001. After completing the training process, the model was tested. The CNN achieves a test accuracy of 75%. To provide a better insight into the results, we list the confusion matrix for the test set in table 6.

Tab. 6: Confusion matrix for test data [%]

		Predicted label	
		Water bodies	Roads
True label	Water bodies	72.45	27.55
	Roads	22.68	77.32

As observed in table 6, 72.45% of water body images were classified correctly while the remaining 27.55% were misclassified as roads. In case of road images, 77.31% were classified correctly and 22.68% were misclassified as water bodies. There are two main reasons for the misclassifications. First, the training data set is small and deep neural network models usually require many training examples. The second reason for the misclassification is probably in the way the training images have been clipped. While clipping grids of 100×100 pixels from the DTM, some of the examples of water bodies might contain roads, and vice versa. Figure 7 shows some example inputs and their predictions by our model.

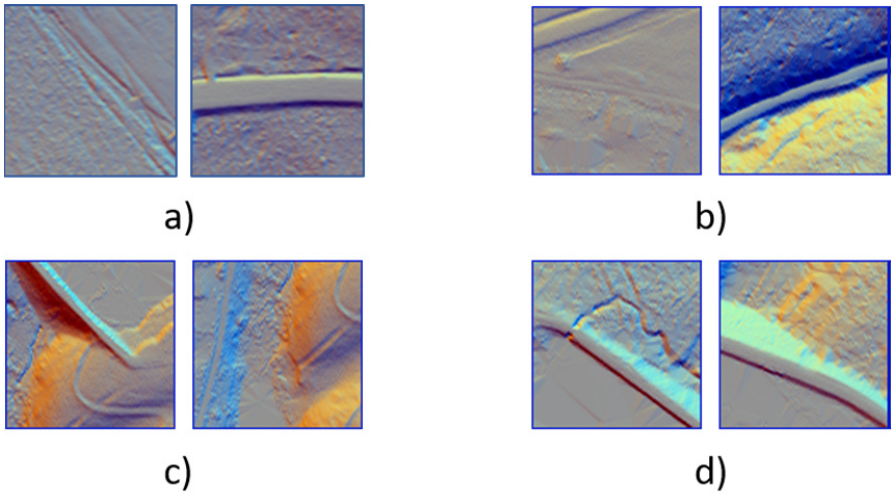


Fig. 7: Example inputs and predictions by our model. The input size is 100 x 100 pixel. a) Road input classified correctly as road. b) Road input classified incorrectly as water body. c) Water body input classified correctly as water body. d) Water body input classified incorrectly as road.

5 Conclusion and Outlook

In this research, we explored and discussed different methods to classify airborne laser scanning. In the first method, each point in the point cloud is classified. The second method deals with object classification using digital terrain models, which are a derived product from airborne laser scanning point clouds. A convolutional neural network (CNN) was applied as classifier in both methods. A CNN requires a regular structure for their input, which irregular point clouds from airborne laser scanning does not provide.

For this reason, the 3D point clouds were projected into 2D images in the first method. For each point, the normalized minimal, average and maximal height is taken into account as input images. The trained network classifies the point cloud into the classes ground, non-ground and car and achieves an overall accuracy of 91,14% while testing, which is sufficient for practical use to derive a digital terrain model. From the scientific point of view, this is not enough. As mentioned in experiments in chapter 4.1.3 there are still some errors left in the trained network resulting in around 10% of the non-ground points being misclassified as ground points. In future work, we will reduce those misclassifications. Equally, we will increase the number of classes in our network in order to distinguish buildings from vegetation in the non-ground class as well as streets from simple ground points in the ground class.

The object-based classification also achieves reasonable results. Only by considering the elevation values in a grid containing the whole object in our data set, the model learns to identify objects with an accuracy of around 75%. In summary, the object-based classifier implemented in this research proves to be appreciable. With more training data, the model is expected to learn better and gain a higher accuracy. To this end, additional training data will be included from other areas; furthermore, also perturbations of the existing training data (e.g. rotation or scaling) will be applied to create more data with a higher variation. Additionally, we will extend the model to detect more classes such as hollow ways, pits, heaps, shafts, sink holes, etc. After classification, the next task is to delineate the object geometry and create 3D object models.

6 References

BISHOP, C., 2006: Pattern Recognition and Machine Learning. Springer Science+Business Media, LLC. Singapore. ISBN 978-0-387-31073-2, 198.

DORREN, L.K.A., MAIER, B. & SEIJMONSBERGEN, A.C., 2003: Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. Forest Ecology and Management 183 (1-3), 31-46.

GIRSHICK, R., DONAHUE, J., DARRELL, T. & MALIK, J., 2013: Rich features hierarchies for accurate object detection and semantic segmentation. Extended version of the published conference paper at CVPR 2014. arXiv: 1311.2524v5.

HU, X. & YUAN, Y., 2016: Deep-Learning-Based Classification for DTM Extraction from ALS Point Cloud. Remote Sens. 2016 (8), 730.

KERAS-TEAM, 2018: Keras: Deep Learning for humans. GitHub repository, https://github.com/keras-team/keras

KINGMA, D.P & BA, J.L., 2015: Adam: A method for stochastic optimization. Published as a conference paper at ICLR 2015. arXiv:1412.6980v9.

KRIZHEVSKY, A., SUTSKEVER, I., & HINTON, G.E., 2012: ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information (NIPS) 2012 (1), 1097-1105.

MAAS, A., HANNUN, A.Y. & NG, A.Y, 2013: Rectifier Nonlinearities Improve Neural Network Acoustic Models. ICML'13 Proceedings of the 30th International Conference on Machine Learning. JMLR: W&CP (28).

MALLINIS, G., KOUTSIAS, N., TSAKIRI-STRATI, M. & KARTERIS, M., 2008: Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. Published in ISPRS Journal of Photogrammetry and Remote Sensing, **63**(2), 237-250.

MANTURANA, D. & SCHERER, S., 2015a: 3D Convolutional Neural Networks for Landing Zone Detection from LiDAR. Published in IEEE International Conference on Robotics and Automation (ICRA) 2015, 3471-3478.

MANTURANA, D. & SCHERER, S., 2015b: VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition. Published in IEEE International Conference on Intelligent Robots and Systems (IROS), 922-928.

NAIR, V. & HINTON, G. E. 2010: Rectified Linear Units Improve Restricted Boltzmann Machines. ICML'10 Proceedings of the 27th International Conference on Machine Learning, 807-814.

PALAFOX, L. F., HAMILTON, C. W., SCHEIDT, S. P., & ALVAREZ, A. M., 2017: Automated detection of geological landforms on Mars using Convolutional Neural Networks. Computers and Geosiences, **101**, 48-56.

PROKHOROV, D., 2010: A convolutional Learning System for Object Classification in 3D Lidar Data. IEEE Transactions on Neural Networks, **21**(5), 858-863.

SCHERER, D., MÜLLER, A. & BEHNKE, S., 2010: Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition. International conference on artificial neural networks. Springer, Berlin, Heidelberg, 92-101.

SIMONYAN, K. & ZISSERMAN, A., 2015: Very Deep Convolutional Networks For Large-Scale Image Recognition. Published as a conference paper at ICLR 2015. arXiv:1409.1556v6.

SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I., & SALAKHUTDINOV, R., 2014: Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research, **15**, 1929-1958.

QI, C.R., SU, H., MO, K. & GUIBAS, L.J., 2017: PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. Proc. IEEE Computer Vision and Pattern Recognition (CVPR), **1**(2), 4, arXiv: 1612.00593v2.

Object Proposals for Pedestrian Detection in Stereo Images

UYEN D. X. NGUYEN[1], FRANZ ROTTENSTEINER[1] & CHRISTIAN HEIPKE[1]

Abstract: Pedestrian detection is an active research field in computer vision and photogrammetry today due to its importance for applications related to autonomous driving, machine human interaction, and surveillance. Object proposals play a significant role in guiding a classifier where to examine image regions that may contain pedestrians. In this paper, we present a proposal framework employing both 3D information derived from stereo images and RGB cues to generate pedestrian bounding box proposals with high recall and relatively small number of candidates. Our proposal generator has two stages: (1) generating a large number of proposals to achieve a good recall value; and (2) eliminating unreasonable initial candidates based on 3D cues and pedestrian geometric constraints. Preliminary experiments on the Kitti benchmark dataset show that our proposal framework is comparable to state-of-the-art methods.

1 Introduction

Today, pedestrian detection, a specific case of object detection, is one of the most vigorous research areas in computer vision and photogrammetry, especially in applications related to autonomous driving, robotics, safety, surveillance, etc. With the support of these systems, human effort in processing huge amounts of images, which is time-consuming, expensive and subjective, can be reduced or completely avoided.

In recent years, Convolution Neural Networks (CNNs) have undergone a dramatic development, which enabled object detection to become more and more accurate and reliable (REDMON et al. 2016; GIRSHICK 2015; REN et al. 2015; GIRSHICK et al. 2014). However, every CNN object detector depends on a region proposal method to determine possible locations in an image where desired objects can appear. A typical paradigm for this is exhaustive search in the entire image with sliding windows of multiple scales (HARZALLAH et al. 2009; VEDALDI et al. 2009). This approach is easy and simple, but results in a very large number of candidates, which makes it computationally expensive to adopt a complex classifier for reliable recognition. Recently, algorithms were suggested which exploit low level features to generate a set of proposals to reduce the number of regions which have to be examined more closely (ALEXE et al. 2012; UIJLINGS et al. 2013; CHENG et al. 2014; ZITNICK & DOLLÁR 2014). These algorithms proved their effectiveness on different datasets like PASCAL VOC (EVERINGHAM et al. 2010) or ImageNet (DENG et al. 2009). However, for a dataset containing objects with large scale variation, occlusion, and truncation like the Kitti benchmark for autonomous driving (GEIGER et al. 2012), most of these methods do not achieve a satisfactory recall with a relatively small number of proposal bounding boxes (HOSANG et al. 2016).

In this work, with the focus on generating object proposals for pedestrians using a stereo camera, we tackle the problem of region proposals in two steps, and we aim to obtain a high recall with a

[1] Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Str. 1, D-30167 Hannover, E-Mail: [nguyen, rottensteiner, heipke]@ipi.uni-hannover.de

limited number of proposals. First, a large number of proposals is produced in an initial step to recover a high recall. Then, a filtering step eliminates proposed regions that most certainly do not contain pedestrians. For that purpose, we employ geometric constraints relating to pedestrians and scene context information, and we re-rank all bounding boxes produced from the first step with a scoring function. The candidates with the highest scores are then selected for further analysis and classification. Two main assumptions are made about pedestrians: they all stand on the ground and have restricted size. Further assumptions concern the scene, which we take as being composed of a horizontal ground plane and a number of objects with primarily vertical extent. To demonstrate the efficiency of our approach, we conduct experiments on the Kitti dataset. The performance of our region proposal is analysed and compared to other state-of-the-art methods.
The remainder of this paper is arranged as follows: Section 2 reviews related works. Section 3 presents the methodology and algorithms used in our work. Section 4 illustrates our experimental results, followed by the conclusion in Section 5.

2 Related Works

In this chapter we discuss previous work concerning region proposal methods. It is the aim of a proposal generator to determine all instances of interesting objects in an image with as few false alarms as possible, delineating them, e.g., by bounding boxes. The result allows more sophisticated and accurate classifiers like CNNs to be adopted to detect the objects. Currently, several paradigms are widely applied for the region proposal task.
Objectness scoring: this approach uses a sliding window in multiple scales to examine every area in an image. Each region is then assigned an "objectness" score based on combining a number of cues such as normalised gradients, structured edges, texture, etc. Then the regions with the highest score are selected as proposal candidates. Several well-known algorithms in this group are objectness (ALEXE et al. 2012), edge boxes (EB) (ZITNICK & DOLLÁR 2014), BING (CHENG et al. 2014). EB is one of the most successful methods in this group, which efficiently evaluates a proposal candidate based on the contours located inside and on the edge of the bounding box.
Super pixel grouping: representative proposal generators of this group are selective search (SS) (UIJLINGS et al. 2013) and multiscale combinatorial grouping (MCG) (ARBELAEZ et al. 2014). This technique exploits a hierarchical segmentation to obtain a set of super pixels, which are subsequently merged and scored by a ranking function. SS can deliver high quality proposals and is widely used by many detectors (WANG et al. 2013; GIRSHICK et al. 2014) due to its comparably short execution time and high recall. However, the results are sensitive to large size variation, occlusion, and truncation of objects. In addition, small objects lead to the requirement to select small super pixels, which can lead to high computation cost.
3-D object proposal: the methods mentioned above mostly focus on solving the object proposal task based on RGB images. However, depth information can also provide a valuable cue for evaluating candidate regions. A few methods like MCG-D (GUPTA et al. 2014) consider depth information to improve performance. But their complicated scoring functions make the method rather slow. Using depth cues and energy minimization to infer potential objects with 3D boxes,

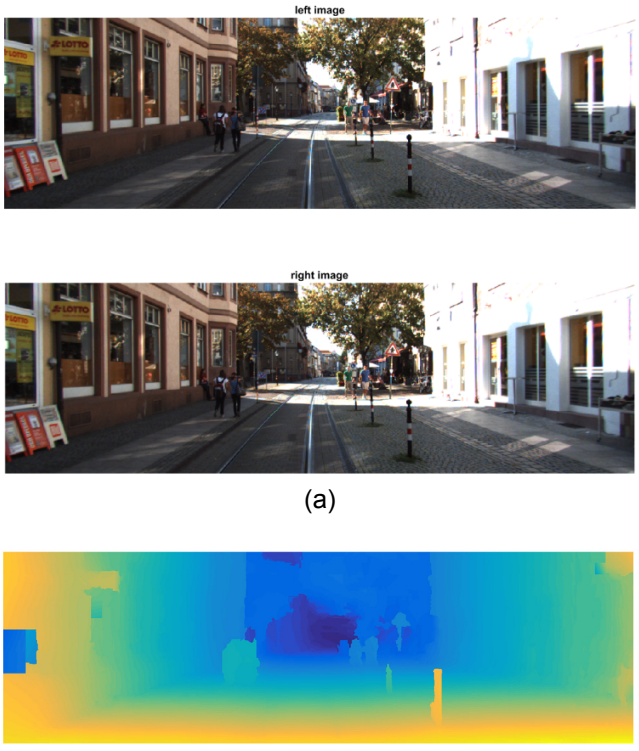
CHEN et al. (2015) can achieve a good recall value for the Kitti benchmark. Nevertheless, considering only 3D position and dismissing RBG information is not necessarily a good idea because many object features can only be distinguished in RGB images.

3 Methodology

In this study, we aim to develop a pedestrian proposal framework. After appropriate 3D scene modelling we generate initial proposals based on RGB image information using the EB method (ZITNICK & DOLLÁR, 2014). Then, we use 3D information obtained from the stereo image pair and geometric constraints corresponding to pedestrians to filter out unreasonable bounding boxes.

3.1 Modelling the Scene

Given a stereo image pair with known orientation parameters, the disparity map D of all pixels with respect to the left image is first estimated using a state-of-the-art dense matching approach (YAMAGUCHI et al., 2014). Then, a point cloud P is computed using the disparity values (Fig. 1).



(a)

(b)

Fig. 1: The stereo image pair (a) is used to compute the disparity map (b).

We assume our scene to be mainly composed of a more or less horizontal ground plane (e.g. a road), vertical planes (e.g. building facades) and the sky (which we don't consider further). In addition, other objects such as pedestrians are present, which are our objects of interest. Many objects in an urban scene can be considered as vertical planar surfaces supported by the ground

plane. Reconstructing these vertical objects and the ground plane in object space would provide additional evidence for pedestrian detection.

3.1.1 Potential obstacles

We define potential obstacles, i.e. non-pedestrian objects, as regions in an image which are perpendicular to the ground plane. Under this definition, a number of pixels in each image column that have the same disparity value belong to an obstacle if we assume images to be taken with horizontal and parallel optical axes. A binary obstacle mask for each input image is estimated using the following steps proposed by HU & UCHIMURA (2005):

Step 1 From the disparity map D, the vertical or v-disparity image V_{dis} is computed such that each column in V_{dis} is a disparity histogram of the corresponding column in D.

Fig. 2: v-disparity image estimated from the disparity map shown in Fig. 1b.

Step 2 The intensity of a pixel $v(x, y)$ in V_{dis} represents the number of pixels in column y of D that have approximately disparity x. Hence, the binary obstacle mask O_{mask} can be built by finding pixels in the disparity map D, which correspond to pixels with the entry in V_{dis} larger than a threshold value. Those pixels in D are considered as obstacle regions in O_{mask}.

Fig. 3: The binary obstacle mask is built from the v-disparity image; obstacle pixels are shown in white.

Step 3 Morphological closing is used to join small obstacle regions together. Then, all none-obstacle regions smaller than a threshold are considered to be caused by errors and, thus, are included as obstacle areas.

In fact, we produce two obstacle masks O_{mask1} and O_{mask2} by applying two different thresholds (Fig. 4). Based on a small threshold, O_{mask1} contains nearly all vertical objects in the scene. In contrast, O_{mask2} is computed with a higher threshold, so small objects (e.g. pedestrians) should not be part of this mask.

3.1.2 Ground plane extraction

Unlike vertical object pixels, ground plane pixels should have similar depths per row. Using this assumption, the ground plane is estimated as follows:

Step 1 We use the obstacle mask O_{mask1} to eliminate most pixels related to vertical objects in the disparity map, so that the remaining pixels in a new disparity map:

$$D_{nonobs} = O_{mask1} .* D \ , \tag{1}$$

mostly belong to the ground plane. In (1), $.*$ denotes a pixel-wise multiplication of the grey values.

Step 2 Ground pixels in D_{nonobs} are determined in a similar way as the obstacle mask. However, instead of using the v-disparity image, we compute the disparity histogram of each row in D_{nonobs} to generate a horizontal or h-disparity image. As a result, ground pixels in D_{nonobs} with their 3D positions are collected in a set $P_{ground} = \{p_1(\mathrm{x},\mathrm{y},\mathrm{z}), \dots, p_n(\mathrm{x},\mathrm{y},\mathrm{z})\}$.

Step 3 Any 3-D point $(\mathrm{x},\mathrm{y},\mathrm{z})$ lying on the ground plane (Ω) must satisfy the planar equation

$$(\Omega): a\mathrm{x} + b\mathrm{y} + c\mathrm{z} + d = 0, \tag{2}$$

where (a, b, c) is the normal vector with length 1 and d is the distance from the origin to (Ω).

The ground plane (Ω) is determined using the 3-D ground points P_{ground}, together with RANSAC to remove outliers. Then, the ground pixels in the image which form a ground mask G_{mask} are determined as those which have an absolute distance in object space to (Ω) smaller than a value ε.

(a)

(b)

Fig. 4: Binary obstacle mask O_{mask1} which covers almost all vertical objects (red part) (a). Binary obstacle mask O_{mask2} which covers large vertical objects only (b).

Fig. 5: The disparity map without vertical objects in the obstacle mask O_{mask1} (Fig. 2.a).

3.1.3 Areas of interest

We define areas of interest as those areas where pedestrian can appear in the image. Assuming that pedestrians are not taller than a threshold $high_{pedes}$, pixels corresponding to 3-D points having a distance to (Ω) in object space smaller than $high_{pedes}$ and larger than ε belong to the area of interest, which are represented by a binary mask PD_{mask}. Assuming a pixel $pd(c,r)$ in PD_{mask} , where c and r are the pixel's column and row indices, respectively, corresponds to the 3-D position $pd(\mathrm{x},\mathrm{y},\mathrm{z})$ in object space, PD_{mask} is derived by (3):

$$PD_{mask}(c,\mathrm{r}) = \begin{cases} 1 \text{ if } \varepsilon < dist(pd(\mathrm{x},\mathrm{y},\mathrm{z}),\ \Omega) \le high_{pedes} \\ 0 \qquad otherwise \end{cases} . \tag{3}$$

In (3), $dist$ is the Euclidean distance from a point to a plane in 3-D space.

Fig. 6: Extracted ground plane (green) and areas of interest (red) in the reference image of a stereo pair.

3.2 Object Proposal and Re-ranking

In order to search for pedestrians in the image, we apply an object proposal framework consisting of two stages to the areas of interest with the purpose of achieving a high recall while only producing a small number of candidate bounding boxes. A huge number of initial candidates are produced in the first stage; we then employ additional cues related to geometrical knowledge of pedestrians and scene information to eliminate unreasonable proposals.

3.2.1 Initial proposals

Instead of combining different features, the EB algorithm (ZITNICK & DOLLÁR 2014) directly generates object proposal from edge maps generated by the structured edge detector (DOLLÁR & ZITNICK 2013). EB computes objectness as the likelihood of a bounding box (BB) to contain an object based on the number of contours lying within and on that BB, where a contour is a set of edges forming a coherent boundary, curve or line. The reader is referred to (ZITNICK & DOLLÁR 2014) for more details on how the objectness score is computed. The desired number of BB can be influenced by setting an objectness score threshold. There are three variants of EB based on different parameter values: EB 50, EB 70, and EB 90, in which the step size of the sliding window decreases and the density of sampling increases, respectively. Therefore, depending on the trade-off between accuracy, efficiency, and the number of desired proposals, one wants to achieve for a specific purpose, a suitable variant can be chosen.

In this study, the variant EB 70, which has nearly similar running time as EB 50 but can produce more precise BB, is employed to generate regions that potentially contain any type of object. In order not to eliminate pedestrians in this early stage, we choose a small threshold for the objectness score, which will result in a relatively large number of regions. However, the grey values of image pixels that are outside the areas of interest according to PD_{mask} are set to 0 before passing the image on to the EB 70 algorithm, so that no proposal regions that are completely outside the area of interest will be generated.

3.2.2 Proposal ranking

The initial proposal step produces candidate regions for unspecific types of objects, so that we want to filter out *BB* that are very unlikely to contain pedestrians. For that purpose, we exploit scene information and prior knowledge about pedestrians, which is not considered in the first

stage. We define features that encode this additional information and use them to define a combined scoring function, which is used to rank all proposals. Only *BB* with a high score are maintained because they are very likely to contain a pedestrian, whereas the others are discarded. The high-scoring BB can be fed to a sophisticated classifier for the final decision whether they contain a person or not. Details about that procedure are given in the following paragraphs.

Overlap with the area of interest: this criterion expresses the fact that a good candidate region should lie inside or at least cover a part of the area of interest. We determine an overlap rate f_{int_area} for each bounding box *BB*:

$$f_{int_area} = \frac{\sum_{p(c,r)\in BB}(PD_{mask}(c,r))}{|BB|}, \quad (4)$$

where $p(c,r)$ is a pixel of the proposal bounding box BB and $|BB|$ is the number of pixels covered by BB. BB having a high value of f_{int_area} are more likely to contain a pedestrian than *BB* with a low value of that ratio.

Overlap with the road plane and obstacles: these criteria encode the fact that a bounding box containing a pedestrian should not cover too many pixels belonging to the road or to obstacles such as buildings, fences or tree trunks. We derive two ratios f_g and f_o for the overlap of a BB with the road plane and the obstacles, respectively, to evaluate this constraint, using O_{mask2} to obtain f_o:

$$f_g = \frac{\sum_{p(c,r)\in BB}(G_{mask}(c,r))}{|BB|}, \quad f_o = \frac{\sum_{p(c,r)\in BB}(O_{mask2}(c,r))}{|BB|}. \quad (5)$$

Bounding box ratio: for a specific dataset, a pedestrian appearing in an image should have a size and a ratio between height and width within specific ranges; we assume this to be true even in the case of occlusions. This constraint is related to a two features f_{BBr} and f_{BBw}:

$$f_{BBr} = \frac{BB_{height}}{BB_{width}}; \qquad f_{BBw} = BB_{width}, \quad (6)$$

where BB_{width} and BB_{height} are width and height of the bounding box BB in image space.

Standing on the ground: this criterion is related to the knowledge that a pedestrian should stand on the ground. Even if just a pedestrian's upper part is visible (due to occlusion), the distance of the lowest points of the related BB part from the ground should not be too large. We define the average distance f_{feet} of the lower part of a BB from the ground to evaluate this criterion:

$$f_{feet} = \frac{1}{|BB_{feet}|}\sum_{p\in BB_{feet}} dist(p,\Omega), \quad (7)$$

where BB_{feet} is a square area in the centre of the lower boundary of the bounding box BB (Fig. 7).

Pedestrian height: this criterion is related to the knowledge about pedestrians' heights. The average distance f_{head} of the upper part of a BB from the ground is used as a feature:

$$f_{head} = \frac{1}{|BB_{head}|}\sum_{p\in BB_{head}} dist(p,\Omega), \quad (8)$$

where BB_{head} is a square area in the centre of the upper boundary of the bounding box BB (Fig. 7).

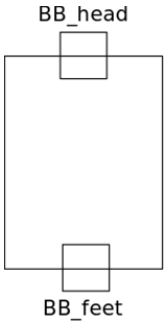


Fig. 7: Head (BB_head) and feet (BB_feet) area of a bounding box *BB*.

Depth consistency: a bounding box BB covering a person should contain a large number of pixels which correspond to that person and thus have similar depth values. This criterion is represented as the highest percentage of pixels having similar depth in BB. We use the depth histogram of all pixels in BB to derive the corresponding value f_{depth} :

$$h_{depth}(\mathrm{i}) = \frac{1}{|BB|}\sum_{i}^{n} 1(p \in BB, i * bin_{size} < depth(p) \leq (i+1) * bin_{size}) \,, \qquad f_{depth} = \max(h_{depth}) \,, \tag{9}$$

where $h_{depth}(\mathrm{i})$ is the histogram entry the i^{th} bin, bin_{size} is the size of each histogram bin, n is the number of bins, and $depth(p)$ is the depth value of pixel p.

Scoring function: our scoring function uses the features f_i described in Eqs. (3) to (8) to compute a score for each bounding box BB delivered by EB 70. We assume these features f_i to be independent and to follow normal distributions. Thus, our scoring function $p(C_{ped}|F)$ delivering a likelihood for a *BB* to contain a pedestrian given the feature vector *F* which contains all the features f_{i} defined earlier in this section is based on the following probabilistic model:

$$\begin{aligned} p(C_{ped}|F) &\propto p(C_{ped})p(F|C_{ped}), \\ &\propto p(C_{ped}) \prod_{i=1}^{n} p(f_i|C_{ped}) \,, \\ &\propto \prod_{i=1}^{n} p(f_i|C_{ped}) \quad . \end{aligned} \tag{10}$$

$$p(f_i|C_{ped}) = \frac{1}{\sqrt{2\pi\sigma^2_{f_i(C_{ped})}}} e^{-\frac{(f_i - \mu_{f_i(C_{ped})})^2}{2\sigma^2_{f_i(C_{ped})}}} \,. \tag{11}$$

In Eq. 10, $p(f_i|C_{ped})$ is a likelihood function for f_i assuming that the BB contains a pedestrian. The factorisation can be made due to our assumption of the features to be independent. The prior $p(C_{ped})$ for a BB to contain a pedestrian is supposed to be uniform and, thus, is neglected. The index i indicates one of the features defined in Eqs. (3) to (8); n is the number of features evaluated. The individual likelihood terms are modelled as Gaussians (Eq. 11), where $\mu_{f_i(C_{ped})}$

and $\sigma^2_{f_i(C_{ped})}$ are mean and variance of the features f_i for a *BB* supposed to contain a pedestrian. They are determined in a training step. We rank the BB according to the scoring function $p(C_{ped}|F)$ and select the N_{opt} best BB as the final proposed regions.

4 Experimental results

Dataset: We evaluate our approach on the Kitti object detection benchmark (GEIGER ET AL., 2012), which has 7481 images in the training set. For these images, *BB* containing pedestrians are available. Just as the benchmark organisers, the *BB* are split into three sets (*hard*, *moderate*, *easy*) according to the level of difficulty to be expected for detecting the respective person based on the BB size, occlusion, and truncation levels in the input image. As the ground truth labels are not provided for the test set of the benchmark, we use only the training set for our evaluation. In order to be able rank our proposal, we first need to determine the means and variances of the normal distributions used to compute the ranking score according to Eqs. (10) and (11). Consequently, we apply cross-validation for the evaluation: we split the available images into five independent sets. In each test run, we use four sets (80% of the data) for learning the parameters of the score function and 20% for testing. We repeat this procedure five times, each time using a different image set for testing, so that in the end, each image contributes to the test set once. We report combined evaluation metrics over all test runs.

We compare the performance of our new techniques to other state-of-the-art proposal methods, namely BING, SS, EB, MCG-D, and 3DOP from (CHEN et al., 2015). Among these methods, BING, SS, and EB take RGB images as input, whereas MCG-D and 3DOP use depth as an additional cue.

Evaluation metrics: The quality of a *BB* proposal is estimated using its intersection over union (IoU) with the ground truth:

$$\mathrm{IoU}(A,B) = \frac{A \cap B}{A \cup B}$$

Following the criteria of the Kitti benchmark (GEIGER ET AL., 2012), a *BB* proposal is counted as a true positive if its IoU is equal to or larger than 0.5. We are mostly interested in the *recall*, which is the percentage of BB containing pedestrians in the reference that were detected by our proposal method and, thus, correspond to true positives. We present recall as a function of the IoU threshold and assess the impact of the number N_{opt} of selected *BB* proposals on the results.

Parameters settings: To estimate the obstacle masks O_{mask1} and O_{mask2}, we set the values of the two thresholds corresponding to the intensity of vertical disparity image to 40 and 200, respectively. We set the value ε in equation (2) to 0.2 (m). Further, a pedestrian is supposed not to exceed 2 meters in height. The side lengths of the squares containing the head and feet areas in Eqs. (7) and (8) (Fig. 7) are set to $1/3$ of the width of the bounding box. For estimating the depth constancy feature, we use a bin size of 1 meter for the depth histogram (see Eq. 9).

Results: We use the EB 70 method (ZITNICK & DOLLÁR 2014) to generate initial BB and select 20000 best candidates per input image based on the objectness score. The average recall of

pedestrians with IoU of 0.5 or better are 99.6%, 96.3%, and 88.6% for the *easy*, *moderate*, and *hard* sets, respectively; these numbers are upper bounds for the recall that our method can achieve because we do not define new BB proposals. Fig. 8 shows the recall of EB 70 and our method as a function of the IoU threshold. The figure shows that our ranking function works very well on the *easy* and *moderate* sets. When ranking the BBs according to our scoring function and selecting the best N_{opt} = 2000 proposals, we discard hardly any correct BB for the *easy* and *moderate* sets, so that we achieve nearly the same recall as EB 70. Thus, we can reduce the number of region proposals delivered by EB 70 by a factor of 10 while losing only a very small percentage of regions corresponding to pedestrians. For the *hard* set, our method is not as successful; only when using the N_{opt} = 5.000 best proposals we can achieve a recall close to EB 70 (Fig. 8); however, this still would reduce the number of BBs to be inspected by more complex processes by a factor of 4. Fig. 9 shows recall as a function of N_{opt}. Again it shows that using N_{opt}= 5.000 results in a reasonably good performance independently from the level of difficulty. In contrast, the state-of-the-art methods need 10.000 proposals to achieve approximately 90% recall for *easy* and *moderate* sets. To obtain our recall at 1.000 proposals, those methods need at least 5.000 bounding boxes (see Figs. 9 and 10). Only 3DOP outperforms our method when N_{opt} is lager than 1.000. This is mainly because they focus on 3D object proposals using deep learning techniques together with additional data. However, when N_{opt} is smaller than 1.000, our method produces even better result than the 3DOP.

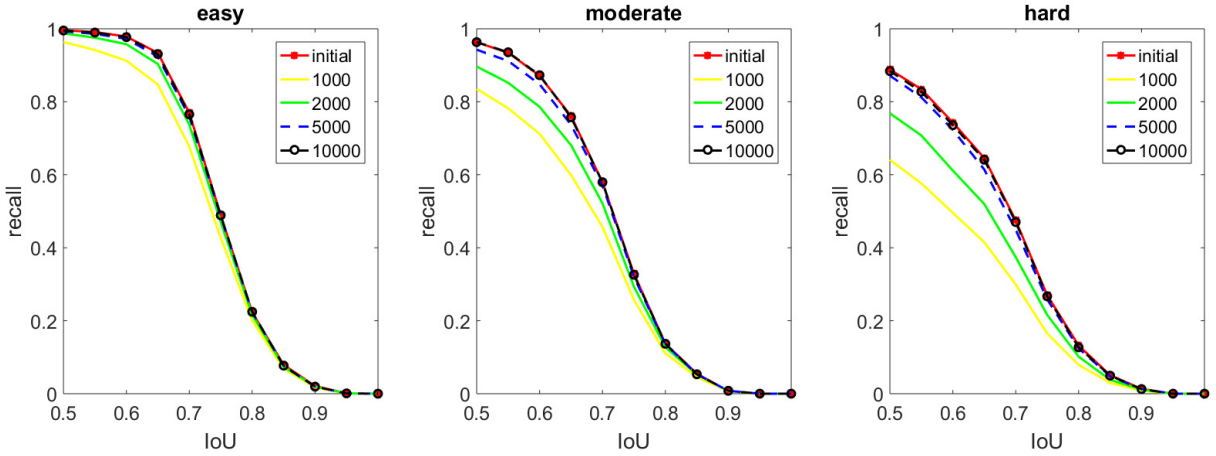


Fig. 8: Recall as a function of the IoU threshold. Init: Evaluation of the 20000 proposals generated by EB 70. The remaining curves were derived for different values for N_{opt} $(1000, 2000, 5000, 10000)$ for selecting the final candidates after ranking.

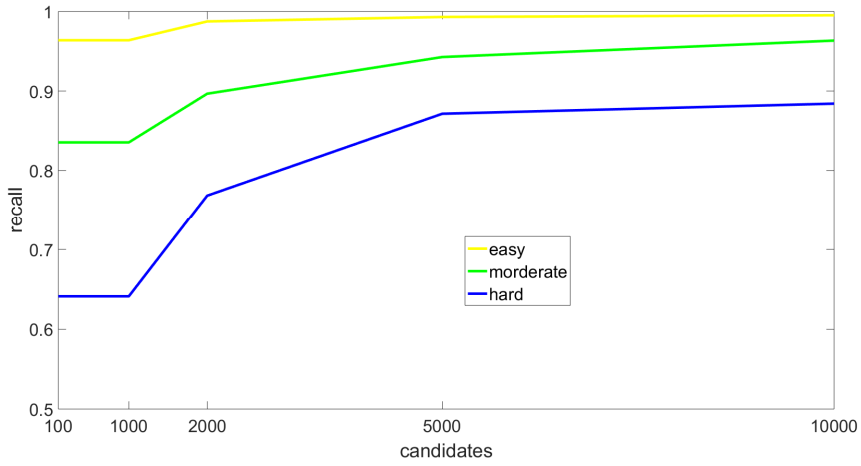


Fig. 9: Recall vs. number of candidates for pedestrian proposals on Kitti dataset of our approach.

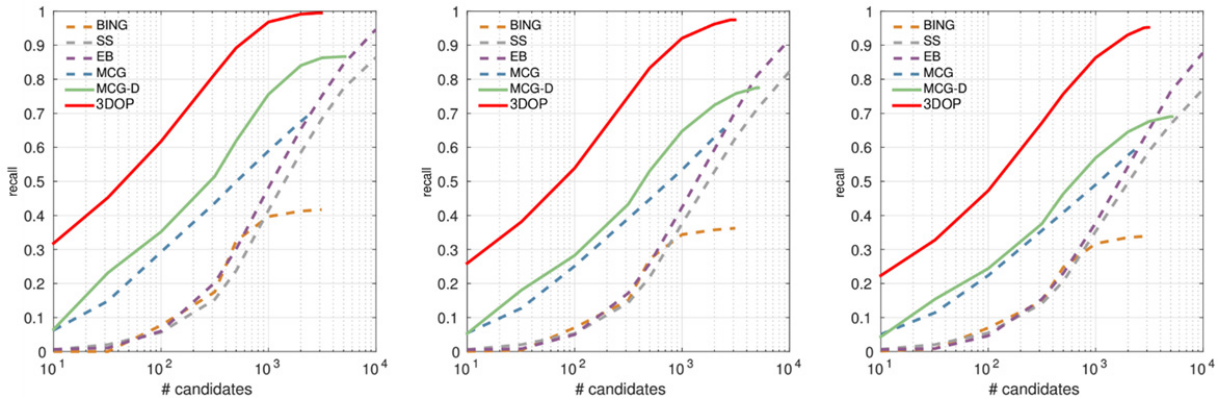


Fig. 10: Recall as a function of the number of candidates for pedestrian proposals on Kitti dataset of state of the art methods: BING, SS, EB, MCG, MCG-D, 3DOP.The graphs are adapted from CHEN et al. (2015).

5 Conclusion

In this paper, we have presented our simple but efficient framework for proposing regions that may contain pedestrians. It combines both, RGB and depth cues in two stages: initial proposal generating and filtering with a new score function. The experimental results show that by exploiting additional 3D cues, our approach can produce promising results and outperform most of the state-of-the-art methods. Specifically, we can achieve approximately 90% recall for the *easy* and *moderate* sets of pedestrians of the Kitti benchmark with 2.000 region proposals, and we can achieve reasonably good results when using 5000 proposals, thus reducing the computational costs of subsequent methods by a factor of 10 and 4, respectively. Moreover, with a very small number of proposals (less than 1.000), our proposal framework can achieve even better recall than the state-of-the art method 3DOP which is based on a sophisticated deep learning method. In the future, we will take advantage of CNN and our proposal method to develop a complete pedestrian detector, which will form the basis of pedestrian tracking.

Reference

ALEXE, B., DESELAERS, T. & FERRARI, V., 2012: Measuring the Objectness of Image Windows. IEEE Transactions on Pattern Analysis and Machine Intelligence, **34**(11), 2189-2202.

ARBELAEZ, P., PONT-TUSET, J., BARRON, J.T., MARQUES, F. & MALIK, J., 2014: Multiscale Combinatorial Grouping. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 328-335.

CHEN, X., KUNDU, K., ZHU, Y., BERNESHAWI, A.G., MA, H., FIDLER, S. & URTASUN, R., 2015: 3D Object Proposals for Accurate Object Class Detection. Advances in Neural Information Processing Systems, 424-432.

CHENG, M. M., ZHANG, Z., LIN, W.Y. & TORR, P., 2014: BING: Binarized Normed Gradients for Objectness Estimation at 300fps. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3286-3293.

DENG, J., DONG, W., SOCHER, R., LI, L.J., LI, K. & FEI-FEI, L., 2009: Imagenet: A Large-scale Hierarchical Image Database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 248-255.

DOLLÁR, P., & ZITNICK, C.L., 2013: Structured forests for fast edge detection. In: Proceedings of the IEEE International Conference on Computer Vision, 1841-1848.

EVERINGHAM, M., VAN GOOL, L., WILLIAMS, C.K., WINN, J. & ZISSERMAN, A., 2010: The Pascal Visual Object Classes (voc) Challenge. International Journal of Computer Vision, **88**(2), 303-338.

GEIGER, A., LENZ, P. & URTASUN, R., 2012: Are We Ready for Autonomous Driving? the Kitti Vision Benchmark Suite. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 3354-3361.

GIRSHICK, R., 2015: Fast R-CNN. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1440-1448.

GIRSHICK, R., DONAHUE, J., DARRELL, T. & MALIK, J., 2014: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 580-587.

GUPTA, S., GIRSHICK, R., ARBELÁEZ, P. & MALIK, J., 2014: Learning Rich Features from RGB-D Images for Object Detection and Segmentation. In European Conference on Computer Vision, 345-360.

HARZALLAH, H., JURIE, F. & SCHMID, C., 2009: Combining Efficient Object Localization and Image Classification. The 12th International Conference on Computer Vision, 237-244.

HOSANG, J., BENENSON, R., DOLLÁR, P. & SCHIELE, B., 2016: What Makes for Effective Detection Proposals? IEEE Transactions on Pattern Analysis and Machine Intelligence, **38**(4), 814-830.

HU, Z. & UCHIMURA, K., 2005: U-V-disparity: An Efficient Algorithm for Stereovision Based Scene Analysis. The IEEE Intelligent Vehicles Symposium, 48-54.

REDMON, J., DIVVALA, S., GIRSHICK, R. & FARHADI, A., 2016: You Only Look Once: Unified, Real- time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779-788.

REN, S., HE, K., GIRSHICK, R. & SUN, J., 2015: Faster R-CNN: Towards Real-time Object Detection with Region Proposal Networks. In Advances in Neural Information Processing Systems, 91-99.

UIJLINGS, J.R., VAN DE SANDE, K.E., GEVERS, T. & SMEULDERS, A. W., 2013: Selective Search for Object Recognition. International Journal of Computer Vision, **104**(2), 154-171.

VEDALDI, A., GULSHAN, V., VARMA, M. & ZISSERMAN, A., 2009: Multiple Kernels for Object Detection. The 12th International Conference on Computer Vision, 606-613.

WANG, X., YANG, M., ZHU, S. & LIN, Y., 2013: Regionlets for Generic Object Detection. In Proceedings of the IEEE International Conference on Computer Vision, 17-24.

YAMAGUCHI, K., MCALLESTER, D. & URTASUN, R., 2014: Efficient Joint Segmentation, Occlusion Labeling, Stereo and Flow Estimation. In European Conference on Computer Vision, Springer, 756-771.

ZITNICK, C.L. & DOLLÁR, P., 2014: Edge Boxes: Locating Object Proposals from Edges. In European Conference on Computer Vision, Springer, 391-405.

Augmented Reality an Fenstern: Herausforderungen und Lösungskonzepte

MARCO GRAF[1] & MARTIN CHRISTEN[1]

Abstract: Augmented Reality an Fenstern bietet vor allem aufgrund des hohen möglichen Dynamikumfangs der Bilder einige Herausforderungen. So sind Marker, die bei gewöhnlicher Beleuchtung funktionieren, bei Gegenlicht-Bildern von Fenstern untauglich. Weiter kam erschwerend dazu, dass der Marker von jeder Person ohne Infrastruktur erstellt werden können sollte. Die Lösung wurde in Form eines leeren DIN-A4-Blattes gefunden. Da zudem beispielsweise Fensterdimensionen mit hoher Genauigkeit in den Bildern gemessen werden können sollten, mussten Kalibrierwerte von Smartphone-Kameras verwendet werden. Entgegen ersten Annahmen, gab es keine tauglichen bestehenden Datenbanken solcher Parameter, sodass diese selbst ermittelt werden mussten. Dies geschah mittels einer Aluminium-Kalibrierplatte mit aufgedruckten, codierten Markern, der Software Agisoft PhotoScan und einer Automatisierung über Python. Mit diesen Grundlagen konnten erfolgreich Position und Orientierung der Bilder in Relation zum Marker berechnet, Messungen mit einer Standardabweichung von knapp 2cm durchgeführt, sowie einfachste Objekte ins Bild augmentiert werden.

1 Einleitung

Die Bestimmung der Dimensionen von Objekten wie Fenstern, Wandaussparungen und Ähnlichem ist für verschiedene Anwendungen im Feld der Inneinrichtung ein wichtiger Schritt. Herkömmlich geschieht dies mittels Zollstock oder Laserdistanzmesser, was jedoch die entsprechende Ausrüstung und zudem Zugänglichkeit zum gesamten Objekt erfordert.
In diesem Beitrag wird eine Methode vorgestellt, mittels einer Smartphone-Bildaufnahme solche Distanzmessaufgaben durchzuführen. Während in der Photogrammetrie gewöhnlicherweise mit mehreren Bildern gearbeitet wird, war hier die Randbedingung gegeben, nur eine einzige Aufnahme verwenden zu können. Da zum Messen in Einzelbildern eine Grössenreferenz notwendig ist, soll zudem ein für den Nutzer möglichst einfach erhältlicher Marker gefunden werden. Üblicherweise sind dies gedruckte Marker, die spezifische Muster zur Erkennung beinhalten. Diese Anwendung hingegen sollten auch Nutzer verwenden können, die keine solche Infrastruktur besitzen. Weiter soll die Aufnahmeposition und -ausrichtung der Bilder bestimmt werden, sodass anschliessend mittels Augmented Reality (AR) virtuelle Objekte ins Bild gerendert werden können.

[1] Fachhochschule Nordwestschweiz, Hochschule für Bau, Architektur und Geomatik, Institut Geomatik, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: [martin.christen, marco.graf]@fhnw.ch

2 Stand der Technik

Sowohl zur Messung in Einzelbildern, sowie zur Smartphone-Kamerakalibrierung gibt es bestehende Lösungen, die folgend kurz erläutert werden. Zudem wird der Begriff der Augmented Reality umrissen.

2.1 Messungen in Einzelbildern

Zur Messung in Einzelbildern gibt es bereits verschiedene Apps und Webapplikationen. Den meisten gemein ist die notwendige manuelle Eingabe einer Referenzlänge. Weiter gibt es Lösungen, die nur in orthogonal aufgenommenen Fotos funktionieren, wie etwa „KLONK Image Measurement“ von IMAGE MEASUREMENT CORPORATION (2013). Somit können mit dieser Lösung keine perspektivischen Verzerrungen berücksichtigt werden. Etwas mehr Möglichkeiten bietet beispielsweise „ImageMeter“ von FARIN (2012). In dieser App kann ein Viereck aufgezogen werden, dessen Masse bekannt sind. Als Beispiel werden oft die vier Seiten einer Wand aufgeführt. Dadurch kann die perspektivische Verzerrung berücksichtigt werden, entsprechend fällt die Bedingung der orthogonal zum Objekt liegenden Aufnahmeposition weg. Sehr ähnliche Funktionalität weist auch „On 3D-CameraMeasure” von POTATOTREESOFT (2017) auf. Darin sind aber bereits bekannte Referenzobjekte, wie Kreditkarte, A3- und A4-Papier vordefiniert. Alternativ kann ein benutzerdefiniertes Referenzobjekt angegeben werden. All diesen Apps gemein ist, dass das Referenzobjekt oder der Referenzmassstab manuell angewählt werden müssen. Das Wegfallen dieser manuellen Auswahl ist ein Alleinstellungsmerkmal des Webdienstes «smartpicture3d» von SMARTPICTURE TECHNOLOGIES (o. J.). Im Gegenzug muss für die Nutzung dieses Dienstes aber ein spezieller Marker ausgedruckt und am zu messenden Objekt angebracht werden. Anschliessend müssen gemäss Benutzeranleitung Fotos aus drei unterschiedlichen Standpunkten (im Falle eines Fensters aus der Nähe, von Weitem und eine Schrägansicht) aufgenommen werden. Anschliessend kann man die gewünschten Masse im Bild markieren und die Daten einsenden. Nach einigen Tagen wird das Resultat als PDF per Mail zugestellt. Sämtliche dieser Anwendungen funktionieren ohne Berücksichtigung der Kamera-Kalibrierung. Dies bringt den Nachteil mit sich, dass beispielsweise die radiale Verzeichnung die Messergebnisse verfälscht. Weiter ist eine manuelle Markierung des Markes umständlich, fehleranfällig und teilweise schwierig durchzuführen.

2.2 Augmented Reality (AR)

Unter Augmentierter Realität (Augmented Reality, AR) versteht man allgemein die Anreichung der Realität durch künstliche virtuelle Inhalte. Dabei kommt es zu einer Verschmelzung der Realität mit der Virtualiät (JUNG 2013). Es besteht eine Vielzahl unterschiedlicher Definitionen von AR. Bereits in den 60er Jahren wurde AR von SUTHERLAND (1965) realisiert. Etabliert hat sich der Begriff AR jedoch erst mit der Definition von AZUMA (1997). Danach versteht man unter AR eine Überlagerung von Realität und Virtualität, welche interaktiv ist und bei welcher die Darstellung in Echtzeit erfolgt und 3D-Objekte geometrisch registriert sind.

2.3 Smartphone-Kamerakalibrierung

Die Kamerakalibrierung ist ein essentieller und wichtiger Teil jedes AR-Systems (ABDULLAH & MARTINEZ 2002). Dabei werden die Parameter der inneren Orientierung, also die Lage des Bildhauptpunktes, die Kamerakonstante und Verzeichnungsparameter geschätzt. Entscheidend ist sie ebenso für die vorliegende Arbeit, da neben dem ins Bild rendern von virtuellen Objekten ebenso mit hoher Genauigkeit darin gemessen werden können sollte. Die Annahme, es würden bereits frei zugängliche Sammlungen solcher Kalibrierdaten geben, konnte nicht bestätigt werden. Es gibt zwar einige Ansätze solcher Sammlungen wie bei RAY (2016), HTW-INKA (2014) oder LENSFUN (2015). Allerdings erfüllte keine dieser existierenden Sammlungen die Anforderungen dieser Arbeit. Entweder waren kaum Daten vorhanden, der Fokus lag auf Spiegelreflex- und Systemkameras, oder es waren keine aktuellen Geräte erfasst, da die Datenbank bereits zu einem früheren Zeitpunkt kommerzialisiert wurde. Eine mögliche Abhilfe würde das Augmented Reality Toolkit (ARToolkit) bieten. ARToolkit ist eine Open-Source Bibliothek für das Tracking in AR-Anwendungen und bietet eine App an, die die Kalibrierung von Smartphones ermöglicht (ARTOOLKIT 2016a). Dazu wird ein ausgedrucktes Schachbrettmuster als Referenz verwendet. Leider ist die Auflösung auf maximal 1920×1080 Pixel beschränkt, weshalb es für die vorliegende Arbeit nicht eingesetzt werden konnte, da Bilder in der vollen Auflösung der Kamera verwendet werden mussten.

3 Evaluation eines geeigneten Markers

Soll in Einzelbildern einer monokularen Kamera gemessen werden können, ist zwingend eine Grössenreferenz notwendig. Zudem war die Bestimmung der Ausrichtung der Kamera eine weitere Anforderung. Dazu können AR-Marker verwendet werden. Mittlerweile gibt es verschiedenste Systeme von AR-Markern, die mittels Drucker erstellt werden können. Dazu gehören klassische quadratische Marker, die zudem bei einigen Systemen, wie etwa ARTOOLKIT (2016b), individuell gestaltet werden können. Bedingung für die Nutzung solcher Marker ist allerdings der Besitz eines Druckers und die Kenntnis der Bedienung eines solchen. Deshalb musste ein Marker gefunden werden, der auch ohne Infrastruktur durch jede Person nutzbar ist. Da die Marker auf Fenstern angebracht werden, war die Robustheit gegenüber verschiedensten Lichtverhältnissen ein wichtiges Merkmal. So entspricht beispielsweise die Innenraumbeleuchtung bei Nacht einer Direktbeleuchtung, während bei Tageslicht starkes Gegenlicht auf den Marker fällt.

3.1 Geldschein als Marker

Eine Variante eines Markers, die leicht zu beschaffen ist, praktisch jede Person Zugriff darauf hat und die Dimensionen exakt bekannt sind, wäre ein Geldschein. Konkret wurden Versuche mit einer Schweizer 20-Franken-Note als Marker durchgeführt. Dabei zeigte sich allerdings, dass sich das Erscheinungsbild mit ändernden Lichtverhältnissen stark unterschiedlich präsentieren kann. Dabei kommen im Gegenlicht Merkmale stark zum Vorschein, die bei Direktbeleuchtung kaum erkennbar sind (siehe Abb. 1 und Abb. 2). Entsprechend extrahieren Algorithmen wie SIFT, SURF oder ORB in verschiedenen Lichtverhältnissen unterschiedliche Merkmale, was die Verwendung als Marker erschwert. Weiter führt eine grössere Distanz zum Marker dazu, dass kaum mehr Merkmale erkennbar sind (siehe Abb. 3). Daher eignet sich ein Geldschein nicht als Marker. Allgemein wurde erkannt, dass Marker, die auf Merkmalsextraktion beruhen, in stark ändernden Lichtverhältnissen kaum praktikabel sind.

Abb. 1: Geldschein mit Direktbeleuchtung

Abb. 2: Geldschein im Gegenlicht

Abb. 3: Geldschein aus einigen Metern aufgenommen

3.2 DIN-A4 als Marker

Schlussendlich konnte ein tauglicher Marker in Form eines leeren DIN-A4-Blattes gefunden werden. Obwohl sich dieses mit ändernden Lichtverhältnissen ebenso visuell stark verändert, bleibt es dennoch eine relativ homogene Fläche, die bei ungefähr frontaler Sichtweise zudem die Proportionen eines A4 ($1:\sqrt{2}$) aufweist (siehe Abb. 4).

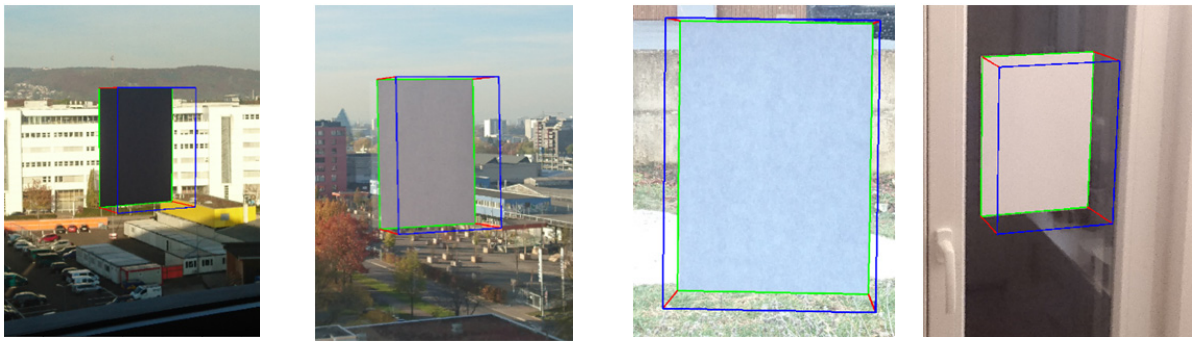


Abb. 4: DIN-A4-Marker in unterschiedlichsten Lichtverhältnissen mit augmentiertem Quader

Diese homogene Fläche erlaubt eine Detektion des Markers über Canny-Kantenfilter (CANNY 1986) und anschliessender Umriss-Erkennung in OpenCV. Dabei werden die vier Eckpunkte

extrahiert, die den Bezug zwischen Bild- und Objektraum herstellen. Konkret wurde im Objektraum ein Koordinatensystem mit Nullpunkt in der unteren linken Ecke des A4-Blattes definiert.

4 Kalibrierung der Kameras

Aufgrund der teils starken Verzeichnungen der günstigen, sehr kleinen Smartphone-Kameras ist für Augmented Reality, sowie zur genauen Messung in Bildern, eine Bestimmung der inneren Orientierung der Kamera unabdingbar. Mit nur einer Bildaufnahme und den vier Referenzkoordinaten des A4-Markers ist eine Simultankalibrierung der Kamera, wie sie oft in der Photogrammetrie durchgeführt wird, nicht möglich. Deshalb sind die Parameter der inneren Orientierung wichtig, um gute Resultate zu erhalten. Wie erwähnt, konnte keine der existierenden Datenbanken von Kalibrierungsdaten die Anforderungen erfüllen. Deshalb war eine eigenständige Erstellung einer solchen unumgänglich.

4.1 Untersuchung bestehender Verfahren

Ein mögliches Werkzeug dazu bietet AGISOFT LLC (2011) mit der Software „Agisoft Lens“, die einzig dem Zweck der Kamera-Kalibrierung dient. Dazu wird ein auf dem Bildschirm dargestelltes Schachbrettmuster aus unterschiedlichen Perspektiven aufgenommen und daraus die innere Orientierung geschätzt.
Allerdings wurde bemerkt, dass die Ergebnisse sehr inkonsistent waren. Beispielsweise variierte die Verzeichnung in jedem Kalibriervorgang, teilweise sogar sehr stark (am Bildrand über 20 Pixel). Dies ist in Abb. 5 ersichtlich. Ein ähnliches Verhalten konnte bei der Kamerakonstante f beobachtet werden. Die Gründe dafür konnten nicht abschliessend eruiert werden. Eine Veränderung der Kamerakonstante könnte daher herrühren, dass die Kamera aufgrund der relativ kurzen Distanz zum Schachbrettmuster bei jeder Aufnahme neu fokussierte.

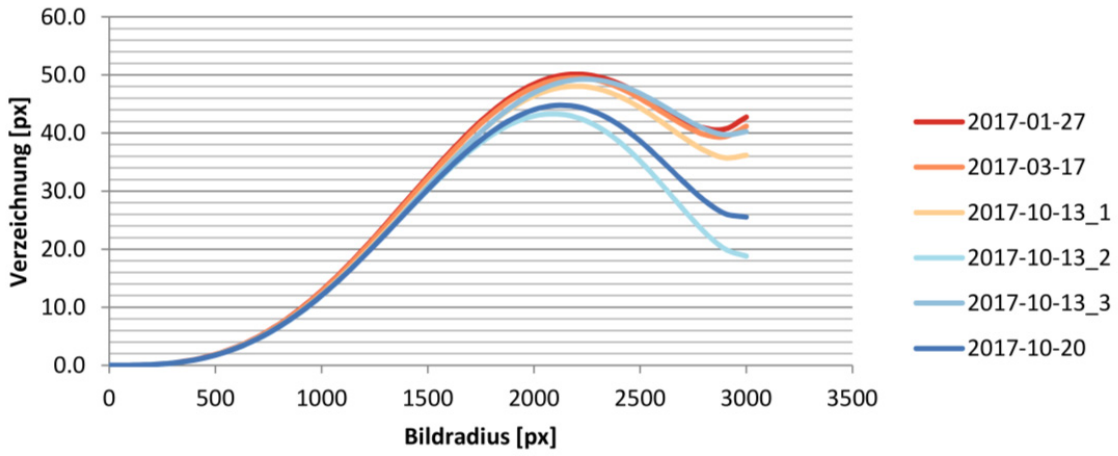


Abb. 5: Ermittelte Verzeichnung einer Xperia Z5 Smartphone-Kamera in verschiedenen Kalibrierungsvorgängen in der Software Agisoft Lens

4.2 Entwickeltes Verfahren

Aus den vorhergehend genannten Gründen und der Annahme, dass die Aufnahmedistanz der Kalibrierbilder etwa der späteren Einsatzdistanz entsprechen sollte, wurde nach einer alternativen Lösung gesucht. Da Versuche in der kompletten Photogrammetrie-Software von AGISOFT LLC (2016) – Agisoft PhotoScan Professional – konsistentere Ergebnisse lieferten und die Entwicklung einer eignen Kalibriersoftware aufwändig gewesen wäre, wurde versucht, auf dieser bestehenden Software aufzubauen. Grundlage dafür ist eine bedruckte Aluminium-Kalibrierplatte im Format A0 (Abb. 6) mit codierten Markern (Abb. 7). Diese Codierung beinhaltet die Nummer des Markers, womit sie automatisch erkannt und korrekt zugeordnet werden können.

Um die Soll-Koordinaten der Marker zu erhalten, wurden einerseits die Soll-Positionen, der in Druck gegebenen Marker, andererseits Tachymeter-Messungen auf alle vier Marker in den Ecken, verwendet. Eine Transformation zeigte auf, dass die gedruckten Marker einen leichten affinen Verzug aufwiesen. Dieser wurde geschätzt und auf sämtliche Marker-Koordinaten angewandt. Diese bildeten die Soll-Koordinaten für die Kalibrierung. Aufgenommen wurden jeweils acht Bilder mit unterschiedlichen Positionen und Ausrichtungen, wie dies (LUHMANN 2010) empfiehlt. Damit wäre bereits eine Kalibrierung der Kameras möglich gewesen. Allerdings wäre diese mit viel manuellem Aufwand verbunden gewesen (Projekt erstellen, Bilder und Soll-Koordinaten importieren, etc.), sodass dieser Ablauf mittels Python automatisiert wurde. Damit kann mittels eins Sets von acht Bildern automatisch eine Kalibrierung der Kamera durchgeführt werden kann, die in einer Datenbank abgespeichert wird.

Abb. 6 Alu-Kalibrierplatte mit codierten Markern

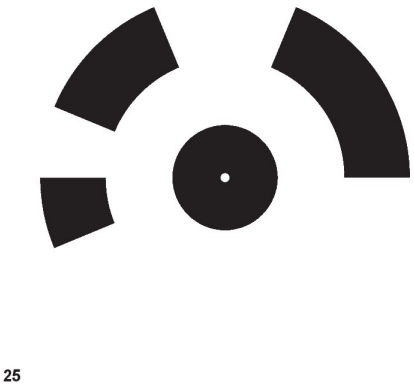


Abb. 7: Detailansicht eines Markers

5 Untersuchungen zu Messungen in Bildern

Die Grundlagen aus den vorhergehenden Kapiteln ermöglichen eine Bestimmung der Kameraposition- und Ausrichtung relativ zum Marker. Dazu wird in den EXIF-Informationen des Bildes der Kameratyp ausgelesen, entsprechende Kalibrierungswerte aus der Datenbank gelesen, und

anschliessend über solvePnP von OPENCV (2017) die Kameraposition und -orientierung berechnet. Dies ermöglicht einerseits, virtuelle Objekte ins Bild zu rendern (siehe einfachste Beispiele in Abb. 4), andererseits werden damit Messungen im Bild möglich. Wichtig ist festzuhalten, dass diese Messungen nur auf der Ebene des Markers korrekt sind. Soll beispielsweise die Breite der Wandaussparung auf Höhe Vorderkante Wand gemessen werden, so ist der Tiefenunterschied zwischen dem Marker und der Messebene zu berücksichtigen. Durchgeführte Untersuchungen mit einem Sony Xperia Z5 Smartphone zeigten eine beträchtliche Genauigkeit auf. Mittels 12 Bildern von verschiedenen Fenstern wurden aus Distanzen von gut einem Meter bis knapp drei Meter Masse berechnet. Diese wurden mit Referenzmassen, gemessen mit Laserdistanzmesser, verglichen (je vier Masse pro Bild). Dabei resultierte eine mittlere Standardabweichung von 1.7cm.

6 Fazit & Ausblick

Der hohe Dynamikumfang von Bildern von Fenstern führte zu Herausforderungen, vor allem bei der Auswahl des Markers. Die gefundene Lösung in Form eines A4-Blattes, scheint aber geeignet, trotz solchen stark unterschiedlichen Lichtverhältnissen zu funktionieren. Die entwickelte, mittels Python automatisierte, Kalibrierung von Smartphone-Kameras ist effizient und liefert gute Ergebnisse, wie in der anschliessenden Untersuchung von Fenster-Massen in Bildern gezeigt werden konnte. Die erreichte Standardabweichung von knapp 2cm ist beachtlich, bedenkt man die starke Extrapolation vom A4-Marker auf ein raumhohes Fenster. Augmentiert wurden nur einfachste Objekte, wie Quader, die Grundlage für weitere Anwendungen ist aber vorhanden. Dank der Kameraposition- und Ausrichtung können beliebige Objekte ins Bild gerechnet werden.

Mit im letzten Jahr vorgestellten „ARKit“ von APPLE (2017) bieten sich neue Möglichkeiten für die Messung von Objektdimensionen. So ist es möglich, die Messpunkte direkt mit dem Gerät anzufahren und damit auch ein präzises Anzielen der Messpunkte zu ermöglichen. Ein solches Beispiel zeigt (MILLER 2017). Solche, auf visueller Odometrie beruhende Verfahren, könnten sich für die Bestimmung von Objektdimensionen an Fenstern eignen. Wie robust sie bei gleichmässigen Texturen (weisse Wand) und überbelichteten Bildbereichen funktionieren, müsste geprüft werden.

Gerade für die Augmentierung von komplexeren Objekten an Fenstern, wäre es zudem notwendig, die Lichtverhältnisse erkennen und entsprechend adaptieren zu können. Verfahren dazu bestehen bereits (WANG & SAMARAS 2002; LOPEZ-MORENO et al. 2013; WEHRWEIN et al. 2015). Die Anwendbarkeit auf Fenster wäre in einem weiteren Schritt zu evaluieren.

7 Literaturverzeichnis

ABDULLAH, J. & MARTINEZ, K. 2002: Camera self-calibration for the ARToolKit. In: First IEEE International Augmented Reality Toolkit Workshop. Proceedings. First IEEE International Augmented Reality Toolkit Workshop, Proceedings, Darmstadt, DOI: 10.1109/ART.2002.1107017.

AGISOFT LLC, 2011: Agisoft Lens User Manual - Version 0.4.0. http://downloads.agisoft.ru/lens/doc/en/lens.pdf [Stand: 15.12.2017].

AGISOFT LLC, 2016: Agisoft PhotoScan User Manual. Professional Edition, Version 1.2. http://www.agisoft.com/pdf/photoscan_1_2_en.pdf [Stand: 20.07.2016].

APPLE, 2017: ARKit | Apple Developer Documentation. https://developer.apple.com/documentation/arkit [Stand: 19.01.2018].

ARTOOLKIT, 2016a: Camera Calibration App for Android. https://archive.artoolkit.org/documentation/doku.php?id=4_Android:android_camera_calibration [Stand: 17.01.2018].

ARTOOLKIT, 2016b: Creating and Training Traditional Template Square Markers. https://artoolkit.org/documentation/doku.php?id=3_Marker_Training:marker_training [Stand: 17.01.2018].

AZUMA, R.T., 1997: A Survey of augmented reality. In: Presence **6**(4). 355-385.

CANNY, J., 1986: A Computational Approach to Edge Detection. In: IEEE Transactions on Pattern Analysis and Machine Intelligence. **PAMI-8**(6), 679-698. DOI: 10.1109/TPAMI.1986.4767851.

FARIN, D., 2012: ImageMeter - Manual. http://dirk-farin.net/projects/ImageMeter/manual.html [Stand: 17.01.2018].

HTW-INKA, 2014: Automatic camera intrinsics finder and database. https://github.com/htw-inka/cam-intrinsics-db [Stand: 17.01.2018].

IMAGE MEASUREMENT CORPORATION, 2013: Image Measurement App. http://www.imagemeasurement.com/en/image-measurement-app [Stand: 17.01.2018].

JUNG, B., GRIMM, P., BROLL, W. & DÖRNER R., 2013: Virtual und Augmented Reality (AR/VR) Grundlagen und Methoden der Virtuellen und Augmentierten Realität, eXamen.press, DOI 10.1007/978-3-642-28903-3_8, Springer-Verlag Berlin Heidelberg

LENSFUN, 2015: Lensfun. http://lensfun.sourceforge.net/ [Stand: 17.01.2018].

LOPEZ-MORENO, J., GARCES, E., HADAP, S., REINHARD, E. & GUTIERREZ, D., 2013: Multiple Light Source Estimation in a Single Image. In: Computer Graphics Forum, **32**(8), 170-182. DOI: 10.1111/cgf.12195.

LUHMANN, T., 2010: Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen. 3. Auflage Berlin: Wichmann, H.

MILLER, C., 2017: Latest ARKit app turns iOS into a powerful measuring tool. https://9to5mac.com/2017/09/06/latest-arkit-app-turns-ios-into-a-powerful-measuring-tool-video/ [Stand: 19.01.2018].

OPENCV, 2017: Camera Calibration and 3D Reconstruction - solvePnP. https://docs.opencv.org/3.4.0/d9/d0c/group__calib3d.html#ga549c2075fac14829ff4a58bc931c033d [Stand: 19.01.2018].

POTATOTREESOFT, 2017: ON 3D-CameraMeasure. https://play.google.com/store/apps/details?id=com.potatotree.on3dcamerameasure&hl=de [Stand: 17.01.2018].

RAY, D., 2016: Camera Calibration database – argus : 3D for the people. http://argus.web.unc.edu/camera-calibration-database/ [Stand: 17.01.2018].

SMARTPICTURE TECHNOLOGIES, 2018: Smart Picture 3D – Snap. Measure. Share. http://smartpicture3d.com/ [Stand: 17.01.2018].

SUTHERLAND, I., 1965: The ultimate display. In: Multimedia: From Wagner to virtual reality.

WANG, Y. & SAMARAS, D., 2002: Estimation of multiple directional light sources for synthesis of mixed reality images. In: 10th Pacific Conference on Computer Graphics and Applications, 2002. Proceedings, 38-47. DOI: 10.1109/PCCGA.2002.1167837.

WEHRWEIN, S., BALA, K. & SNAVELY, N., 2015: Shadow Detection and Sun Direction in Photo Collections. In: 2015 International Conference on 3D Vision, Lyon, 460-468. DOI: 10.1109/3DV.2015.58.

Weed Detection in Close-range Imagery of Agricultural Fields using Neural Networks

AVISHEK DUTTA[1], JOSEPH M. GITAHI[1], PRASHANT GHIMIRE[1], ROBIN MINK[2], GERASSIMOS PETEINATOS[2], JOHANNES ENGELS[1], MICHAEL HAHN[1] & ROLAND GERHARDS[2]

Abstract: Modern day agriculture is becoming an endeavour where precision is highly desired and drones with imaging capabilities are contributors of big data to this field of research. Our focus in this paper is on precise weed control, in particular, the optimisation of yield and cost while having minimum impact on the environment. The information about in-field variability of weed patches can be exploited for sub-plot specific weed control, which leads to a restrained yet flexible use of herbicides. We use close-range imagery of weeds recorded with handheld cameras, having a resolution of only a few millimetres per pixel in their natural environment. In this paper we adapt Convolutional Neural Networks with the goal of separating weeds from the main crops in close-range imagery. We explore two ways to design the networks: pixel-wise classification and object-based detection. For both approaches, we use pre-trained networks, which are fine-tuned with the recorded weed images. The deep learning libraries used are Caffe and Tensorflow. The study demonstrates successful adaptation of pre-trained CNNs for weed classification in close-range imagery which could be extended to UAV imagery in future.

1 Introduction

Modern agricultural techniques focus on high yields, low costs and eco-friendly practices. While frugal use of herbicides is desired, weed control still remains important for the increase in the productivity of the crops. The common approach so far is the uniform application of herbicides to a field, neglecting the spatial variability of weed species and densities. This results in higher costs, pollution of soil and water resources. In addition, the herbicides may adversely affect the crops if applied in high concentrations. By mapping different species of weeds, their density and distribution, herbicide spraying can be adjusted as opposed to uniform application.

As weeds and crops are spectrally similar at an early stage of growth, pixel-based classifications of the two do not always provide reliable accuracy and hence popular rule-based methods include features such as the shape of leaf and texture to increase the achievable accuracy of classification (SØGAARD 2005; ÅSTRAND & BAERVELDT 2002; GOLZARIAN & FRICK 2011). However, these classification methods rely on handcrafted feature extraction techniques which are not robust enough to discern complexities that exist in the natural environment.

Our ultimate objective is to classify weeds in aerial images acquired from a height of less than 10 meters but for this study, we restrict ourselves to terrestrial images i.e. images acquired with hand-held devices. In this paper, we propose Convolutional Neural Networks (CNNs) for weeds classification using two approaches, object detection and pixel-wise labelling. We adapt pre-trained CNN models trained on expansive datasets and fine-tune with the weeds' images.

[1] Hochschule für Technik Stuttgart, Schellingstrasse 24, D-70174 Stuttgart, E-Mail: avishek.dutta@hft-stuttgart.de

[2] Universität Hohenheim, Institut für Phytomedizin, Otto-Sander-Str. 5, D-70599 Stuttgart, E-Mail: robin.mink@uni-hohenheim.de

2 Data

The data used for this study comprises images of five different weed species, as shown below.

Tab. 1: A list of the weeds classes available for the study

EPPO Code	Scientific name	English name
MATCH	Matricaria chamomilla	Wild chamomile
POAAN	Poa annua	Annual meadowgrass
STEME	Stellaria media	Common chickweed
VIOAR	Viola arvensis	Field pansy
AMARE	Amaranthus retroflexus	Common amaranth

The images were taken at Heidfeldhof farm, located in Plieningen, Stuttgart. These images were photographed vertically downwards from approximately 50cm above the ground by a cell phone camera having a resolution of 3024×4032 pixels. The weeds were cultivated in natural environment and were at an early stage of growth. Additionally some weed images from a dataset created by GISELSSON et al. (2017) were used for testing the models in the pixel-based approach. This dataset consists of images taken in controlled conditions with the soil covered by small stones to prevent green moss layer.

3 Object-based approach

Object detection deals with localization of objects in addition to the classification of the same objects. This approach of the study utilizes Faster R-CNN (REN et al. 2015) architecture for localizing weeds along with the recognition of the species to which they belong, by fine-tuning a pre-trained Resnet-50 model (HE et al. 2016) trained on 90 different object categories of COCO dataset (LIN et al. 2015).

Resnet-50 is a 50-layer architecture, consisting of the building blocks as shown in Table 2. A building block, for example, Conv2_x consists of three units with each unit having three convolutional layers of 1×1, 3×3 and 1×1. A shortcut connection is added from the result of the previous block (which in this example is 3×3 max pool) to the output of the first unit of Conv2_x. This shortcut connection is implemented as element-wise addition of the outputs of 3×3 max_pool to the output of the first sub-block of Conv2_x. Shortcut connections are also provided within the adjacent units in a block. For example, a shortcut connection is provided between the output of 1×1, 256 convolution of first unit of block Conv2_x and the output of 1×1, 256 convolution of the second unit of the same block. The shortcut connections between the units in adjacent block decrease the image size. Batch normalization has been performed after each convolution to overcome the problem of exploding and vanishing of the gradients during backpropagation.

Tab. 2: Resnet Architecture for an input size of 224*224 (adapted from HE et al. (2016))

Building block	Output size	Units
Conv1	112*112	7*7, 64, stride 2
	56*56	3*3 max pool, stride 2
Conv2_x	56*56	$\begin{bmatrix} 1*1, 64 \\ 3*3, 64 \\ 1*1, 256 \end{bmatrix} * 3$
Conv3_x	28*28	$\begin{bmatrix} 1*1, 128 \\ 3*3, \ 128 \\ 1*1, 512 \end{bmatrix} * 4$
Conv4_x	14*14	$\begin{bmatrix} 1*1, 256 \\ 3*3, 256 \\ 1*1, 1024 \end{bmatrix} * 6$
Conv5_x	7*7	$\begin{bmatrix} 1*1, 512 \\ 3*3, 512 \\ 1*1, 2048 \end{bmatrix} * 3$
	1*1	Average pool, 5-d fc

For the detection of weeds species, Faster R-CNN algorithm has been used. The blocks up to Conv4_x were used for Region Proposal Network (RPN) whereas the final block Conv5_x was used as an object detector to predict the class of weeds along with the refinement of the bounding boxes proposed by the RPN. Region Proposal Network consists of a sliding window of kernel size n, over the feature maps generated by the block Conv4_x. At each location of the sliding window, a number of object proposals are computed. These object proposals are computed on the basis of reference boxes, called grid anchors. A number of grid anchors of varying scales and aspect ratios are generated at each position of the sliding window. Though the anchors are generated in feature maps, the coordinates of these anchors corresponds to the image coordinates. The correspondence of the feature maps to that of the image can be calculated as the feature maps are generated through a series of convolutions and pooling.

A grid anchor is labelled as positive if the Intersection over Union (IoU) of the grid anchor to the bounding box of the ground truth is over a certain threshold. Intersection over Union is calculated by the area of intersection of the boxes divided by the area of union. All the grid anchors below the threshold are labelled as negative. These proposals are passed to the classifier and regressor in the RPN where they are classified as objects and background by the classifier and the regressor computes the bounding box regression.

The proposals that have been classified as objects and non-objects along with their regressed bounding box are sent to the detector network, consisting of the final convolution block of the network. Here, the objects are classified into their classes and the coordinates of the bounding box are computed.

3.1 Data Preparation

The images were cropped due to the memory limitations of the computer before being added to the training and test sets. The bounding boxes were created in an XML format using LabelImg tool (LIN 2015). Table 3 shows the number of training and test images used for the study.

Tab. 3: List of Training and Test images

S.N.	Images containing	Number of Training Images	Number of Test images
1	MATCH	78	20
2	POAAN	132	33
3	STEME	174	40
4	VIOAR	126	30
5	AMARE	130	33

The training images contain at least one object and at most 8 objects of an individual class. The bounding box of each object include four coordinates, minimum and maximum x and y coordinate values. In addition, class information is provided.

3.2 Implementation

The weights of the network were initialized by the pre-trained model on COCO dataset and only the weights of the fully connected layer were initialized by Xavier initialization (GLOROT & BENGIO 2010).

The training of the Region Proposal Network (RPN) was done simultaneously along with the training of the object detector, as opposed to the four-step alternating training proposed by REN et al. (2015). A batch size of 1 image was used because the training images vary in their dimensions. Since the training data was sparse, augmentation of images and corresponding bounding boxes was done by randomly flipping them horizontally and vertically, rotating by 90 degrees, altering brightness, contrast, hue and saturation along with random cropping, padding, and scaling.

Grid anchors were generated at scales 0.25, 0.5, 1 and 2 with aspect ratios 1:2, 1:1 and 2:1. The reason for using these is to generate object proposals with varying scales and aspect ratios. This is believed to make the network robust in predicting objects that vary in scale and/or are obscured by other objects. In our case, since the weeds have various stages of development even in the same image and may be obscured by crops, the use of various scales and aspect ratios is important.

For generating the region proposals, a window of size 3x3 was convolved on the results of Conv4_x, followed by the activation step. This output was simultaneously passed to the classification and regression layers by a convolution of 1x1 producing scores for the "objectness" and the four coordinates of a bounding box of each object. The weights of the Conv4_x were initialized by the pre-trained model; however, the weights of the subsequent layers of the RPN were initialized by the truncated normal initialization with a standard deviation σ of 0.01.

The Intersection over Union (IOU) threshold for distinction between objects and non-objects was set to 0.7 which means those proposals whose IOU values were less than 0.7 in the ground truth bounding boxes, were considered as background. As the number of proposals for the background would be particularly high for the images that contain few objects of interest, only a subset of such proposals is considered. In this study, a ratio of 1:1 for the object and background was specified.

The classification of the objects was done by cropping and resizing the convolutional feature maps proposed by RPN and passing it through the Conv5_x layer. The output of this layer was passed through the softmax activation function to predict the class probabilities of the object and the same output was passed to a different regression layer for the final regression of the bounding box. The implementation was done on Tensorflow Object Detection API (HUANG et al. 2016).

3.3 Results and Discussion

Initially, the training was run with up to 200,000 iterations; the accuracy was simultaneously calculated along with the training by evaluating the models created during the training at various iterations using test data. The training loss started from a higher value in the first iteration and then reduced abruptly after few thousand iterations, after that it remained stagnant (see Figure 1). The test accuracy started from a low value and increased abruptly and then remained stagnant after few thousand iterations. Neither the loss nor accuracy improved or degraded even if the training was done for 450,000 iterations (see Figure 1).

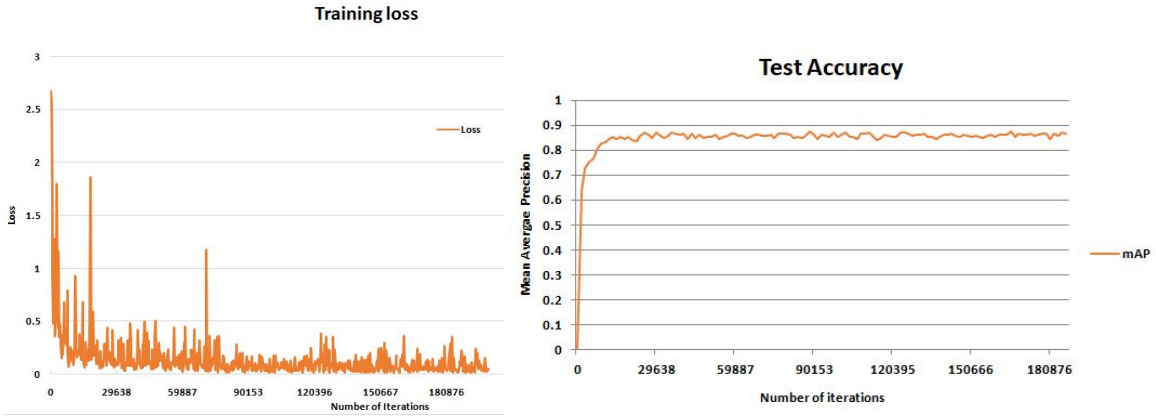


Fig. 1: Training Loss and Test Accuracy

In this setup, the training and test set were mutually exclusive sets of images where the weeds are large. We refer to such images as “large weed images”. It was expected that the network, after sufficient training, would learn the complex features such as shape and texture from the training images and would generalize even on the images where the weeds are significantly small. Hence, another set of such images, which we refer to as “small weed images” were added to the test set to assess the quality of the model. This set of images had 97 images with MATCH and STEME photographed along with wheat. The development stage of the weeds in this image set was similar to the previous set of images, the differences being the size of the weeds, the inclusion of the crops and occurrence of above-specified weeds in a single image.

The evaluation of the detection has been performed according to the PASCAL VOC metrics (EVERINGHAM et al. 2007). The Average Precision (AP) of each class and a mean Average Precision (mAP) for the entire class has been computed. The AP has been calculated by computing precision at various recall values, ranging from zero to one and then taking the weighted mean of such computed precisions, where the weights are the increase in recall values

from the previous steps. The mAP values are the average values of AP over all the classes. The precision is calculated by the number of true positives divided by the sum of true positives and false positives, whereas the recall is calculated by the number of true positives divided by the sum of true positives and false negatives. In our case, true positives refer to the detections that have an Intersection over Union value of at least 0.5. In case of multiple detections of the same class, a single detection is considered as true positive while all the others are considered as false positives. Other false positives are the ones which the model detects although there are no corresponding objects in the ground truth. The false negatives are the ground truth objects which the model did not detect.

Table 4 shows the AP and mAP of each of the weed class which was tested on two different datasets. Dataset 1 refers to the test data set as described in Table 2 and Dataset 2 refers to the test data set including the "small weed images". A third setup (Dataset 3) of the experiment was done in which the "small weed images" were randomly split in training and test data in ratio 80% and 20% respectively, and were added to the respective datasets of Table 2.

Tab. 4: Average Precision and mean Average Precision for three sets of test

	AP(%)					mAP(%)
	AMARE	**STEME**	**MATCH**	**POAAN**	**VIOAR**	**All weeds**
Dataset 1	89.4	74.6	86.3	74.2	96.3	84.2
Dataset 2	87.2	30.2	23.3	40.8	94.0	55.1
Dataset 3	88.5	70.1	78.9	78.9	95.4	82.3

It can be seen that the AP for AMARE and VIOAR does not differ much since the test data of dataset 2 does not contain weeds of small size. However, there has been a massive reduction in AP for STEME, MATCH, and POAAN, and therefore a reduction of the mAP for all weeds. This is due to the wheat in test data images being incorrectly classified as POAAN, as seen in Figure 2 (top row, left). The reason for the misclassification might be the similarities in leaf structure as well as spectral values of wheat and POAAN. Moreover, the size of STEME and MATCH was significantly small. This increased the false positives of POAAN and also the false negatives of STEME and MATCH were increased resulting in low precision and a low recall for the respective classes.

The results show that the accuracy of all the weed species improved and is comparable to the results obtained for the Dataset 1. The justification for the above results comes from the fact that since the "small weed images" are also trained, the Region Proposal Network is able to propose even the smaller objects, shown in Figure 2 (top row, right). Furthermore, inclusion of these images in the training dataset allows the network to assign regions with the wheat to background and hence decrease in false positives for POAAN. The slight decrease in the AP for STEME and MATCH as compared to the first experimental setup might be due to some misclassification among these two classes. The inclusion of "small weed images" in training has negligible effect on the detection of weeds in "large weed images", as shown in Figure 2 (bottom row).

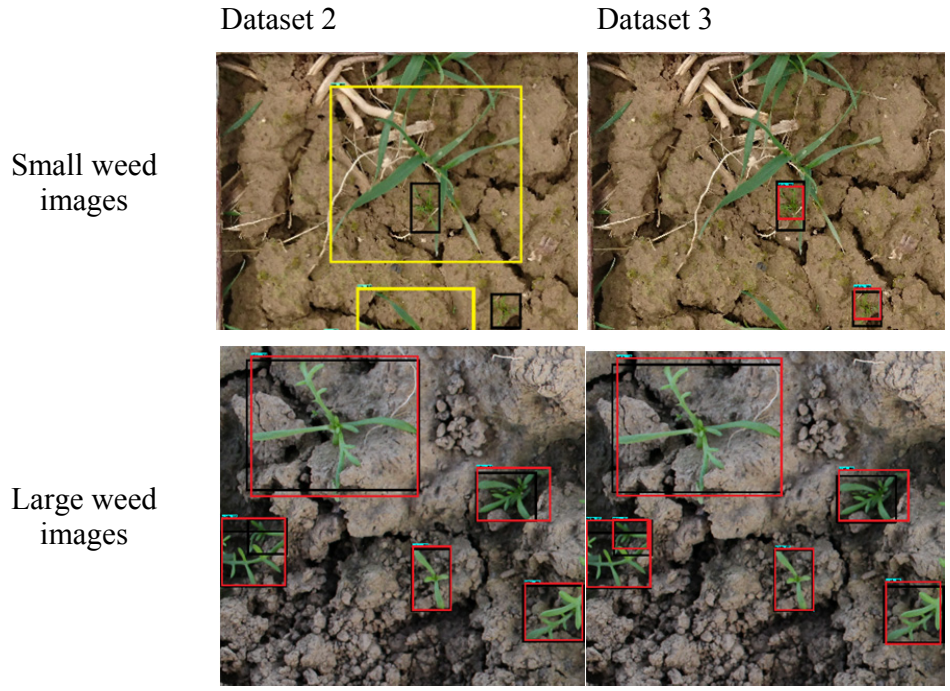


Fig. 2: Detections on test images (network trained on Dataset 2 and 3) with large and small weeds. The ground truth boxes are shown in black. The yellow boxes correspond to misclassification of wheat as POANN and the red boxes correspond to MATCH. The inclusion of "small weed images" for training in Dataset 3 improves the detection but has negligible effect in detection of weeds in "large weed images".

4 Pixel-based approach

Semantic segmentation aims at producing a classification map of the same size as the input image. The map assigns one out of a set of defined classes to each pixel as opposed to categorization networks which assign a single class to an entire image. Several architectures have been created for pixel-wise classification which consists of two main stages, the encoder for classification and the decoder for pixel-wise prediction and output of the segmentation results.

LONG et al. (2014) proposed fully convolutional networks (FCNs) for semantic segmentation. The architecture transforms common classification CNNs such as VGG, AlexNet and GoogLeNet into FCN by rewriting their fully connected layers into convolutional layers which form the encoder stage of the network. The encoder generates low-resolution feature maps which are then passed to the decoder for upsampling to get prediction maps of the same size as the input image. The upsampling is performed using transposed convolution layers also referred to as deconvolution layers through bilinear interpolation. The decoder refines the upsampled outputs by merging them with features from different stages in the encoder stage which are coarse but of high resolution.

Segnet (BADRINARAYANAN et al. 2015), another semantic segmentation architecture, differs from FCN in regards to the implementation of both the encoding and decoding stages. In the encoder stage, all the fully connected layers of a classification network are discarded. The lower

resolution feature maps from the encoder are upsampled in the decoder using unpooling layers as opposed to using deconvolutional layers in the FCN. The resulting encoder-decoder network is efficient in memory usage and computational time. This is due to fewer parameters after discarding the fully connected layers in the encoder and use of max-pooling indices from the encoder for non-linear upsampling in the decoder. The low memory usage, however, results in loss of accuracy compared to FCN which preserves the feature maps in the encoder stage by re-writing the fully-connected layers instead of discarding them. For this reason, FCN was used in this approach.

4.1 Data Preparation

For training a FCN, image pairs as data input and a corresponding ground truth mask as the label input is required. Tab. 5 shows the number of training image samples created from the field images and the plant seedlings dataset (GISELSSON et al. 2017)

Tab. 5 : The training Dataset

EPPO Code	Name	No. of Images	No. of Images + Augmentations
MATCH	Wild chamomile	179	1611
POAAN	Annual meadow grass	212	1908
STEME	Common chickweed	200	1800
VIOAR	Field pansy	168	1512
	Wheat	201	1809
TOTAL		960	8640

The image samples were cropped to contain exactly one weed or wheat and the background unlike in the object-based approach in Section 3.1 where a training image contained one or more plants. Therefore, the number of images for each species in this approach is higher than the one used in the former approach. The reason for cropping in the indicated way was to allow segmentation through maximum likelihood classification using ArcGIS Desktop. Symbolic values corresponding to the six classes, five classes in the training dataset and the background, were assigned to the pixels of the segmented images. These symbolic images were used in the sequel as ground truth labels. In order to expand the training dataset effectively, data augmentation was performed where both geometric and pixel value changes were applied. Increasing the dataset has been used to reduce overfitting and speed up convergence while training models. The geometric augmentation techniques included rotating by 90º, 180º and 270º, horizontal and vertical flips of both the images and the labels. In the second set of augmentations, the pixel values were altered by performing colour enhancement, colour equalization and applying a Gaussian blur. Unlike in the object-based approach where the augmentations were done on the fly, these augmentations were performed prior to the training process. The image dataset was split with 80% as the training dataset and 20% as the test dataset for each class and stored in light memory-mapped databases (LMDBs).

4.2 Implementation

The FCN architecture was implemented by leveraging models trained on bigger datasets through transfer learning. The deep learning framework Caffe (JIA et al. 2014), developed by the Berkeley Vision and Learning Centre (BVLC), was used for the implementation while the model definitions and pre-trained weights were obtained from Caffe's Model Zoo.

In the first set up, a VGG-16 model by SIMONYAN AND ZISSERMAN (2014) used for the ImageNet Challenge 2014 was transformed into a fully convolutional network. The VGG-16 network has three fully connected layers at the end, FC6, FC7 and FC8. The last layer FC8, used as the classifier for ImageNet's 1000 classes, was discarded while FC6 and FC7 were re-written into convolutional layers and the weights transferred using the net surgery script provided in Caffe. By reshaping and retaining FC6 and FC7, the feature extraction weights from the VGG-16 are preserved. The weights were used to initialize the encoding stage of the "at-once" FCN-8s architecture used for training with PASCAL VOC 2011 dataset (EVERINGHAM et al. 2011). Gaussian and bilinear weight fillers were used to initialize the weights of the convolutional and deconvolutional layers respectively in the decoder. A combination of a low learning rate of 10^{-10}, a high momentum of 0.99 and a batch size of one was used for the training. This setup is referred to as VGG16-FCN in the next sections.

In the second experimental setup, the weights of a FCN-8s model trained on 20 classes of the PASCAL VOC dataset were fine-tuned using the weeds dataset. This was done by a full transfer of weights and adding a convolution layer as a classifier with the six classes. The new classifier was initialized with a Gaussian weight filler. A low learning rate of 10^{-14} has been used together with a high momentum of 0.99 to fine-tune the entire network. The low learning rate was necessary to prevent fast distortions to the fully transferred weights.

4.3 Results & Discussion

In the training phase, choosing and optimizing training parameters required monitoring by plotting the training loss, test loss and accuracy which indicate the performance in real time. The goal of this is to decrease the training and test losses while increasing the model's accuracy. The learning rate remains the most important parameter and estimating an optimal value involves testing a range of values and different learning rate policies. Very low learning rate means the training takes long to converge or gets stuck at a local minimum while a very high learning rate leads to the loss function rising very fast. In both setups, a combination of a fixed learning policy rate, a low learning rate and a high momentum gave better results.

Caffe offers the flexibility to specify which layers to either fine-tune or freeze the weights during training. Fine-tuning the whole network gave better results which can be attributed to the difference in nature of training classes of our dataset from the ImageNet (RUSSAKOVSKY et al. 2014) and PASCAL VOC (EVERINGHAM et al. 2011) datasets used to train the VGG-16 and FCN-8s models. Attempt to perform training without initializing new layers with weight fillers resulted in high loss values. Gaussian and Xavier weight fillers were both tested to initialize new convolution layers both yielding similar performance.

The performance of the models from the two experiments was evaluated on the test dataset using mean Intersection over Union (mIU) metric. The metric also known as Jaccard similarity coefficient is the ratio of correctly classified pixels to the total number of ground truth and

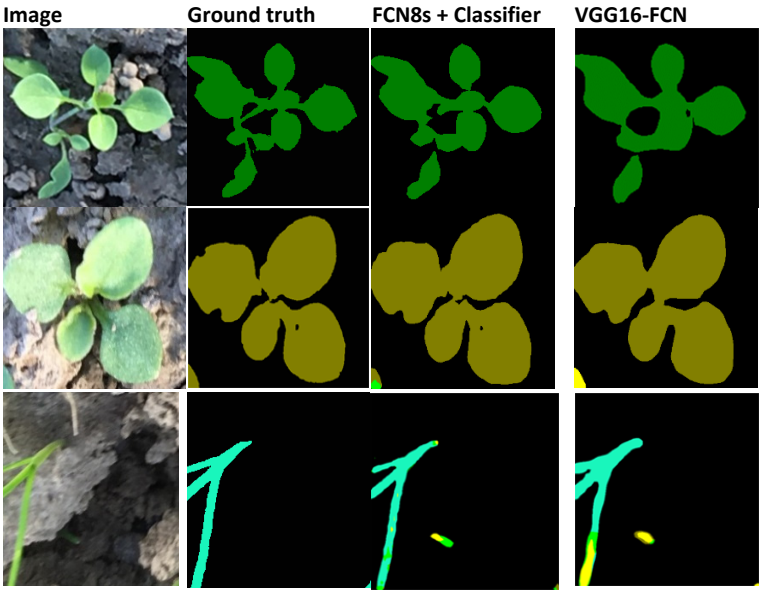
predicted pixels. It provides an accurate measurement that penalizes false positives making it suitable for evaluating semantic segmentation models. The pixel accuracy can be a misleading measure of model performance if one class has a higher number of pixels in the images compared to other classes. In the training and test datasets, most of the pixels belong to the soil which is represented in the class background. The frequency weighted IU, considers the IU of each class weighted by the number of pixels in that class. This gives an overall accuracy with larger classes receiving larger weights hence the high results in Tab. 6, due to the influence of the background class which does not give an accurate performance of the models. The models predict most of these background pixels correctly but misclassify some of the weed and wheat pixels. In addition to the metrics, visualization of the segmented images was compared to the ground truth labels.

- Mean accuracy: $(\mathbf{1}/\boldsymbol{n_{cl}})\sum_i \boldsymbol{n_{ii}} / \boldsymbol{t_i}$
- Mean Intersection over Union (IU): $(\mathbf{1}/\boldsymbol{n_{cl}})\sum_i \boldsymbol{n_{ii}} / (\boldsymbol{t_i} + \sum_j \boldsymbol{n_{ji}} - \boldsymbol{n_{ii}})$
- Frequency weighted IU: $(\sum_k \boldsymbol{t_k})^{-1} \sum_i \boldsymbol{t_i}\, \boldsymbol{n_{ii}} / (\boldsymbol{t_i} + \sum_j \boldsymbol{n_{ji}} - \boldsymbol{n_{ii}})$

Here $\boldsymbol{n_{cl}}$ is the number of classes, $\boldsymbol{t_i}$ total number of pixels of class $\boldsymbol{i}$ in the ground truth label, $\boldsymbol{n_{ii}}$ is the number of correctly predicted pixels of class $\boldsymbol{i}$, and $\sum_j \boldsymbol{n_{ji}}$ is the total number of pixels predicted to belong to class $\boldsymbol{i}$ (LONG et al. 2014).

Tab. 6 : A comparison of evaluation metrics after 30,000 training iterations of each experimental setup calculated on the test dataset with 1,725 Images

Models	Mean accuracy	Mean IU	Frequency weighted IU
VGG16-FCN	77.60	24.75	91.7
FCN8s + Classifier	77.43	25.32	92.84



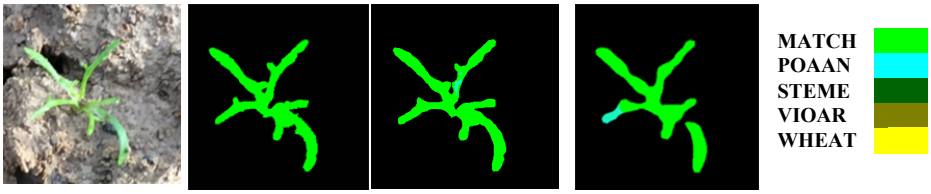


Fig. 3: Visual comparison of segmentation results of images from the test dataset taken in the same conditions as the training dataset

From the metrics in Tab. 6 and inference results in Fig. 3, the *FCN8s + Classifier* setup yields slightly better results. This method employs full weights transfer with only one layer being randomly initialised compared to the first method where the entire decoder weights are randomly initialized. Weeds sharing similar leaf structures such as elongated leaves in classes POAAN, Wheat and MATCH are susceptible to misclassifications. This can be seen in Fig. 3 where some POAAN pixels get misclassified as wheat or MATCH and the same occurs in MATCH segmentation. The models were further tested on images from the plant seedlings dataset (GISELSSON et al. 2017) as a control check for overfitting. The results in Fig. 4 show the ability of the models to predict a majority of the pixels correctly with the *FCN8s + Classifier* setup yielding better spatial detail and less misclassified pixels.

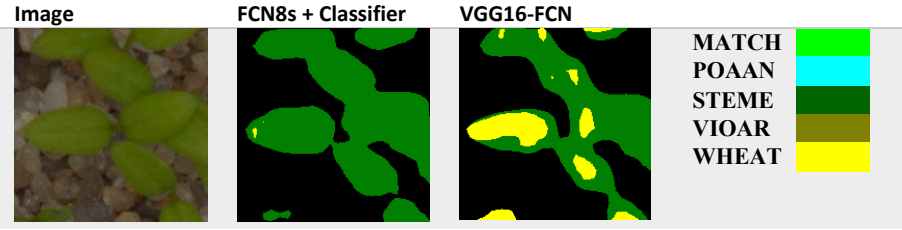


Fig. 4: Segmentation results for a STEME image from the plant seedlings dataset

5 Conclusion

The experiments have demonstrated success in transferring pre-trained CNNs to address weed classification with limited training images. Further improvements could be achieved by expanding image datasets. A large collection of weed and crop images varying in growth stages, soil textures, resolutions and conditions would help in training robust CNNs for use in precision agriculture.

The resolution of training images poses some restrictions both due to the memory requirements for the computational process as well as in inferring the generated model. The models trained in images where the weeds are large, have difficulties in detecting the weeds when the size of the weeds is significantly small. However, the inclusion of some images with small weeds in training seems to improve the performance of such models. Further experimentation is required to be able to detect objects that vary significantly in scale as compared to the trained objects so as to create more robust models.

There are still challenges to detect objects when they are obscured by other objects. Various parameters such as initialization of the weights, regularization of the network affect the learning process and have to be further investigated. Other semantic segmentation networks have been proposed either as extensions to the FCN models or entirely different architectures focusing on reducing computational resources and time, increasing accuracy or a combination of the three. These could be explored and compared.

6 Acknowledgements

The project was supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme.

7 References

ÅSTRAND, B. & BAERVELDT, A.J., 2002: An Agricultural Mobile Robot with Vision Based Perception for Mechanical Weed Control, Autonomous Robots, **13**, 21-35.

BADRINARAYANAN, V., KENDALL, A. & CIPOLLA, R. 2015: SegNet. A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. arXiv:1511.00561 [cs.CV].

EVERINGHAM, M., VAN GOOL, L., WILLIAMS, C.K.I., WINN, J. & ZISSERMAN, A., 2007: The Pascal Visual Object Classes Challenge 2007 (VOC 2007) results (2007), available at: http://www.pascal-network.org/challenges/VOC/voc2007/workshop/index.html.

EVERINGHAM, M., VAN GOOL, L., WILLIAMS, C.K.I., WINN, J. & ZISSERMAN, A., 2011: The PASCAL Visual Object Classes Challenge 2011 (VOC 2011) Results, available at: http://www.pascal-network.org/challenges/VOC/voc2011/workshop/index.html.

GISELSSON, T.M., DYRMANN, M., JØRGENSEN, R.N., JENSEN, P.K. & MIDTIBY, H.S., 2017: A Public Image Database for Benchmark of Plant Seedling Classification Algorithms. arXiv:1711.05458 [cs.CV].

GLOROT, X. & BENGIO, Y., 2010: Understanding the Difficulty of Training Deep Feedforward Neural Networks, in Teh, Y.W. and Titterington, M. (Eds.). Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, PMLR, Chia Laguna Resort, Sardinia, Italy, 249-256.

GOLZARIAN, M.R. & FRICK, R.A., 2011: Classification of Images of Wheat, Ryegrass and Brome Grass Species at Early Growth Stages Using Principal Component Analysis, Plant Methods, **7**(1), 28.

HE, K., ZHANG, X., REN, S. & SUN, J., 2016: Deep Residual Learning for Image Recognition. arXiv:1512.03385 [cs.CV]

HUANG, J., RATHOD, V., SUN, C., ZHU, M., KORATTIKARA, A., FATHI, A., FISCHER, I., WOJNA, Z., SONG, Y., GUADARRAMA, S. & MURPHY, K., 2016: Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors. arXiv:1611.10012 [cs.CV].

JIA, Y., SHELHAMER, E., DONAHUE, J., KARAYEV, S., LONG, J., GIRSHICK, R., GUADARRAMA, S. & DARRELL, T., 2014: Caffe: Convolutional Architecture for Fast Feature Embedding, ACM, available at: http://dl.acm.org/ft_gateway.cfm?id=2654889&type=pdf.

LIN T., 2015: LabelImg. Github Repository, https://github.com/tzutalin/labelImg, Accessed on 01/27/2018.

LIN, T.-Y., MAIRE, M., BELONGIE, S., HAYS, J., PERONA, P., RAMANAN, D., DOLLÁR, P. & ZITNICK, C.L., 2015: Microsoft COCO. Common Objects in Context. arXiv: 1405.0312 [cs.CV].

LONG, J., SHELHAMER, E. & DARRELL, T., 2014: Fully Convolutional Networks for Semantic Segmentation. arXiv:1411.4038 [cs.CV].

REN, S., HE, K., GIRSHICK, R. & SUN, J., 2015: Faster R-CNN. Towards Real-Time Object Detection with Region Proposal Networks. Advances in Neural Information Processing, 91-99.

RUSSAKOVSKY, O., DENG, J., SU, H., KRAUSE, J., SATHEESH, S., MA, S., HUANG, Z., KARPATHY, A., KHOSLA, A., BERNSTEIN, M.S., BERG, A.C. & LI, F.-F., 2014: ImageNet Large Scale Visual Recognition Challenge. arXiv:1409.0575 [cs.CV].

SIMONYAN, K. AND ZISSERMAN, A., 2014: Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556 [cs.CV].

SØGAARD, H.T., 2005: Weed Classification by Active Shape Models. Biosystems Engineering, **91** (3), 271-281.

Object-Based Mobile Augmented Reality for a 3D Model

ROSANNY SIHOMBING[1] & VOLKER COORS[1]

Abstract: Augmented Reality (AR) is a convincing technology for manipulating user's visual perception of the real world. A single image-based tracking approach is not sufficient to recognise a 3D object from different viewpoints. The basic issue of this approach is it can no longer recognise a 3D model when the viewing perspective is changed since the image captured by the camera is no longer match the reference image. This paper presents an object recognition approach to track markerless 3D building blocks as the AR target. This was achieved by extracting the feature points of the building blocks through a preliminary scanning process using Vuforia Object Scanner. Then a dataset of these feature points is used to recognize the building blocks so that the tracking process can be performed and virtual models can be imposed. To make the AR scene more realistic, the virtual model shadow is added using shadow mapping approach. The experiment shows the presented method is able to recognize the building blocks from side to side, which leads to a robust tracking performance, while the shadow mapping successfully adds realistic shadows to the scene. This robust performance also resulting in a stable appearance of the augmented virtual model.

1 Introduction

A complex and large physical 3D city model is often used as a tool for urban planning, since it maps important characteristics of the city, such as green areas, main roads, buildings and towers, and another significant landmarks (Figure 1). However, due to its complexity, any changes made to the city model would be a costly process in terms of money, time, and resources. Therefore, an inexpensive and simple way to display possible addition or modification of the physical city model would give the city planners efficiency and effectiveness they need when doing urban planning.

The basic concept of AR is to manipulate human visual perception of the real world by mixing real object and computer-generated virtual object in the real world, in such a way that they are accurately aligned in real-time (AZUMA et al. 2001). Over the last decade, both operating system and hardware quality of mobile devices are increasing rapidly, making them capable of doing all heavy computations needed for AR (BILLINGHURST et al. 2015).

1.1 Previous Work

On our previous work (SIHOMBING 2016), we have implemented an AR mobile application using a markerless image-based approach for indoor urban planning. It enables multiple city planners to view and assess city plans directly on a 2D city map or an indoor 3D physical city model in real-time using an android device (Figure 2). It was done by obtaining the feature points of the

[1] Hochschule für Technik Stuttgart, Schellingstr. 24,
D-70174 Stuttgart, E-Mail: [rosanny.sihombing, volker.coors]@hft-stuttgart.de

reference image and utilise them to recognize the object target before performing the tracking and adding the virtual model to the AR scene.

While this method gives a robust tracking performance and a convincing result with the 2D city map, it does not perform well with the 3D city model (Figure 3). The application ability is limited to the scene recognition from a particular viewpoint when using the 3D city model as a target. This limitation happened due to a perspective issue when using a single image as a reference for a 3D object. When the viewing perspective changes, so does the 3D object image captured by the camera. As the result, the 3D object cannot be recognised anymore through its reference image. Therefore, on this case a single reference image is not sufficient to have a robust tracking performance. To solve this problem and improve what has been achieved, a more suitable recognition approach needs to be applied.

The main goal of this paper is improving the result we have achieved (SIHOMBING 2016) by (1) using a 3D city model as the object target for urban planning and (2) rendering the shadow of the virtual 3D model in order to add realism to the whole scene. The remaining sections of this paper is structured as follows. Section 2 explains the methods to achieve the objective. The experiments and corresponding evaluations are presented on section 3, while the conclusion and future works can be found on section 4.

Fig. 1: Physical city model of Hamburg [source: www.hamburg.de/stadtmodell]

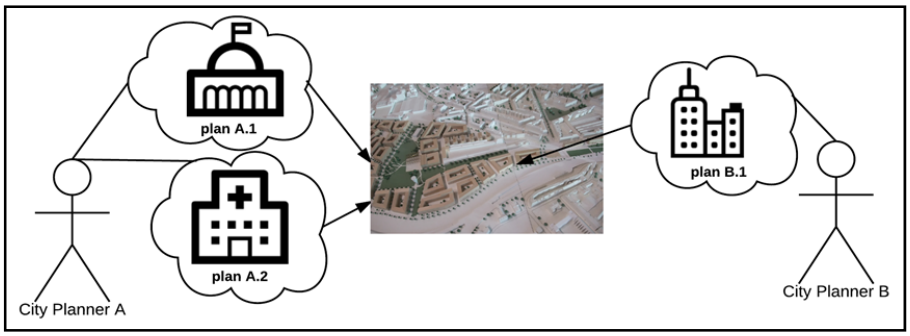


Fig. 2: City planners with their own ideas on plans on the physical city model

Fig. 3: 2D image-based approach for both city map and physical building blocks. (above) An A4 size 2D map is used as the target, which can be recognised quickly and the virtual building blocks can be imposed without being disappear. (below) A 2D reference image of untextured building blocks is used as the target. The physical building blocks can be recognised only from a very limited perspective and the virtual model was clearly misplaced.

2 Methods

The recognition method presented on this paper aim to recognise and track untextured building blocks, which act as the object target, so that the augmented virtual model can be seen from a broader perspective. On the next step, the shadow of the virtual model will be casted to the ground to level up the realism.

2.1 Recognising 3D Object

The more complex geometric shape a 3D object has, the more feature points it has. These feature points determine how quick a 3D object can be recognised. We utilize Vuforia Object Scanner (PTC INC. 2018) in order to obtain 3D object feature points by scanning the object. After gathering the feature points, a quick test can be performed to check whether or not the 3D object can be recognised through those feature points. If the desired performance has not achieved yet, then more feature points must be gathered by continuing the previous scanning process (Figure 4). The detail of this process can be found on (PTC INC. 2018).

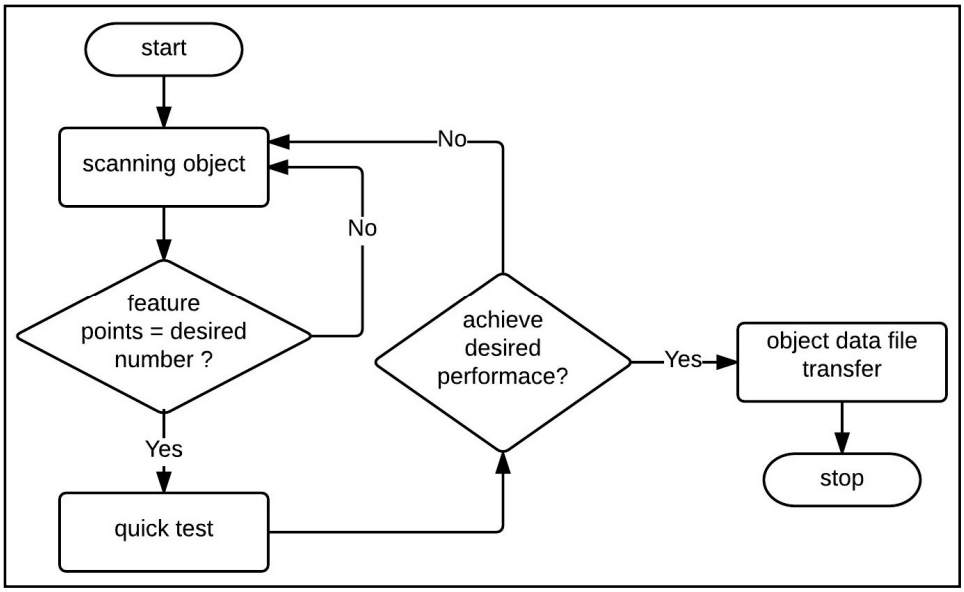


Fig. 4: Scanning process to obtain feature points

2.2 Tracking The Target and Placing The Virtual Model

Tracking is a looping process which starts as soon as the object target is recognised. When the tracking is running, the pose matrix is always updated depending on the camera pose (Figure 5). Pose matrix represents the pose of an object target as seen from the camera. In eq. 1 this matrix is denoted by $\boldsymbol{P}$ and calculated based on the camera pose towards the object target. This 3x4 matrix consists of a 3x3 rotation matrix, denoted by $\boldsymbol{r}$, and a 3x1 translation matrix, denoted by $\boldsymbol{t}$. The rotation matrix shows how the object target is rotated, while the translation matrix indicates the object target position relative to the camera.

$$P = \begin{bmatrix} r11 & r12 & r13 & tx \\ r21 & r22 & r23 & ty \\ r31 & r32 & r33 & tz \end{bmatrix} \quad (1)$$

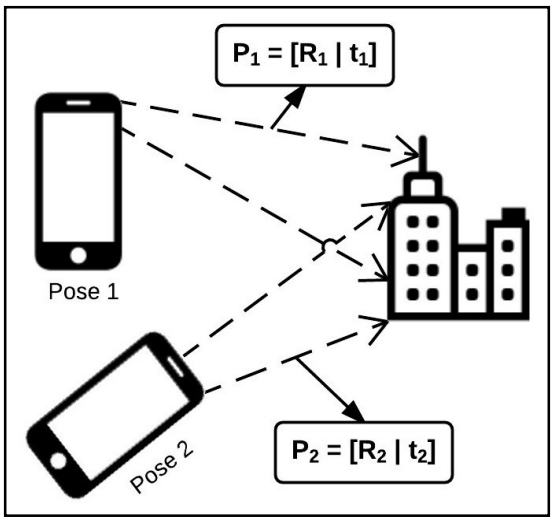


Fig. 5: Camera pose and pose matrix

To place the virtual model in the desired location, it must be transformed so that its centre point match the centre point of the 3D object target. The same requirement goes to the scale and orientation of the virtual model, they need to be transformed so the virtual model fits the 3D object target. These transformations process can be done manually in Unity3D (UNITY 2018) using its Graphical User Interface. The virtual model used on this paper is provided by the Hamburg City and reconstructing one is outside the scope of this paper.

2.3 Shadow Mapping

As shadow is important for human spatial perception, casting the shadow of the virtual is expected to level up the realism of the AR scene (SUGANO et al. 2003). The technique used to cast the shadow is the common shadow mapping technique (WILLIAMS 1978). To implement this technique, it needs to be adjusted to AR where a shadow receiver that is invisible to the user must be added to the scene. Therefore, an invisible planar object which acts as the ground must be created and placed in world space centre (0, 0, 0) under the virtual model, while the existing virtual model is assigned as the shadow caster and also receiver. The light source's position is configured manually. This invisible shadow receiver approach on AR is similar to (CASTRO et al. 2012) and also (HALLER et al. 2003) even though shadow volume technique was used to cast the shadow. The following steps describe the approach to map virtual shadow on this paper:

1) Create an invisible planar plane as the ground. Put it on coordinate (0, 0, 0) and assign it as shadow receiver.
2) Define light source position.
3) Assign the virtual model as both shadow caster and receiver.

3 Experiments and Evaluations

The mobile application has been implemented using Vuforia SDK (PTC INC. 2018) as the AR framework and Unity3D (UNITY 2018) as the 3D engine on a Nexus 9 tablet running Android 6.0.1. After the implementation, we had 2 major experiments which aim to (1) demonstrate the tracking performance of the proposed approach and (2) evaluate the shadows in regard to making the whole AR scene more realistic. Both experiments used untextured 3D model provided by The City of Hamburg as the object target (Figure 6). This 3D model is a copy of some part of Hamburg physical city model as pictured in fig. 1.

Fig. 6: 3D object target

3.1 Tracking

3.1.1 Experiment

The virtual model for this experiment consists of 8 textured virtual building blocks in 3DS file format and provided by The City of Hamburg. These 8 virtual objects are 1 entity which depends on each other. The virtual model was imposed to the object target just like a scenario of urban planning which already explained on Section 1 and pictured in fig. 2. Two building blocks of the virtual model correspond to 2 building blocks of the object target. It is expected that these 2 virtual building blocks are completely projected to the corresponding object target building blocks, while the other 6 virtual building blocks are projected to their corresponding map location. The result and demo video of this experiment is available at the following URL: https://vimeo.com/221005614. Some images extracted from the video are presented in this paper as well (Figure 7).

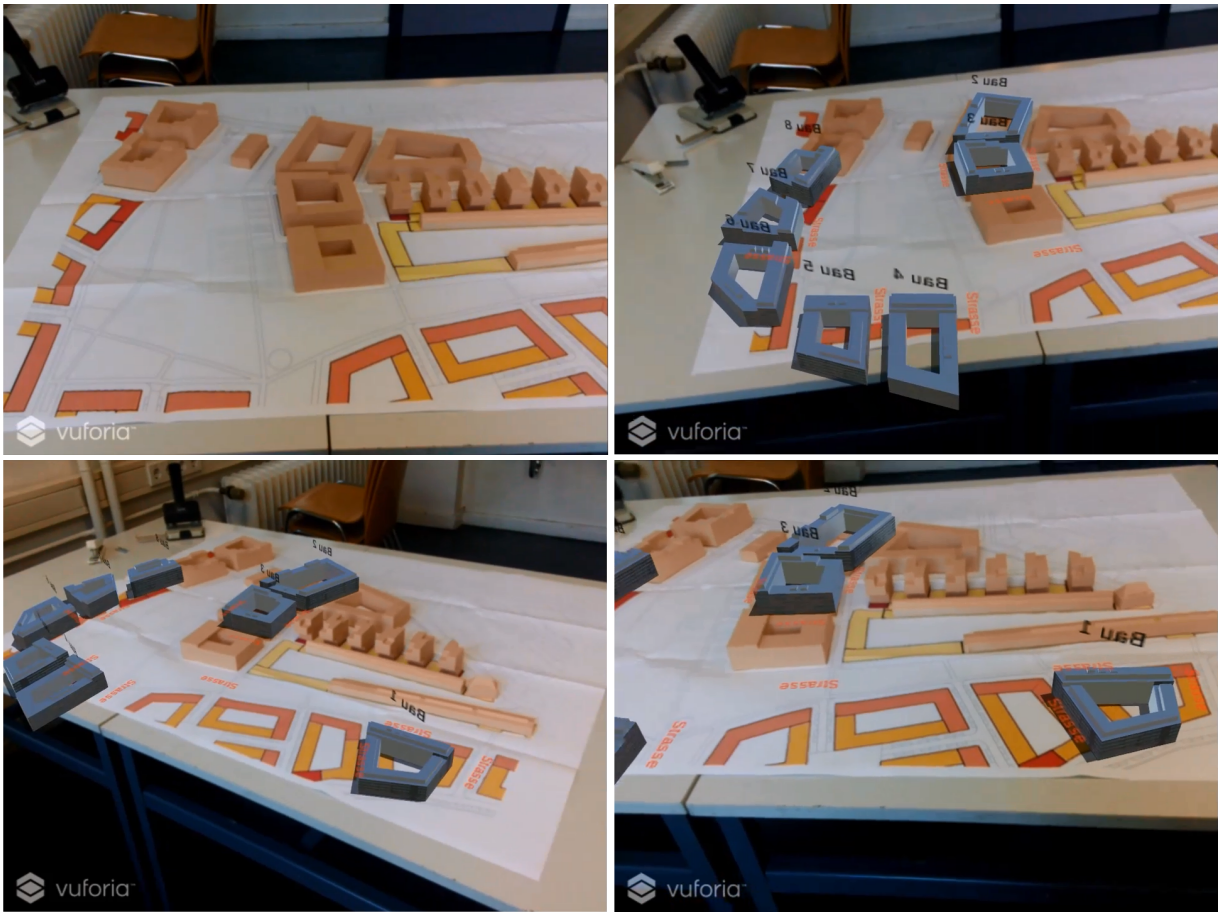


Fig. 7: (top-left) Mobile application is getting started. (top-right) The virtual is imposed to the object target as soon as the target is recognized. (bottom-left) Tracking the object target from another viewpoint. (bottom-right) Getting closer to the object target.

3.1.2 Evaluation

As shown in the video, the experiment shows the untextured 3D model can be recognised as the object target and the virtual model can be imposed to the target. The tracking performance was convincing enough since the virtual model can be viewed from side to side without flickering or being disappeared during the tracking. This performance gives the users more flexibility in observing the virtual model since they can move around to see it from different viewpoints. However, while 2 virtual buildings in the middle can be projected accurately, the other 6 virtual buildings were slightly misplaced.

To inspect this accuracy issue, we manually map the virtual model to the 90cm x 110cm 2D map of the target object location using a GUI tool in order to see the desired location of each virtual building on the map and how accurate the corresponding location is. The mapping (Figure 8) shows 6 virtual buildings do not overlap perfectly with their corresponding map location, while the other 2 virtual buildings in the middle match very well. Therefore, this accuracy issue is still acceptable since the issue lies on the virtual model rather than an issue with the tracking or locating the virtual model.

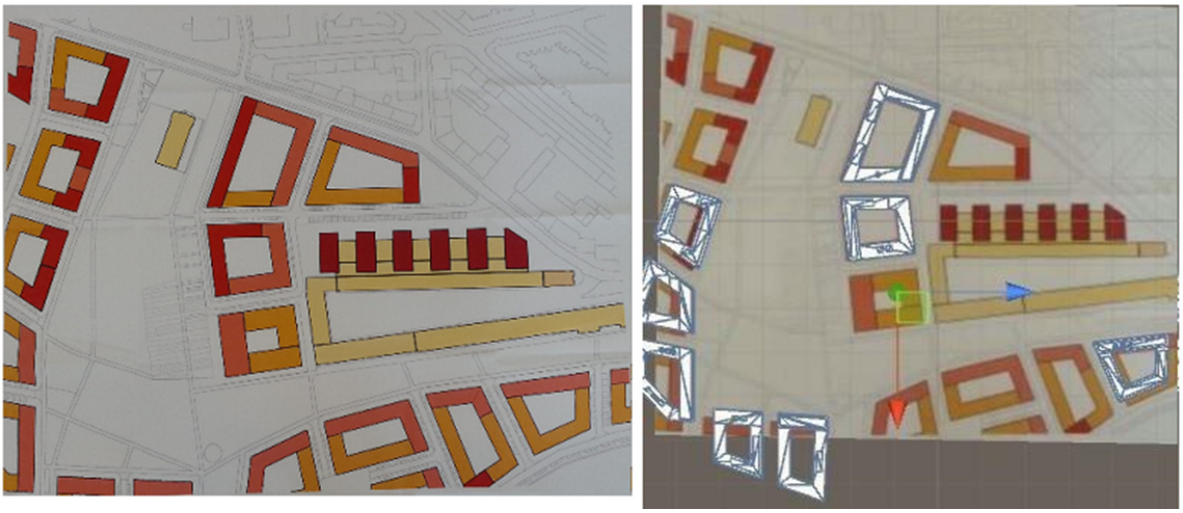


Fig. 8: (left) The location map of the 3D model. (right) Six virtual model do not correspond accurately to their location, while the other 2 can be located accurately.

3.2 Shadow Mapping

3.2.1 Experiment

Shadow Mapping was implemented by using Unity3D. On the previous experiment, the shadow of the virtual model was already casted to the ground. However, the shadow was not so visible due to the height of the virtual model. To make it more obvious, we used a cube and a cylinder imported from Unity3D library as the virtual models. The cube and the cylinder are higher than the actual virtual model and were placed next to each other in the AR scene. These addition virtual models were assigned as both shadow caster and shadow receiver (Figure 9). There was no particular point registration needed in this experiment since the objective is only to see how well the shadow mapping works when the virtual models are placed on the ground (x, y, 0).

3.2.2 Evaluation

As pictured in fig. 9, the shadow is clearly visible and mapped realistically to the invisible ground and to the visible virtual object. This promising performance is also shown in the tracking experiment, in which the shadows of the virtual buildings can be seen as well. The shadow appearances are different which reflect the shadow behaviour in the real world. Furthermore, the shadow also helps to emphasize the spatial relationship between the objects.

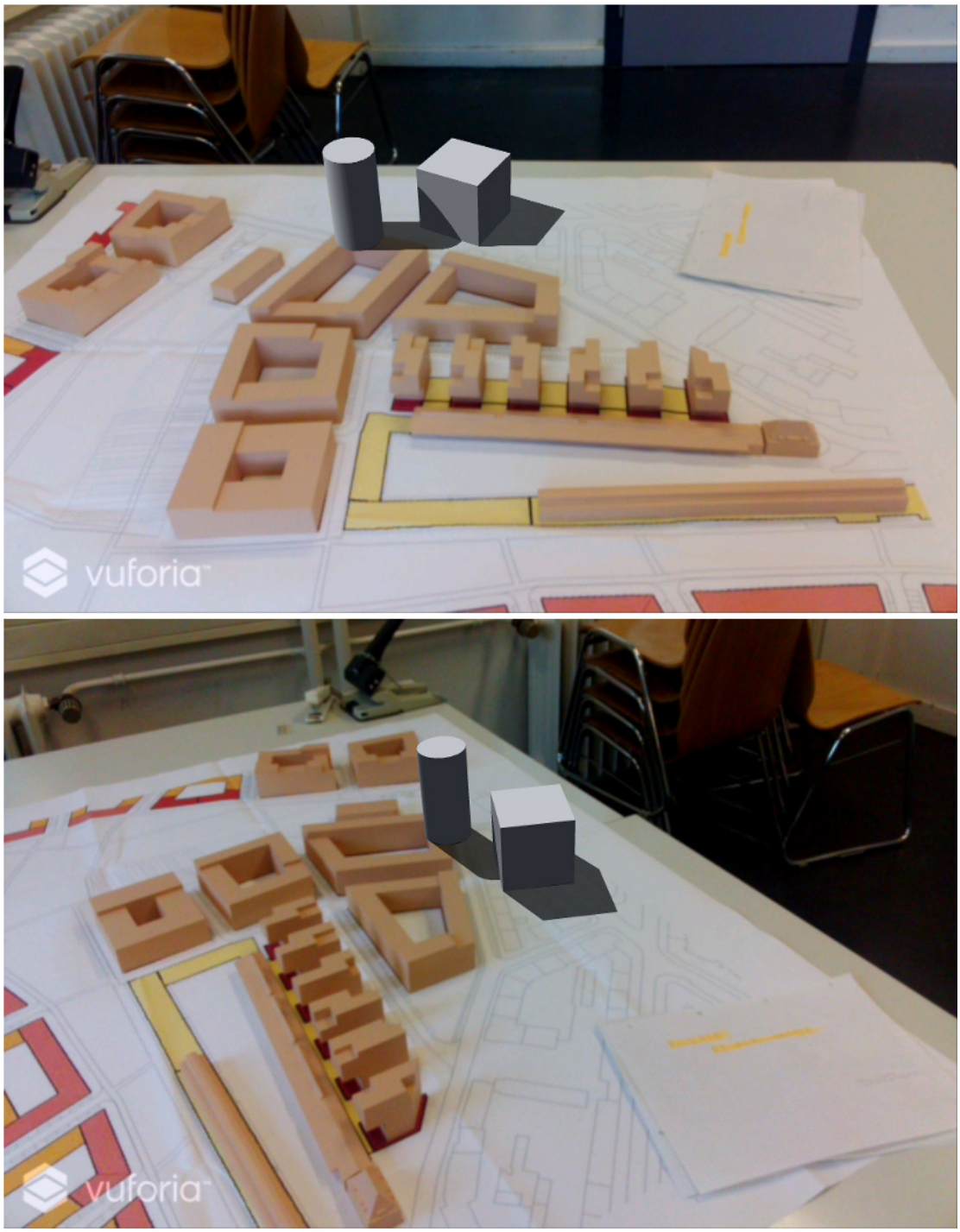


Fig. 9: A cylinder and a cube are assigned as both shadow caster and receiver, so each can also receive shadow from another virtual object.

4 Conclusion and Future Work

Object-based mobile AR for a 3D model and shadow mapping have been implemented. The 3D object recognition approach gives a convincing performance and result when working with 3D model. It is able to track the 3D model as the object target from broader viewpoint. As the result, the augmented 3D model can be viewed without being disappeared. The augmented shadow made the AR scene more realistic by emphasizing the spatial relationship between the object target, the virtual object and the ground, which is essential for urban planning.
As already mentioned, locating the virtual model on this paper were done manually. Therefore, in our further research we would like to make this process automatic so that it would be easier to add another virtual models to the AR scene. It is also our interest to have a real-time shadow and map it to the real objects so that the interaction between real and virtual object can be more realistic.

5 Acknowledgement

We would like to thank Dipl.-Ing. Walter Sieh and his team at Landesbetrieb Geoinformation und Vermessung of Hamburg city for their support and also for providing the 3D model, so that it was possible to have the experiments in Stuttgart.

6 References

AZUMA, R., BAILOT, Y., BEHRINGER, R., FEINER, S. K., JULIER, S. & MACINTYRE, B., 2001: Recent Advances in Augmented Reality. IEEE Computer Graphics and Applications, **21**(6), 34-37.

BILLINGHURST, M., CLARK, A. & LEE, G., 2015: A Survey of Augmented Reality. Foundations and Trends® in Human–Computer Interaction, **8**(2–3), 73–272.

CASTRO, T. K. D., FIGUEIREDO, L. H. D. & VELHO, L., 2012: Realistic shadows for mobile augmented reality. In Proceedings - 2012 14th Symposium on Virtual and Augmented Reality, SVR 2012, 36–45.

HALLER, M., DRAB, S. & HARTMANN, W., 2003: A real-time shadow approach for an augmented reality application using shadow volumes. Virtual Reality Software and Technology, 56.

PTC INC., 2018: Vuforia Object Scanner. Accessed August 2017. https://library.vuforia.com/articles/Training/Vuforia-Object-Scanner-Users-Guide.

SIHOMBING, R., 2016: Development of Image-Based Mobile Augmented Reality Application For City Map and City Model. MSc Thesis, Hochschule für Technik Stuttgart.

SUGANO, N., KATO, H. & TACHIBANA, K., 2003: The effects of shadow representation of virtual objects in augmented reality. In Proceedings - 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2003, 76–83.

UNITY, 2018: Unity Game Engine. Accessed August 2017. https://unity3d.com/de.

WILLIAMS, L., 1978: Casting Curved Shadows on Curved Surfaces. ACM SIGGRAPH Computer Graphics, **12**(3), 270-274.

Automatisierte Auswahl von Methoden zur Überlagerung von Geodaten mit unterschiedlichen Bezugsgeometrien

MAIKE SCHOTTEN[1] & JOCHEN SCHIEWE[1]

Zusammenfassung: Durch die Verknüpfung verschiedener Datensätze können oft vielfältige neue Informationen gewonnen werden. Zu dieser Überlagerung gehört nicht nur die geometrische Verschneidung, sondern auch die Neuverteilung der Attributwerte.
Da es trotz einer Reihe existierender Überlagerungsalgorithmen an einer kompakten und homogenen Bereitstellung der unterschiedlichen Methoden fehlt, werden in diesem Paper die Erstellung eines aufgabenorientierten und datenbasierten Regelapparates zur Methodenauswahl sowie die Dokumentation und Darstellung der erzielten Überlagerungs-Genauigkeiten beschrieben.

1 Problemstellung

Oft können erst durch die Verknüpfung von unterschiedlichen Datensätzen interessante und räumlich differenzierte Zusammenhänge erkannt und daraus weiterführende Ergebnisse gewonnen werden. Bei solchen Überlagerungen tritt allerdings häufig das Problem auf, dass die zugrunde liegenden Daten für unterschiedliche Bezugsgeometrien (z. B.: Landkreisebene vs. PLZ-Bereiche) vorliegen und somit die Zuordnung der jeweiligen Attributwerte nicht ohne Weiteres möglich ist.
Es existiert bereits eine Reihe von Überlagerungs-Algorithmen (vgl. FLOWERDEW & GREEN 1994; LI et al. 2007; TOMIO et al. 2016), allerdings fehlt es den Nutzern an einer leicht handhabbaren Software, die die notwendigen Methoden kompakt zusammenstellt, sowie an wissenschaftlich fundierten Empfehlungen zur fallspezifischen Methoden-Auswahl. Ferner findet in der Regel auch keine Dokumentation der resultierenden Interpolationsgenauigkeiten statt, sodass Bewertungen des Überlagerungs-Ergebnisses nicht möglich sind.

2 Methodik

Die beiden Basisoperationen einer Überlagerung sind einerseits die Zusammenführung kleinerer zu größeren Raumeinheiten (Aggregation, Aufsummierung) und andererseits der umgekehrte Prozess, die Zerlegung (Disaggregation) in kleinere Einheiten. Bei den meisten Überlagerungen findet eine Kombination der Operationen statt.
In den gängigen Softwareprodukten ist für diesen Zweck eine einfache Flächengewichtung implementiert. Über die Einführung von Zusatzwissen können die Ergebnisse jedoch an die realen Bedingungen (z. B. Bevölkerungsverteilung) angepasst und damit verbessert werden.

[1] HafenCity Universität Hamburg, Labor für Geoinformatik und Geovisualisierung, Überseeallee 16, D-20457 Hamburg, E-Mail: [Maike.Schotten, Jochen.Schiewe]@hcu-hamburg.de

Die verschiedenen Methoden setzen unterschiedliche Ausprägungen des Zusatzwissens voraus. Daneben gibt es Zusammenhänge zwischen der Interpolationsgüte und (geo-)statistischen Parametern der Eingangsdatensätze. Je stärker der Grad der Disaggregation ist, desto größer ist der zu erwartende mittlere Fehler der Interpolationsergebnisse, da im Gegensatz zur reinen Aggregation eine erhöhte Notwendigkeit zur Interpolation zu erwarten ist. Hierzu wird die Homogenität zwischen Quell- und Zielzonen sowohl über die Anzahl der Polygone wie auch über die Standardabweichung der Flächengrößen der Polygone betrachtet.
Für die Bestimmung der internen Genauigkeit werden die Ergebnisse aller Interpolationsmethoden aus mehreren Testgebieten verglichen und auf Abweichungen untersucht. Durch die Analyse zusätzlicher synthetischer Datensätze ist es zudem möglich, systematische Auffälligkeiten der auftretenden Differenzen zwischen den Methoden zu entdecken.
Im Rahmen der zunehmenden OpenData-Ansätze in Bund und Kommunen liegen für viele Testgebiete (z. B. Hamburg oder Berlin) ausgewählte Statistiken auf mehreren Bezugsebenen vor. Nach Anwendung aller Methoden kann die externe Genauigkeit durch direkten Vergleich mit der ‘ground truth‘ bestimmt werden.

Im Zuge einer besseren Nutzbarkeit werden die erhobenen (geo-)statistischen Kennwerte dokumentiert und daraus (möglichst parametrisierte) Empfehlungen abgeleitet.

3 Ergebnisse und Ausblick

Dieser Werkstattbericht zeigt, dass die daten- und aufgabenspezifische Methodenempfehlung sowie die Dokumentation und Darstellung der Interpolationsungenauigkeiten dem Nutzer eine transparente Entscheidungsgrundlage zur Bewertung und Weiterverwendung der Ergebnisse bieten. Dadurch ist es möglich, den wachsenden Bestand an Geodaten explorativ zu erforschen um neue Erkenntnisse zu gewinnen.

4 Literaturverzeichnis

FLOWERDEW, R. & GREEN, M., 1994: Areal interpolation and types of data. S. Fotheringham & P.Rogerson (eds.), Spatial Analysis and GIS, London: Taylor & Francis, 73-89.

LI, T., PULLAR, D., CORCORAN, J. & STIMSON, R., 2007: A comparison of spatial disaggregation techniques as applied to population estimation for South East Queensland (SEQ), Australia. Applied GIS, **3**(9).

TOMIO, B., SCHIEWE, J., & TRÖGER, J., 2016: Die Kombination von geographischen Daten verschiedener Bezugsräume mittels Flächeninterpolation – Methoden für Datenjournalisten und Datenjournalistinnen. Kartographische Nachrichten, **66**(3), 123-128.

Crowdsourcing zur Informationsgewinnung von Potentialen zur Nutzung von Gebäudeintegrierter Photovoltaik

TIM WENDT[1], ROBERT HECHT[2] & MARTIN BEHNISCH[2]

Zusammenfassung: Im Hinblick auf die Abschätzung des Potentials für Gebäudeintegrierte Photovoltaik (Building-integrated Photovoltaic, kurz BIPV) spielen insbesondere Fassaden- und Konstruktionseigenschaften eine wichtige Rolle. Ziel dieses Beitrages ist die Erprobung von Crowdsourcing zur Informationsgewinnung über Gebäude und deren Fassadeneigenschaften, insbesondere dem Fenster/Türenflächenanteil, basierend auf gegebenen geokodierten Bilddatensätzen. Es werden kommerzielle und nichtkommerzielle Crowdsourcing-Plattformen getestet und die Qualität der abgeleiteten Information hinsichtlich der Genauigkeit durch Vergleich mit Referenzerhebungen bewertet. Die Ergebnisse zeigen qualitative Unterschiede zwischen den kommerziell und den nicht-kommerziell generierten Daten. Als Gemeinsamkeit beider Plattformen konnte eine systematische Unterschätzung der für BIPV nutzbaren Fläche ermittelt werden.

1 Hintergrund und Fragestellung

In der Bundesrepublik Deutschland ist im Vergleich zu anderen europäischen Ländern das Potential der Gebäudeintegrierten Photovoltaik nur sehr gering ausgenutzt (MONTORO et. al. 2009). Diesen Rückstand gilt es in den kommenden Jahren aufzuholen und den Sektor der Gebäudeintegrierten Photovoltaik zu stärken. Insbesondere fällt der Fokus der Nutzung von Flächenpotentialen zur Integration der Photovoltaikmodule auf Gebäude, die bereits existieren. Neben der primären aktiven Energiegewinnung der Photovoltaikmodule bieten zusätzlich multifunktionale Aspekte gebäudeschützender Funktionen weitere Vorteile, wie bspw. Wetter, Wärme, Brand, Schall- und Strahlenschutz (BUNDESVERBAND BAUSYSTEME e.V. 2012). Trotz wirtschaftlicher und gesellschaftlicher Bedeutung fehlt es an ausreichend detaillierten Informationen zum Gebäudebestand, dessen Struktur und räumlicher Verteilung. Zur Realisierung eines effizienten und flächendeckenden Ausbaus der Gebäudeintegrierten Photovoltaik müssen diese Informationslücken ergänzt und zuverlässige Erfassungsansätze entwickelt werden. Die experimentelle Studie steht im Kontext einer Potentialanalyse eines BMWI finanzierten Projektes zur Entwicklung einer vorgefertigten Standard-BIPV-Fassade für ausgewählte Bauwerkskategorien in Deutschland. In diesem Beitrag werden Crowdsourcingansätze getestet, um Flächenpotentiale unter kombinierter Nutzung von 3D-Gebäudedaten und geokodierter Bilddaten zu ermitteln. Fassadeninformationen sind in einer Vielzahl geokodierter Bilder in sozialen Medien, kommerziellen Diensten und Fotoaustauschplattformen (z.B. FlickR, GoogleStreetView, Mapillary, Wikimapia etc.) implizit gegeben und könnten in Zukunft über eine Crowdcourcing-basierte Bildinterpretation explizit erfasst werden.

[1] Technische Universität Dresden, Institut für Kartographie, Helmholzstraße 10, D-01069 Dresden, E-Mail: wendttim@aol.de
[2] Leibniz-Institut für ökologische Raumentwicklung e.V., Weberplatz 1, D-01217 Dresden, E-Mail: [R.Hecht, M.Behnisch]@ioer.de

Crowdsourcing hat sich bereits in zahlreichen Studien als geeigneter Ansatz bewährt, qualitativ hochwertige Daten kostengünstig zu generieren. Zu nennen wären beispielsweise Studien zur Abschätzung der Landbedeckung (BAYAS et al. 2016; SEE et al. 2013) zur Analyse von Strukturmerkmalen von Gebäuden (HECHT et al. 2017) sowie zur Identifizierung von Malaria betroffenen Blutkörperchen in der Medizin (MAVANDADI et. al 2012). Im Kontext von Crowdsourcing von räumlicher Information spielen neben der Klassifizierung (Annotation, Labeling) von Objekten auch komplexere Aufgaben der Digitalisierung von Polygonen oder das Zusammenführen von Datensätze durch Zuordnung (Conflation) eine wichtige Rolle (ALBUQUERQUE et al. 2016). HILLEN UND HÖFLE (2015) entwickelten beispielsweise den Prototypen Geo-reCAPTCHA zur Digitalisierung von Gebäudeumringen.
In der vorliegenden Studie werden auf einer kommerziellen und einer nicht-kommerziellen Crowdsourcing-Plattform Bilder zur Flächenpotentialabschätzung integriert und bezüglich der Datenqualität validiert, indem die durch die Crowd erzeugten Daten mit den Referenzdaten verglichen werden. Unter Flächenpotentialabschätzung wird im Kontext dieser Arbeit lediglich eine Teilanalyse der für BIPV nutzbaren Fläche verstanden. In Kombination mit 3D-Geodaten können jedoch absolute Flächen berechnet und unter Berücksichtigung der Sonneneinstrahlung echte Solarpotentiale in kWh/m² pro Zeiteinheit abgeschätzt werden. Dieser Beitrag thematisiert folgende Fragestellungen:

- Welche Crowdsourcing-Plattformen gibt es?
- Wie genau lassen sich mithilfe von Crowdsourcing die Flächenpotentiale für Gebäudeintegrierte Photovoltaik abschätzen?
- Weisen die kommerzielle und die nicht-kommerzielle Crowdsourcing-Plattform qualitative Unterschiede auf?

Die aufgeworfenen Fragestellungen werden in diesem Positionspapier in den folgenden Kapiteln beantwortet und wichtige Erkenntnisse dieser Studie im Fazit zusammengefasst.

2 Crowdsourcing

2.1 Definitionsansätze

Crowdsourcing versteht sich als eine Wortneuschöpfung, die sich aus den Begriffen „Crowd“ und „Outsourcing“ zusammensetzt. Dabei werden bestimmte Aufgaben bzw. Wertschöpfungsaktivitäten im Auftrag von Unternehmen oder Institutionen ausgelagert und durch einen öffentlichen Aufruf von einer Masse an Menschen durchgeführt (LEIMEISTER 2012). Crowdsourcing nutzt die „Weisheit der Masse“, um kollektives Wissen zusammenzutragen und gezielt einzusetzen (SUROWIECKI 2004). Erstmals wurde der Begriff Crowdsourcing von Jeff Howe im Jahr 2006 genutzt (HOWE 2006). In diesem Beitrag beziehen wir uns allerdings auf die folgende integrierende Definition des Crowdsourcing-Begriffs (ESTELLÉS-AROLAS & GONZÁLEZ-LADRÓN-DE-GUEVARA 2012): *“Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. […]”*. Crowdsourcing wird im Kontext von Citizen Science auch als erste Stufe der Partizipation angesehen, in dem der Mensch als Sensor agiert (HAKLEY 2013).

2.2 Crowdsourcing-Plattformen

Heutzutage existiert eine große Vielzahl verschiedener Crowdsourcing-Plattformen, auf denen Auftraggeber (Unternehmung, Organisation, Gruppe oder einzelne Personen) eine zu lösende Aufgabe veröffentlichen und auf der anderen Seite potentielle Auftragnehmer (Crowd) diese Aufgabe lösen können. Grundsätzlich kann man die Plattformen nach kommerziellen und nicht-kommerziellen Plattformen unterscheiden. Bei kommerziellen Plattformen werden die Nutzer für abgeschlossene Tasks monetär vergütet und so zusätzliche Anreize für die Aufgabenbearbeitung gegeben. Bei den nicht-kommerziellen Plattformen spielen andere motivierende Faktoren eine Rolle, wie z.B. thematisches Interesse, Unterhaltung, Anerkennung oder der wohltätige Gedanke. Die folgenden Tabellen 1 und 2 geben einen Überblick. Tabelle 1 listet die kommerziellen und Tabelle 2 die nicht-kommerziellen Plattformen mit entsprechenden Anwendungsfeldern sowie der dazugehörigen Homepage auf.

Tab. 1: Zusammenstellung von kommerziellen Crowdsourcing-Plattformen

Plattform	Anwendungsfelder	Homepage
Amazon Mechanical Turk	Diverse	https://www.mturk.com
Casting Words	Übersetzung	https://castingwords.com
Clickworker	Diverse	https://www.clickworker.de
Crowdflower	Datenverifizierung und Datenklassifizierung, Datenkategorisierung und Stimmungsanalysen	https://www.crowdflower.com
Crowdsource (OneSpace)	Schriften	https://www.onespace.com
CrowdSPRING	Logo- und Webdesign	https://www.crowdspring.com
Freelancer	Software, Design, Dateneingabe	https://www.freelancer.com
Guru.com	Diverse	http://www.guru.com
iFreelance	Webentwicklung, Design	https://www.ifreelance.com
myGengo	Übersetzung	https://gengo.com
Replace Myself	Diverse	http://www.replacemyself.com
Squadhelp	Diverse	http://www.squadhelp.com
Textbroker	Schriften	https://www.textbroker.de

Tab. 2: Zusammenstellung von nicht-kommerziellen Crowdsourcing-Plattformen

Plattform	Anwendungsfelder	Homepage
BizReef	Webdesign, Übersetzung	http://www.bizreef.com
Crowdcrafting	Bildklassifizierung, Übersetzung, Geokodierung etc.	https://crowdcrafting.org
Samasource	Digital Work	https://www.samasource.org
Zooniverse	Kunst, Biologie, Klima, Geschichte, Sprache, Literatur, Medizin, Natur, Physik, Sozialwissenschaften	https://www.zooniverse.org

Auf Grundlage der zusammengetragenen Informationen wurden für die experimentelle Studie jeweils eine kommerzielle und eine nicht-kommerzielle Crowdsourcing-Plattform ausgewählt, die in den folgenden Abschnitten detaillierter vorgestellt werden.

2.3 Amazon Mechanical Turk

Amazon Mechanical Turk (kurz MTurk) entspringt der Idee von Jeff Bezos (PONTIN 2007) und existiert seit dem Jahr 2005. Es ist eine der bekanntesten kommerziellen Crowdsourcing-Plattformen. Arbeitgeber können dabei Jobs (sog. Human Intelligence Tasks (HITs)) veröffentlichen und die Rahmenbedingungen nach eigenem Belieben gestalten. Arbeitnehmer können auf der anderen Seite dann bestehende Jobs recherchieren und diese gegen eine vom Arbeitgeber festgelegte Geldleistung durchführen. Der Arbeitgeber kann definieren, wieviel Zeit die Bearbeitung der Aufgabe maximal in Anspruch nehmen darf und wie lange das Projekt aktiv ist. Zudem legt er den Preis fest, der für eine abgeschlossene Task zu zahlen ist (CHEN et al. 2011). Laut BUHRMESTER et al. (2011) bietet MTurk alle notwendigen Elemente zur erfolgreichen Durchführung von Crowdsourcingprojekten. Die Stärken dieser Plattform sind das offene Marktdesign und die vielen Teilnehmer.

2.4 Zooniverse

Im Gegensatz zu den kommerziellen Plattformen wird der Anreiz bei Zooniverse nicht des Geldes wegen geschaffen, sondern durch die Motivation, die durch ein Projekt und bei der Beantwortung der Aufgaben entsteht (FREI 2009). Zooniverse startete im Jahr 2007 mit dem Projekt „Galaxy Zoo“. Dabei beteiligten sich mehr als 175.000 freiwillige Menschen. Die daraus gewonnen Daten wurden für 50 Publikationen verwendet (SMITH et al. 2013). Seitdem hat Zooniverse sich zu einer der wichtigsten nicht-kommerziellen Plattform mit vielen unterschiedlichen Anwendungsgebieten (Biologie, Geschichte, Kunst, Medizin usw.) entwickelt.

3 Untersuchungsansatz

Die folgende Abbildung 1 zeigt eine schematische Darstellung des Untersuchungsansatzes. Dabei ist grundsätzlich zu differenzieren in Aktivitäten des Projektleiters (Vorbereitung und statische Analyse) und der von den Arbeitnehmern/Nutzern (Crowd) durchgeführten Interpretation an sich (Crowdsourcing). In der Vorbereitungsphase werden die Bilddatensätze generiert, das Task Design definiert und implementiert sowie die zur Validierung notwendigen Referenzdaten erhoben. Im Anschluss an das Crowdsourcing erfolgt die statistische Analyse anhand der deskriptiven Statistik und einer externen Validierung.

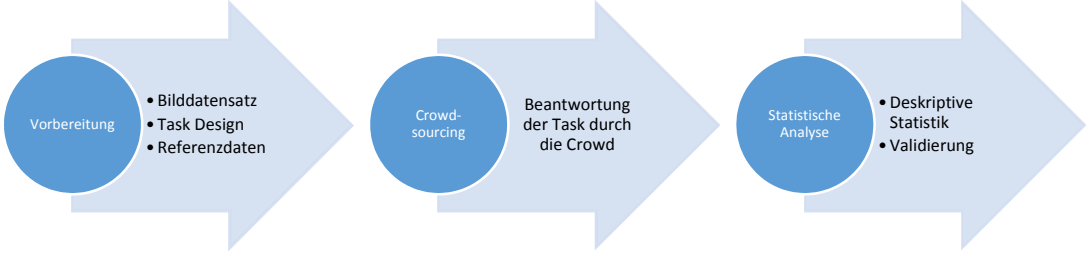


Abb. 1: Workflow der Studie

3.1 Vorbereitung

3.1.1 Bilddatensatz

Grundlage dieser Untersuchung bilden Fotos von Gebäuden. Für die Durchführung wurde ein Bilddatensatz aus dem Projekt „ENOB: dataNWG“ verwendet. Ziel des Projektes ist die Erschaffung einer Nichtwohngebäudedatenbank, „die für valide strukturelle, immobilienwirtschaftliche, energetische und geoinformatische Analysen“ zur Verfügung steht (IWU 2016). Dieser enthielt Fassadenfotos von Gebäuden unterschiedlicher Nutzungskategorien aus den Bundesländern Nordrhein-Westfalen und Thüringen. Die unterschiedlichen Nutzungskategorien können dafür genutzt werden, differenzierte Aussagen über das vorhandene Flächenpotential für BIPV der einzelnen Nutzungskategorien vornehmen zu können. Da pro Gebäude mehrere Bilder aus verschiedenen Blickwinkeln existierten, erfolgte eine manuelle Vorselektion der Bilder. Wichtigstes Kriterien bei der Auswahl des besten Bildes pro Gebäude war eine möglichst vollständige Fassadenabbildung. Aus datenschutzrechtlichen Gründen wurden diese anschließend bearbeitet und alle Kfz-Kennzeichen, Gesichter und Hinweise auf Wohnsitz der Bürger unkenntlich gemacht. Daraufhin konnten die Bilder auf der kommerziellen und der nicht-kommerziellen Crowdsourcing-Plattform integriert werden.

3.1.2 Task Design

Um eine gute Lösbarkeit der definierten Aufgabe (Task) sicherzustellen, wurde eine möglichst einfache Abfrage formuliert, deren Beantwortung unabhängig vom Kulturkreis der Nutzer möglich ist. Zur Abschätzung des prozentualen Anteils der Fassadenfläche für BIPV wurde das Kriterium des Fenster- und Türenflächenanteils verwendet. Realisiert wurde dies über eine Abfrage des Fenster- und Türenflächenanteils über Antwortmöglichkeiten in 10 % Schritten (Auswahlbuttons). Zusätzlich bestand die Möglichkeit im Falle schwieriger Interpretationsverhältnisse keine Bewertung vorzunehmen. Diese Angaben sind nicht mit in die statistische Auswertung eingeflossen. Die Task wurde zusammen mit Handlungsanweisungen und Hilfestellungen auf den Crowdsourcing-Plattformen Amazon Mechanical Turk und Zooniverse identisch umgesetzt (siehe Benutzeroberfläche in Abbildung 3).

3.1.3 Referenzdatenerhebung

Zur Validierung der Daten wurden Referenzwerte mittels Adobe Photoshop erzeugt, indem die Pixelanzahl der Fenster- und Türenfläche gemessen und in Relation zur Gesamtfassade betrachtet wurde. Dieses Verfahren wurde für eine Stichprobe von 25 % des Bilddatensatzes (insgesamt 186 Bilder) durchgeführt. Die Referenzdatenerhebung ist in folgender Abbildung an einem Beispiel mit entsprechender Berechnungsvorschrift aufgezeigt. Den prozentualen Anteil der Fenster- und Türenfläche ist als A_{FT} definiert.

Gebäudepixelanzahl ($P_{Geb.}$) Fassadenpixelanzahl ($P_{Fas.}$)

$$A_{FT} = \frac{PGeb - PFas}{PGeb} \times 100$$

$$A_{FT} = \frac{306490 - 216644}{306490} \times 100$$

$$A_{FT} = 29,31\,\%$$

Abb. 2: Referenzdatenerhebung

Die in diesem Verfahren erstellten Referenzwerte wurden in einer Datenbank zusammengefügt und im Anschluss an das Crowdsourcingverfahren mit den erzeugten Daten verglichen. Es sei darauf hingewiesen, dass geometrische Bildverzerrungen durch Perspektive und Aufnahmesystem bei der Referenzdatenerhebung zu kleinen Abweichungen der gemessenen Referenzwerte gegenüber der Realität führen können, die jedoch im Kontext der Studie vernachlässigbar sind.

3.2 Crowdsourcing

Der Prozess des Crowdsourcing wurde auf einer kommerziellen und einer nicht-kommerziellen Plattform realisiert. Als kommerzielle Plattform wurde Amazon Mechanical Turk (MTurk) und als nicht-kommerzielle Plattform Zooniverse gewählt. Anhand der Abbildung 3 wird der Workflow deutlich, indem Arbeitsanweisungen, Hilfestellungen sowie Antwortmöglichkeiten bereitgestellt werden. Für die korrekte Bearbeitung der Task war es für die Nutzer unerlässlich die Arbeitsanweisung aufmerksam zu lesen sowie die Hilfestellung zu beachten, um bspw. die korrekte Fassadenfläche zu bewerten. Die Benutzeroberflächen unterscheiden sich lediglich grafisch voneinander und sind der Abbildung 3 zu entnehmen.

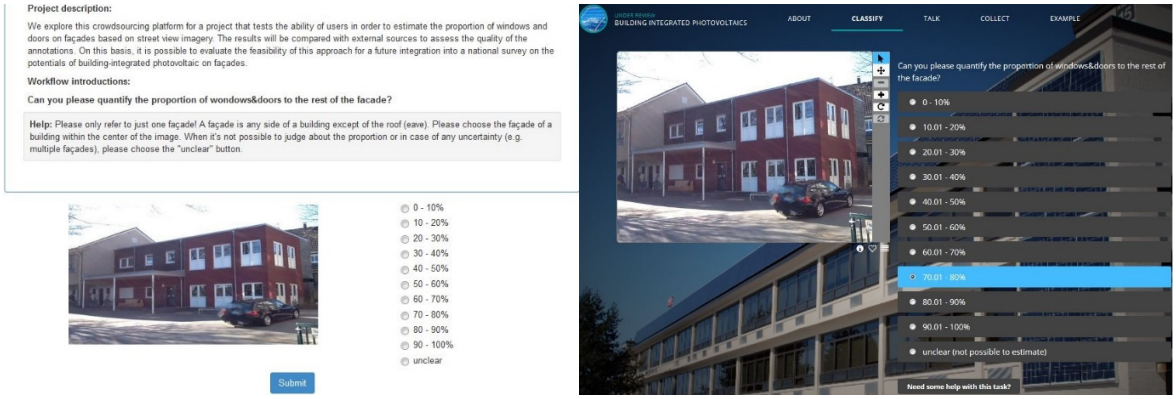


Abb. 3: Implementierte Benutzeroberfläche MTurk (links) und Zooniverse (rechts)

3.3 Deskriptive Statistik

Aus den Ergebnissen der Antworten aus dem Crowdsourcingprozess der jeweiligen Plattform konnten folgende deskriptive Parameter abgeleitet werden.

- Anzahl Bilder
- Annotationen pro Bild (mean)
- Mögliche Kategorien
- Anzahl Annotationen
- Anzahl Annotatoren
- Annotationen pro Annotator (mean)

Diese geben einen ersten Überblick zur Datenstruktur.

3.4 Validierung

Der Validierungsschritt dient dazu, die Qualität des Crowdsourcings in Bezug auf die geschätzten Fenster- und Türenflächenanteile zu messen und zu beurteilen. Hierzu wird der Mittelwert aller Antworten (Annotationen) eines Bildes berechnet und dem jeweils vorliegenden Referenzwert gegenübergestellt. Damit lassen sich Aussagen über die Abweichungen der Ergebnisse aus dem Crowdsourcing vornehmen und ggf. Zusammenhänge feststellen.

4 Erste Ergebnisse und Diskussion

4.1 Deskriptive Statistik

Tabelle 3 zeigt die grundlegenden Kennwerte und statistischen Parameter auf, die sich aus den Daten der Crowdsourcings mittels Amazon Mechanical Turk und Zooniverse ergeben haben. Aus der Tabelle geht hervor, dass die Crowd von Zooniverse die geforderte Anzahl von zehn Annotationen pro Bild in der Projektlaufzeit nicht erreichen konnte. Bei Zooniverse haben zwar deutlich mehr Nutzer (Annotatoren) an dem Projekt teilgenommen, jedoch hat jeder Nutzer im Mittelwert nur 8 Annotationen durchgeführt. Die kommerzielle Plattform von Amazon Mechanical Turk erzielte durch den monetären Anreiz bei den Nutzern eine größere Anzahl von ca. 56 Annotationen pro Annotator.

Tab. 3: Überblick deskriptiver Parameter beider Plattformen

Plattform	Task	Anzahl Bilder	Mögliche Kategorien	Anzahl Annotationen	Anzahl Annotatoren	Annotationen / Bild (mean)	Annotationen / Annotator (mean)
MTurk	Flächenanteil Fenster und Türen in %	743	11	7430	132	10,0	56,3
Zooniverse	Flächenanteil Fenster und Türen in %	743	11	3808	476	5,1	8,0

4.2 Validierung

Die Ergebnisse der externen Validierung zeigt Tabelle 4. Dabei werden die mittleren Werte für den Fenster/Türenflächenanteil von den Nutzern von MTurk mit denen von Zooniverse verglichen. Zudem wird die mittlere Differenz (MEAN Diff) die Standardabweichung der Differenz (SD Diff) dargelegt. Die Standardabweichung gibt die einfache Standardabweichung an und dokumentiert die Streubreite der Werte um das arithmetische Mittel.

Tab. 4: Werte der externen Validierung (Fenster/Türenflächenanteil in %)

	MEAN Referenz (A)	MEAN Wert (B)	MEAN Diff (B-A)	SD Diff
MTurk	27,16	57,36	30,20	13,21
Zooniverse	27,16	41,80	14,64	10,42

Visuell kann ein direkter Vergleich der MTurk-Crowd mit der Zooniverse-Crowd über ein Streudiagramm vorgenommen werden (Abbildung 4). Dabei sind die Mittelwerte der Antwort zum Flächenanteil in % jedes Bildes dem dazugehörigen Referenzwert in % gegenübergestellt. Eingezeichnet ist die Regressionsgerade. Die Anpassungsgüte der Regression wird durch das Bestimmtheitsmaß (R^2) abgebildet. Im Ergebnis ist der Zusammenhang zwischen Referenzwert und abgeschätztem Wert aus dem Crowdsourcing im Falle von Zooniverse stärker.

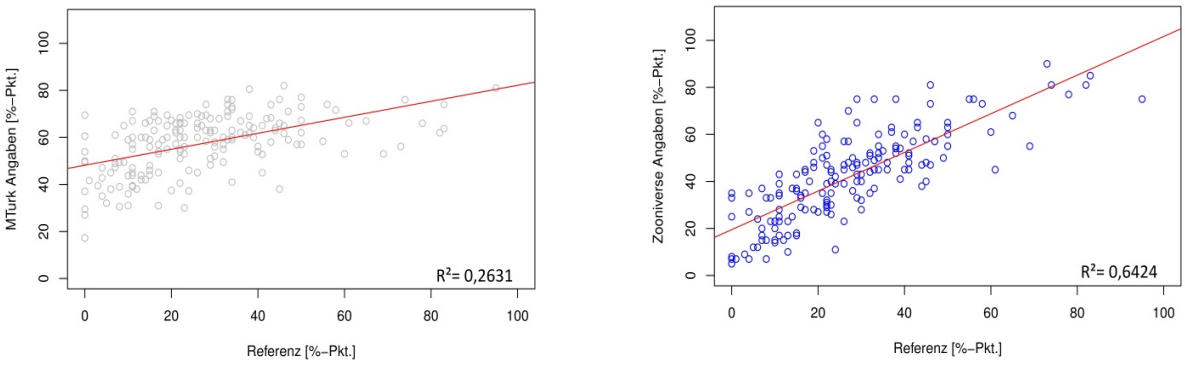


Abb. 4: Gegenüberstellung Crowd Angaben und Referenzwert MTurk (links) und Zooniverse (rechts)

Stellt man auf Grundlage der gegebenen Antworten der Annotatoren und den daraus errechneten Mittelwerten jedes einzelnen Bildes die Stärke der Abweichung in negativer und positiver Ausprägung visuell dar, so kommt die systematische Überschätzung des Fenster- und Türenflächenanteils der Fassaden zur Geltung (Abbildung 5). Allerdings nimmt diese Abweichung kontinuierlich ab, je höher der Fenster- und Türenflächenanteil der Fassaden ist. Diese Besonderheit konnte bei beiden Plattformen beobachtet werden, wobei die Abweichung unter Nutzung von Zooniverse geringer ist. Dieser systematische Zusammenhang stellt eine neue, nutzbare Erkenntnis dar und eignet sich zur Entwicklung eines Korrekturmodells.

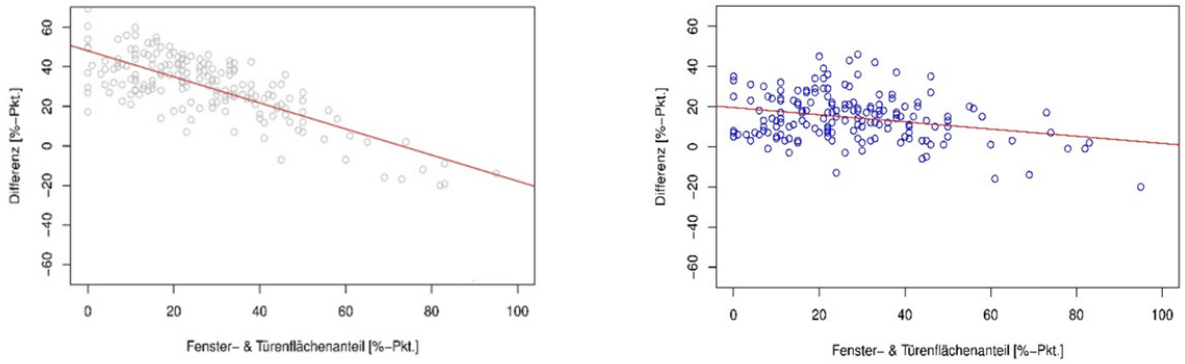


Abb. 5: Abweichung der Crowd - MTurk (links) und Zooniverse (rechts)

Anhand der Konfusionsmatrizen (Tabelle 5 und 6) kann die Stärke der Abweichungen von Amazon Mechanical Turk und Zooniverse gegenüber der Referenz dargestellt werden. Auch hier zeigt sich, dass wesentlich mehr Fehlzuordnungen bei Nutzung der Plattform MTurk im Gegensatz zu Zooniverse auftreten.

Tab. 5: Konfusionsmatrix MTurk

Amazon Mechanical Turk		**Referenz**										
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Summe**
Crowd	**1**	**8**	0	0	0	0	0	0	0	0	0	8
	2	1	**9**	2	0	0	0	0	0	0	0	12
	3	2	4	**3**	0	1	0	0	0	0	0	10
	4	1	5	3	**4**	4	0	0	0	0	0	18
	5	5	4	5	2	**4**	0	0	0	0	0	20
	6	3	2	8	8	2	**0**	1	1	0	0	25
	7	4	4	4	7	1	2	**1**	1	1	0	25
	8	2	10	9	3	6	1	0	**1**	0	0	32
	9	2	2	10	5	5	2	1	0	**2**	1	28
	10	2	2	1	2	0	1	0	0	0	**0**	8
	Summe	28	42	45	32	23	6	3	3	3	1	186

Tab. 6: Konfusionsmatrix Zooniverse

Zooniverse		Referenz										
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Summe**
Crowd	**1**	**16**	3	1	0	0	0	0	0	0	0	20
	2	9	**11**	1	1	1	1	0	0	0	0	24
	3	1	10	**10**	4	1	0	0	0	0	0	26
	4	0	9	12	**3**	1	0	0	0	0	0	25
	5	1	3	8	12	**8**	2	0	0	0	0	34
	6	1	6	5	7	7	**0**	0	0	0	0	26
	7	0	0	4	3	2	1	**1**	0	0	0	11
	8	0	0	4	2	1	2	1	**0**	0	1	11
	9	0	0	0	0	0	0	1	1	**1**	0	2
	10	0	0	0	0	2	0	0	2	2	**0**	7
	Summe	28	42	45	32	23	6	3	3	3	1	186

5 Fazit & Ausblick

Crowdsourcing hält ein großes Potential für die effiziente Annotation von georeferenzierten Bilddaten bereit. In diesem Beitrag wurde die Nutzung von Crowdsourcing für die Abschätzung des Fenster- und Türenflächenanteils unter Nutzung kommerzieller und nicht-kommerzieller Plattformen erprobt. Aus dem Vergleich mit Referenzdaten konnten folgende Erkenntnisse gewonnen werden: (1) Mit der nicht-kommerziellen Plattform Zooniverse konnten bessere Ergebnisse erzielt werden als durch die monetär motivierte Crowd der kommerziellen Plattform Amazon Mechanical Turk. (2) Für beide Plattformen konnte eine systematische Überschätzung des Fenster/Türenflächenanteils beobachtet werden. (3) Mit zunehmenden Fenster/Türenflächenanteil verringert sich die Abweichung des gemessenen Anteils vom Referenzwert.
Schlussfolgernd lässt sich festhalten, dass der Umgang mit der Komplexität der Aufgabenstellung für den Erfolg der Crowdsourcingansätze von großer Bedeutung ist. Im Nachgang sollten vertiefende Auswertungen zu den Ursachen der Abweichungen und möglichen Fehlern im Aufbau des Untersuchungsdesigns durchgeführt werden (z.B. Qualität des Bildmaterials, Didaktik der Aufgabenstellung, unklare Fassadenabgrenzung, Qualifikation der Nutzer). Auch ist zu prüfen, ob eine Potentialabschätzung überhaupt nur für bestimmte Gebäudenutzungskategorien sinnvoll ist. Des Weiteren ist zu klären, ob sich über eine Nachprozessierung mittels Filterung unter Ausschluss der Annotatoren mit einer geringen Vertrauenswürdigkeit die Qualität der Ergebnisse verbessern lässt. Dazu sind entsprechende Gütemaße zur Quantifizierung der Vertrauenswürdigkeit zu berücksichtigen. Zusätzlich könnte auf Basis des systematischen Zusammenhangs von Fenster- und Türenflächenanteil und der Abweichung vom Referenzwert ein entsprechendes Korrekturmodell entwickelt werden.

6 Acknowledgements

Gedankt sei den Kolleginnen und Kollegen des Projektes „Primärdatenerhebung zur Erfassung der Struktur und der energetischen Qualität des Nichtwohngebäudebestands in Deutschland“

(ENOB:dataNWG, Fkz: 03ET1315) mit Förderung durch das Bundesministerium für Wirtschaft und Energie (BMWi) für die Bereitstellung der Bilddaten aus dem Screening (Pilotphase). Ferner danken wir den Kolleginnen und Kollegen des ebenfalls vom BMWi geförderten Projektes „Entwicklung einer vorgefertigten Standard-BIPV-Fassade für ausgewählte Bauwerkskategorien in Deutschland“ (Standard-BIPV, Fkz: 0324063C) für die Unterstützung bei der Umsetzung sowie die wertvollen Hinweise bei der Entwicklung des Task Designs.

7 Literaturverzeichnis

ALBUQUERQUE, J.P., HERFORT, B. & ECKLE, M., 2016: The Tasks of the Crowd: A Typology of Tasks in Geographic Information Crowdsourcing and a Case Study in Humanitarian Mapping. Remote Sensing, **8**, 859.

BAYAS, J.C.L., SEE, L., FRITZ, S., STURN, T., PERGER, C., DÜRAUER, M., KARNER, M., MOORTHY, I., SCHAPESCHENKO, D., DOMIAN, D. & MCCALLUM, I., 2016: Crowdsourcing In-Situ Data on Land Cover and Land Use Using Gamification and Mobile Technology. Remote Sensing, **905**(8), 1-18.

BUHRMESTER, M., KWANG, T. & GOSLING, S. D., 2011: Amazon’s Mechanical Turk: A New Source of inexpensive, Yet High-Quality, Data? Perspectives on Psychological Science, **6**(1), 3-5.

BUNDESVERBAND BAUSYSTEME e.V. (Hrsg.), 2012: Bauwerksintegrierte Photovoltaik-Systeme – Positionspapier. Koblenz.

CHEN, J.J., MENEZES, N.J. & BRADLEY, A.D., 2011: Opportunities for Crowdsourcing Research on Amazon Mechanikal Turk. Seattle.

FREI, B., 2009: Paid Crowdsourcing – Current State & Progress toward Mainstream Business Use.

HAKLAY, M., 2013: Citizen science and volunteered geographic information: Overview and typology of participation. Crowdsourcing Geographic Knowledge, Sui, D., Elwood, S. & Goodchild, M., (Hrsg.), Springer Netherlands, Dordrecht, 105-122.

HECHT, R., KALLA, M. & KRÜGER, T., 2017: Crowd-sourced data collection to support automatic classification of building footprint data. ICC 2017: Proceedings of the 2017 International Cartographic Conference.

HILLEN, F. & HÖFLE, B., 2015: Geo-reCAPTCHA: Crowdsourcing large amounts of geographic information from earth observation data. International Journal of Applied Earth Observation and Geoinformation, **40**, 29-38.

HOWE, J., 2006: The rise of crowdsourcing. Wired Mag 14, 1–4.

IWU, 2016: Forschungsdatenbank Nichtwohngebäude (ENOB: dataNWG).

LEIMEISTER, J. M., 2012: Crowdsourcing: Crowdfunding, Crowdvoting, Crowdcreation. Zeitschrift für Controlling und Management (ZFCM), **56**, 388-392.

MAVANDADI, S., DIMITROV, S., FENG, S., YU, F., SIKORA, U., YAGLIDERE, O., PADMANABHAN, S.; NIELSEN, K. & OZCAN, A., 2012: Distributed Image Analysis and Diagnosis through Crowd-Sourced Games: A Malaria Case Study, PLoSOne **7**(5)

MONTORO, D. F., VANBUGGENHOUT, P. & CIESIELSKA, J., 2009: Building Integrated Photovoltaics: An overview of the existing products and their fields of application. Sunrise (Hrsg.)

PONTIN, J., 2007: Artifical Intelligence, With Help From the Humans. <http://www.nytimes.com/2007/03/25/business/yourmoney/25Stream.html>.

SEE, L., COMBER, A., SALK, C., FRITZ, S., VAN DER VELDE, M., PERGER, C., SCHILL, C., MCCALLUM, I., KRAXNER, F. & OBERSTEINER, M., 2013: Comparing the Quality of Crowdsourced Data Contributed by Expert and Non-Experts, PLoSOne **8**(7)

SMITH, A. M., LYNN, S. & LINTOTT, C. J., 2013: An Introduction to the Zooniverse. AAAI Technical Report Cr-13-01. 103

SUROWIECKI, J., 2004: The Wisdom of Crowds.

Employing OGC's 3D Portrayal Service to Interoperate Hierarchical Data Structures: A Case Study on Visualizing I3S in Cesium

ATHANASIOS KOUKOFIKIS[1] & VOLKER COORS[1]

Abstract: The demand of serving large 3D spatial data, mainly in urban areas, reflected the need of hierarchical 3D data structures. During the last years the OGC community standard I3S (ESRI) and 3D Tiles (Analytical Graphics), emerged in order to deal with this issue. Conceptually, hierarchical 3D structures operate in an analogous manner with web map tiles, differentiating only in the implementation. A prototype implementation focuses on the rendering of I3S in the Cesium client using the 3D Portrayal Service. As a result, the user can query a scene via the 3D Portrayal Service by specifying a spatial region, rather than a specific resource via a URI. The result can be delivered either using I3S or 3D Tiles as a data delivery format, depending on which data is available for the specified region. The Cesium client can both render both the I3S as well as the 3D Tiles content.

1 Introduction

The technological pluralism in 3D web visualization constantly increases the need of interoperability. The availability of a number of different 3D delivery formats and client consumers indicates a need for a generic approach for accessing 3D information. This paper defines an attempt to realize the interoperability between geospatial data web consumers (web globes) and OGC's 3D Portrayal Service regarding the request of hierarchical 3D data storages. 3D Tiles and I3S are the latest implementations of hierarchical data storages. Our goal is to investigate if the 3D Portrayal Service could abstract the access of 3D Tiles or I3S in a dedicated web client/consumer which is designed to be compatible with specific delivery formats (e.g. Cesium, 3D Tiles). Cesium is a web globe API designed to support 3D Tiles, whereas ESRI's ArcGIS API for JavaScript is designed to consume I3S. Our approach attempts to render I3S data in Cesium employing a request scheme that extends 3D portrayal's getScene request.

2 Related Work

Use cases of the 3D Portrayal Service appear to be scarce, since the version 1.0 of the standard was recently released. GAILLARD et al. (2015) utilized the 3D Portrayal Service View conformance class for client side rendering of tiled 3D city models. Building geometry, originating form CityGML, was converted and served as JSON using a GetScene request. GUTBELL et al. (2016) implemented a server side rendering framework to visualize 3D city models using the 3D Portrayal Service GetView request.

[1] Hochschule für Technik Stuttgart, Faculty Geomatics, Computer Science and Mathematics, Schellingstr. 24, D-70174 Stuttgart, E-Mail: [athanasios.koukofikis, volker.coors]@hft-stuttgart.de

Cesium as visualization component is used in several cases. KRÄMER & GUTBELL (2015) produced a surface model visualization using Cesium's Terrain Builder. KYOUNG-SOOK et al. (2017) presented a visual extension of Cesium to visualize moving objects in a space-time cube. LU et al. (2016) implemented an online 3D city model interactive editor using a light-weight viewer based on Cesium.
So far, use cases of Cesium utilize data formats which are compatible with this API.

3 Interoperable visualization using the 3D Portrayal Service (3DPS)

The 3D Portrayal Service is a OGC Standard that abstracts the access of 3D geospatial datasets in various client platforms via the web for visualization purposes (3D PORTRAYAL SERVICE 1.0 2017). The 3D Portrayal Service specifies three methods to access information: GetCapabilities, AbstractGetPortrayal and AbstractGetFeatureInfo. A GetCapabilities request returns information about the available request methods, the extents of the data, data layers (buildings, vegetation etc.), layer styles and streaming formats supported. The retrieval of a scene is implemented by the AbstractGetPortrayal operator. The GetScene method is used for client side rendering and the GetView method for server side rendering. The AbstractGetFeatureInfo implementations are used for requesting metadata of scene features. The supported methods are GetFeatureInfoByObjectId, GetFeatureInfoByPosition and GetFeatureInfoByRay. The 3D Portrayal Service specification does not indicate a content delivery format. The OGC community standard I3S can be used as content delivery for the GetScene request. I3S (Indexed 3d Scene Layer) is a technology for rapidly streaming and distributing large volumes of 3D content across enterprise systems that may consist of server components, cloud hosted components, and desktop, web and mobile clients (I3S-SPEC 2017). The specification is realized by various profiles that describe the behavior of a I3S layer. The spatial extent of the I3S data is split into regions, called nodes, and organized into a hierarchical data structure that allows the client to quickly discover which data is needed and the server to quickly locate the data requested by the client (INDEXED 3D SCENE LAYER FORMAT SPECIFICATION 2017).

4 Methodology

An agile software development approach was used in this study to implement both 3DPS server and web-based client. It follows an iterative incremental process with 4 prototypes in total. Each iteration encapsulates a prototypical objective which contributes to the goal of this case study. In order to prove interoperability, we selected a hierarchical 3D format, i.e., I3S, with a dedicated consumer on the client side and Cesium as an incompatible consumer.
The aim of the first prototype was a preliminary attempt to visualize elements with a geographic component coming from I3S in Cesium. For this, we used the minimum bounding spheres (MBS) of I3S nodes.
The aim of the second prototype was the implementation of the 3D Portrayal Service and visualization of building shells originating from I3S. During this step, it was critical to decide the role of the 3D Portrayal Service in the end to end communication. An important action is the traversal of the I3S node tree, which in the case of an ArcGIS API application is happening on the client.

A Cesium application accessing the I3S REST API for the node traversal, in order to identify the visible nodes in the view frustum, would result to a passive/obsolete 3D Portrayal Service implementation. For this reason, the 3DPS received an extra parameter, i.e., "i3sLayer" and was responsible to access and traverse the node tree of a I3S layer.
The third prototype introduced the "boundingbox" parameter from the GetScene request and replace the "i3sLayer" which was used until this point to identify the requested resource. Even though the access of 3D resources was abstracted after the "i3sLayer" parameter was removed, the "boundingbox" parameter was still a static implementation.
The aim of the fourth prototype was to define a dynamic bounding box selection in the server (3DPS) and client side in order to allow the users to fetch the data they are interested. In this case the "boundingbox" parameter is used to identify which hierarchical dataset is requested and also restrict spatially the tiles/nodes to be rendered.

5 Implementation

The 3D Portrayal Service is implemented utilizing the Play Framework. At a later stage of the application the Dojo Toolkit is introduced in the client side.

5.1 Play Framework

Play is a high-productivity Java and Scala web application framework that integrates the components and APIs needed for modern web application development. Play is based on a lightweight, stateless, web-friendly architecture for highly-scalable applications (PLAY 2.6.X DOCUMENTATION 2017). The source code is compiled into bytecode and executed in the Java Virtual Machine.

5.2 Dojo Toolkit

Dojo Toolkit is an open source JavaScript library designed to ease the rapid development of cross-platform, web applications and web sites (DOJO TOOLKIT 1.13 2017). This tool was used mainly to modularize the client side application.

5.3 First Prototype: Visualization of the minimum bounding spheres of I3S nodes

The first attempt visualizes spatial elements retrieved by a I3S REST API. For this purpose, the minimum bounding sphere (MBS) for each I3S node that lies within a cartographic bounding box is visualized using Cesium's Ellipsoid primitive. The MBS returned by the I3S REST API contains an array with the geographical position of the sphere (longitude, latitude, height) and its radius in meters (Fig. 1).

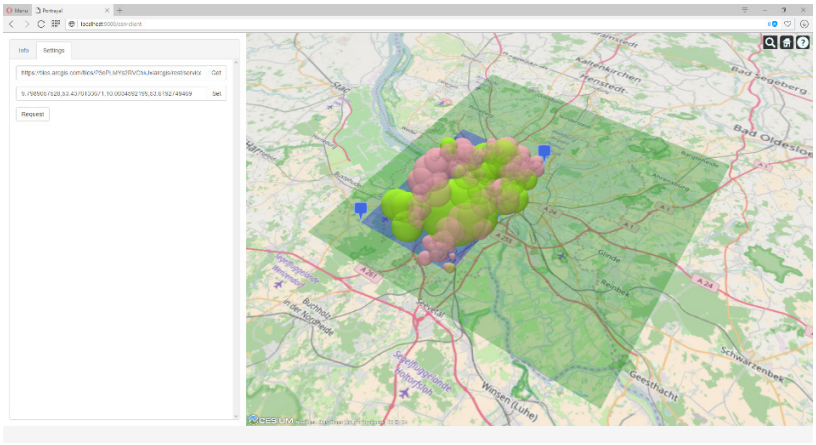


Fig. 1: Visualization of the minimum bounding spheres of I3S nodes within a user defined bounding box. Each sphere is styled by the node level. The green box defines the extents of the I3S layer

5.4 Second Prototype: 3D Portrayal Implementation

In this stage a middleware service (3DPS) is introduced using the Play Framework. This service, which plays the role of the 3D Portrayal Service, acts as a broker that negotiates the I3S nodes that should be rendered in the Cesium viewer (Fig. 2). The "boundingbox" URL parameter of the GetScene request is not present in the server endpoint, yet the I3S layer URL is passed directly as a request parameter to identify the 3D dataset that should be queried.

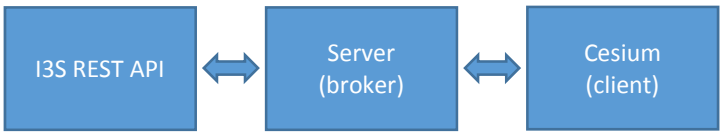


Fig. 2: Rendering I3S in Cesium overview

On the client side, a request is sent to the server when the "camera changed" event is triggered in Cesium. On the server side (broker) the implemented API is responsible to handle any requests from the client. The core action of the broker service is to apply the I3S node selection criteria, then generate a response with the nodes' description that should be rendered in Cesium. Once the broker's response is delivered, the client fetches the I3S payloads using the REST API and prepares the visualization (Fig. 3).

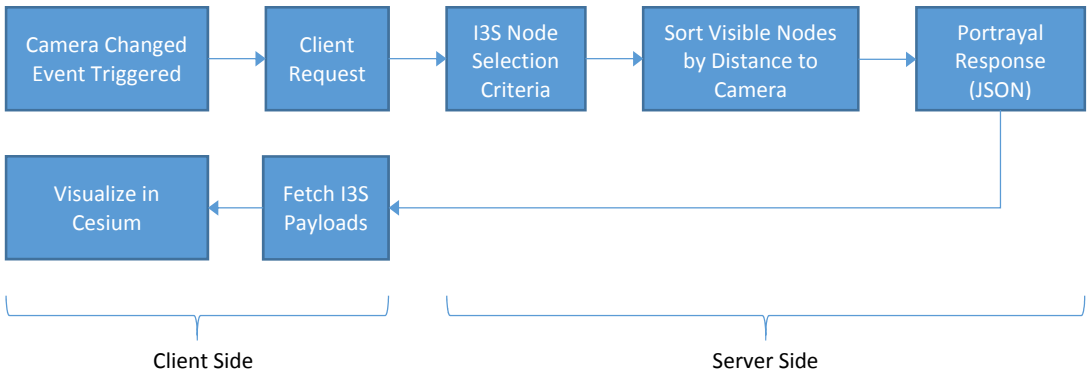


Fig. 3: Communication Synopsis

The I3S node selection is a repetitive process which accesses the root node of the I3S layer and continues by traversing descendant nodes of the tree structure utilizing the breadth-first traversal method. The main node selection criteria are: a) The node's minimum bounding sphere (MBS) visibility b) The node's MBS screen size in pixels and c) The existence of a node's children. The last iteration introduces the intersection of the MBS center with a buffered cartographic bounding box, defined by the user, as an additional criterion. The following pseudo code describes the traversal of an I3S node inside the broker service:

```
if node's mbs in not visible in client's viewport
      break traversal of the node and it's children
else if node's screen size >= node's maxScreenThreshold and node has children
      traverse node's children
else
      add node in the response for rendering
```

When the I3S node traversal is completed, a response is generated which contains the essential information to render a node in Cesium. For rendering optimization, the response nodes are sorted with ascending order by distance to the camera (Fig. 4).


```
[
  + {…},
  - {
        id: "5-1-0-0-0",
        level: 6,
        lng: 9.994545253174223,
        lat: 53.568095227256954,
        radius: 455.4209899902344,
        url: "https://tiles.arcgis.com/tiles/P3ePLMYs2RVChkJx/arcgis/rest/services/Buildings_Hamburg/SceneServer/layers/0/nodes/5-1-0-0-0",
        distanceToCamera: 140.98541019527966,
        time: "1504872740063"
    },
  - {
        id: "5-0-0-2-0",
        level: 6,
        lng: 9.986759068536088,
        lat: 53.57170940098039,
        radius: 563.823974609375,
        url: "https://tiles.arcgis.com/tiles/P3ePLMYs2RVChkJx/arcgis/rest/services/Buildings_Hamburg/SceneServer/layers/0/nodes/5-0-0-2-0",
        distanceToCamera: 162.51702725550808,
        time: "1504872740063"
    },
  - {
        id: "5-1-0-0-1",
        level: 6,
        lng: 9.997446293904332,
        lat: 53.565953434999045,
        radius: 546.78515625,
        url: "https://tiles.arcgis.com/tiles/P3ePLMYs2RVChkJx/arcgis/rest/services/Buildings_Hamburg/SceneServer/layers/0/nodes/5-1-0-0-1",
        distanceToCamera: 214.9592984261286,
        time: "1504872740063"
    },
  + {…},
  + {…},
```


Fig. 4: Broker service response

I3S supports discrete levels of detail (LoD) in the geometry which correspond to the levels in the nodes' hierarchy. Leaf nodes usually contain the original representation with the highest detail, where lower node levels contain simplified geometry of the same features using edge collapse algorithms (Fig. 5).

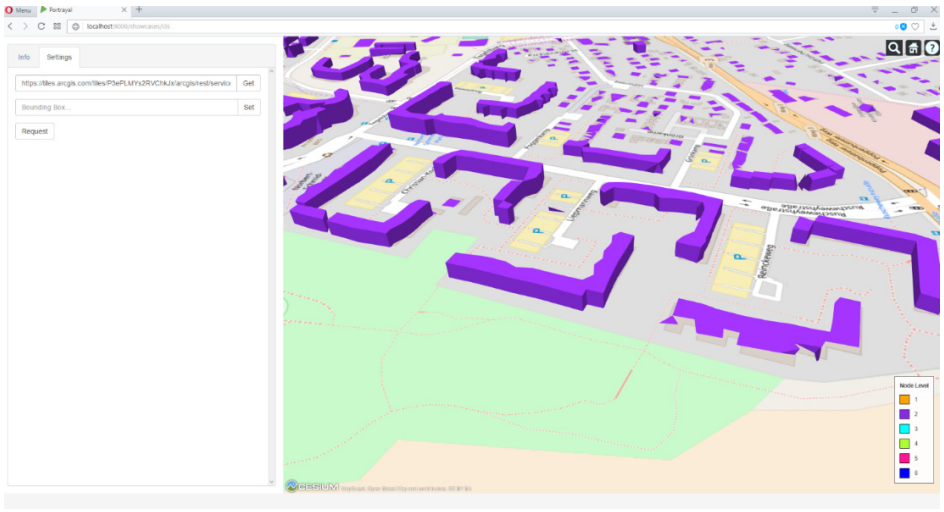


Fig. 5: Decimated geometry of I3S low level nodes rendered in Cesium

5.5 Third Prototype: The “boundingbox” parameter

In this stage of the application the I3S layer URL parameter gets replaced with a static “boundingbox” parameter. Although the user is able to select a bounding box which is available in the 3D Portrayal Service capabilities and corresponds to the extents of various 3D city models, it is not possible to dynamically define its own bounding box (Fig. 6). On the server side a lookup array is used to match the requested bounding box with the I3S layer URL.

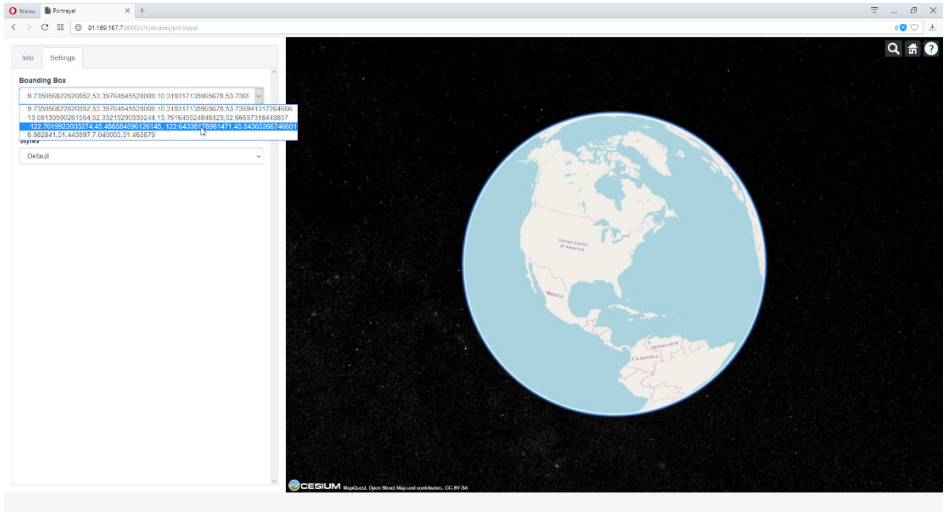


Fig. 6: The user is able to select a bounding box the defines the extents of a 3D city model stored as I3S

5.6 Fourth Prototype: Dynamic bounding box selection

In the code base of the application the Dojo Toolkit is introduced in order to modularize the client application. Even though the client appears the same, many parts of the client were rewritten. The user is able to dynamically define a cartographic bounding box which is spatially intersected by the 3D Portrayal Service available extents in order to identify the requested resource, either

3D Tiles or I3S. Additionally, a cartographic buffer is generated around the requested bounding box to include neighboring I3S nodes (Fig. 7).

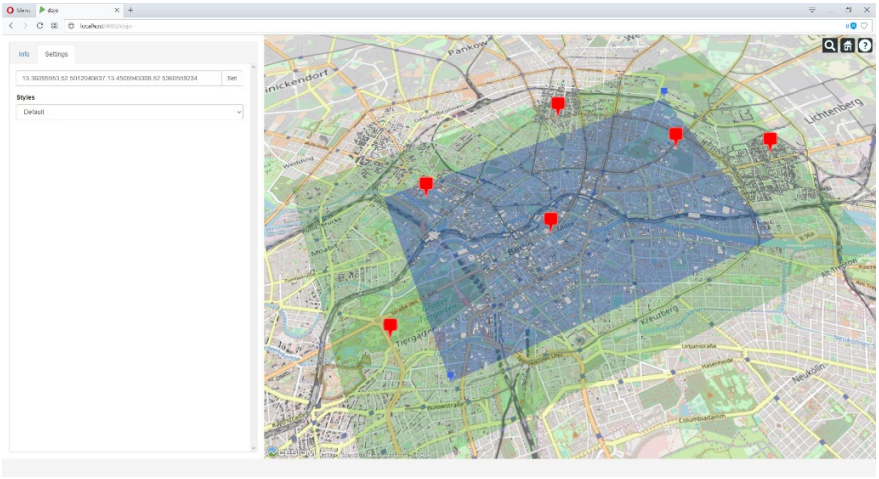


Fig. 7: I3S nodes rendered in Cesium. Blue rectangle: requested bounding box. Green rectangle: buffer. Red pins: centers of the minimum bounding spheres

The 3D Portrayal request is extended with URL parameters which are not defined in the request scheme of the standard. These parameters, originating from Cesium, are used to apply the I3S node selection criteria and to reject outdated responses by the 3D Portrayal Service (Fig. 8).

service: 3DPS
acceptversions: 1.0
request: GetScene
boundingbox: 9.7350568228,53.397645455,10.3193171359,53.7369413172
cullingvolume: -0.489520008514,0.87030950429,-0.054143586849,1533027.01880,-0.161886420120,-0.9853188378,-0.054217845856,1530930.62031,-0.62377256088,-0.110165543187,0.77380316966,-1550108.68011,-0.3480433034,-0.061413811752,-0.9354646987,6121145.3019,-0.97181586470,-0.171579355008,-0.161661529262,4571035.6240,0.97181586470,0.171579355008,0.161661529262,495428963.37
camera: 3737336.29540,658537.61433,5109703.3728,-0.97181586470,-0.171579355008,-0.161661529262
frustum: 1,0.57735026918,0.35572226262
drawingbuffer: 573,930
time: 1516234129317

Fig. 8: Extended 3D Portrayal request with additional parameters highlighted in blue

6 Results & Summary

The 3D Portrayal Service uses a bounding box oriented approach to deliver 3D content to clients while hierarchical 3D datasets are suitable for viewport oriented applications. In this paper, the 3DPS was used as a broker to simplify data access (Fig. 9).

Fig. 9: I3S layer rendered in Cesium via the 3D Portrayal Service. On the left side the default style is used, on the right side the node level is used for styling

It basically translated user bounding box queries to a URI of the relevant 3D data set in I3S. Of course, 3D Tiles data sets can be handled the same way, but the focus was on rendering I3S in Cesium. The cost of using the 3DPS as a broker intermediary is O(log n) if a spatial index is used on server side for searching the relevant data set.

7 Acknowledgements

We would like to thank Tamrat Belayneh for his invaluable experience and help.

8 References

COORS, V., HAGEDORN, B., THUM, S., REITZ, T. & GUTBELL, R., 2017: 3D PORTRAYAL SERVICE 1.0. (15-001r4).

DOJO TOOLKIT 1.13, 2017: dojotoolkit.org. Retrieved 14 January 2018, from https://dojotoolkit.org

GAILLARD, J., VIENNE, A., BAUME, R., PEDRINIS, F., PEYTAVIE, A. & GESQUIÈRE, G., 2015: Urban data visualisation in a web browser. Proceedings of the 20th International Conference on 3D Web Technology - Web3d '15. http://dx.doi.org/10.1145/2775292.2775302

GUTBELL, R., PANDIKOW, L., COORS, V. & KAMMEYER, Y., 2016: A framework for server side rendering using OGC's 3D portrayal service. Proceedings of the 21st International Conference on Web3d Technology - Web3d '16. http://dx.doi.org/10.1145/2945292.2945306

I3S-SPEC, 2017: GitHub. Retrieved 12 January 2018, from https://github.com/Esri/i3s-spec

INDEXED 3D SCENE LAYER FORMAT SPECIFICATION, 2017: GitHub. Retrieved 12 January 2018, from https://github.com/Esri/i3s-spec/blob/master/format/Indexed%203d%20Scene%20Layer%20Format%20Specification.md

KRÄMER, M. & GUTBELL, R., 2015: A case study on 3D geospatial applications in the web using state-of-the-art WebGL frameworks. Proceedings of the 20th International Conference on 3D Web Technology - Web3d '15. http://dx.doi.org/10.1145/2775292.2775303

KYOUNG-SOOK, K., DONGMIN, K., HYEMI, J. & HIROTAKA, O., 2017: Stinuum: A Holistic Visual Analysis of Moving Objects with Open Source Software. Proceedings of the 25th ACM

SIGSPATIAL International Conference on Advances in Geographic Information Systems, 1-4.

LU, Z., GUERRERO, P., MITRA, N., & STEED, A., 2016: Open3D: crowd-sourced distributed curation of city models. Proceedings of the 21st International Conference on Web3d Technology - Web3d '16. http://dx.doi.org/10.1145/2945292.2945302

PLAY 2.6.X DOCUMENTATION, 2017: Playframework.com. Retrieved 12 January 2018, from https://www.playframework.com/documentation/2.6.x/Home

Multi-layer Land Cover Data for Remote-Sensing based Vegetation Modelling for South Korea

CHRISTINA EISFELDER[1], HYUN OK KIM[2] & CLAUDIA KUENZER[1]

Abstract: Land cover data is essential input for vegetation productivity models that are often driven by coarse resolution data. In this study, we analyze how well 1 km land cover data represent land cover at 30 m for South Korea. We derive multi-layer 1 km land cover classes and coverages and analyze how much of land cover heterogeneity is represented by the successive layers. Comparison to global land cover data shows varying agreement. The multi-layer land cover data can be used for example for net primary productivity modelling. Especially, for models that can include more than one vegetation type per pixel, multi-layer land cover data and their corresponding coverages are a major asset.

1 Introduction

Land cover data are important base information for modelling vegetation productivity. They provide information about the distribution of different land cover and vegetation types on the land surface. This knowledge is essential for modelling net primary productivity (NPP), which is the net accumulation of organic matter through photosynthesis by green vegetation per unit area and unit time. NPP is one of the most important components of the carbon cycle, a key variable for ecological monitoring, and a sensitive indicator of climate and environmental change (NIEMEIJER 2002; SCHIMEL 1995).

NPP modelling is commonly based on phenological and meteorological time-series data, which can be derived from remote sensing (PRIETO-BLANCO et al. 2009; YI et al. 2013; ZHAO et al. 2005). One important input dataset for NPP modelling is land cover classification. It is needed to define the type of vegetation to be modelled with its specific parameters for a certain location. Most often, global land cover datasets are employed for NPP modelling (e.g. MATSUSHITA & TAMURA 2002; NIKLAUS et al. 2015; TUM et al. 2016; WIßKIRCHEN et al. 2013), but also regional land cover maps can be used (e.g. BAO et al. 2016; EISFELDER et al. 2014; TUM et al. 2012). The availability of regional land cover maps may be of major advantage, as global land cover data have shown weakness in describing heterogeneity in land cover characteristics on regional scale (e.g. GESSNER et al. 2015; KLEIN et al. 2012; LEINENKUGEL et al. 2014).

Land cover maps commonly define one land cover class per pixel. However, some models, such as the Biosphere Energy Transfer Hydrology model BETHY/DLR (WIßKIRCHEN et al. 2013), are able to include more than one vegetation type per pixel for NPP modelling. In previous studies, each land cover class has been translated to two fixed vegetation types, each weighted by a fixed percentage of coverage (e.g. EISFELDER et al. 2017).

[1] German Aerospace Center (DLR), Earth Observation Center (EOC), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, 82234 Weßling, Germany. E-Mail: Christina.Eisfelder@dlr.de
[2] Korea Aerospace Research Institute (KARI), 169-84 Gwahak-ro, Yuseong-gu, Daejeon, 34133 Korea

The availability of land cover products of higher spatial resolution than required for NPP modelling (e.g. SIMIC et al. 2004), allows to derive a set of land cover information at coarse resolution that provides information about both several land cover classes per pixel and their coverage.
In this study, we derive such multi-layer coarse resolution land cover information for South Korea. We first describe the method applied for deriving multi-layer land cover information including classes and coverages based on a high resolution land cover map of South Korea. We present the results and analyse the information content contained in the successive land cover layers. We also compare the land cover classification to three global land cover datasets.
In this study, we address the following research questions: (a) How well is a high resolution land cover classification represented by a derivative coarse resolution land cover classification? (b) How does this compare to global land cover classifications? (c) How much additional information is contained in successive multi-layer land cover data sets? (d) What can be concluded regarding recommendation for different land cover input data for NPP modelling?

2 Data base and processing of land cover data

2.1 High Spatial Resolution Land Cover Data

For the presented study, the Korean national land cover map issued by the Ministry of Environment was used. The national land cover map is organized with three hierarchical levels of broad (7 classes), middle (22 classes) and detailed classification (41 classes) according to their resolution. In this study, the land cover map at the broad class level, which covers the whole South-Korean territory, was used. This land cover map distinguishes seven land cover classes (forest, urban, agriculture, wetland, grassland, bare soil, and water) and was lastly updated in 2010 based on Landsat-7 data with 30 m spatial resolution (period of image recording 2008–2010). For the Landsat image processing, after the atmospheric correction the coordinate system was projected from the WGS84 UTM to the GRS80 TM (national standard), then the geospatial errors were corrected by in-situ GPS measurements and the national topographical maps (1:25,000). The mosaicked image was calibrated, normalized using the digital elevation model (DEM) to remove the topographic effects of the spectral reflections in mountainous areas. For the classification a hybrid method combining unsupervised and supervised classification algorithm was applied. These classification results were compared with the very high resolution (pan: 1 m, multispectral: 4 m) KOMPSAT-2 (Korea Multipurpose Satellite with optical sensor) images and systematic errors were corrected. The Ministry of Environment ensures an overall accuracy of higher than 75% over the whole nation in case of validation comparing with KOMPSAT-2 images.

2.2 Processing of Land Cover Input Data

For preparation as input for NPP modelling, the high spatial resolution land cover data (30 m) were rescaled to a resolution of 0.008929° (“1 km”) in order to match the resolution of the commonly used LAI input data for BETHY/DLR. The example model is designed to include two land cover types within each pixel. Weight factors (range [0, 1]) define the fractional coverage of the land cover types. In order to keep a maximum of valuable land cover information for NPP modelling, thus three land cover datasets providing information on the primary (largest coverage), secondary (second largest coverage) and tertiary land cover (third largest coverage) within

each 1 km² pixel are required. Each of the land cover dataset layers is to be accompanied by the percentage cover of the land cover class within the respective 1 km² pixel. Three layers can be made use of – one more than land cover types modelled: two land cover types plus one additional layer that is included in case one of the first two land cover classes is not vegetated

An overview on the workflow for deriving the multi-layer land cover information is shown in figure 1. Land cover classes that are not vegetated (water, bare soil, urban) and for which, thus, no NPP is to be calculated were merged into one new land cover class "other". The resulting datasets contains the following classes: agriculture, forest, grassland, wetland, and other.

Finally, based on the Korean national land cover map, this procedure resulted in a total of five successive land cover datasets – each with its supplementary percentage cover dataset. The first three layers can be used as input for BETHY/DLR. In case that the previous land cover class/classes already cover 100% of the 1 km² pixel, the dummy class "no data" was introduced.

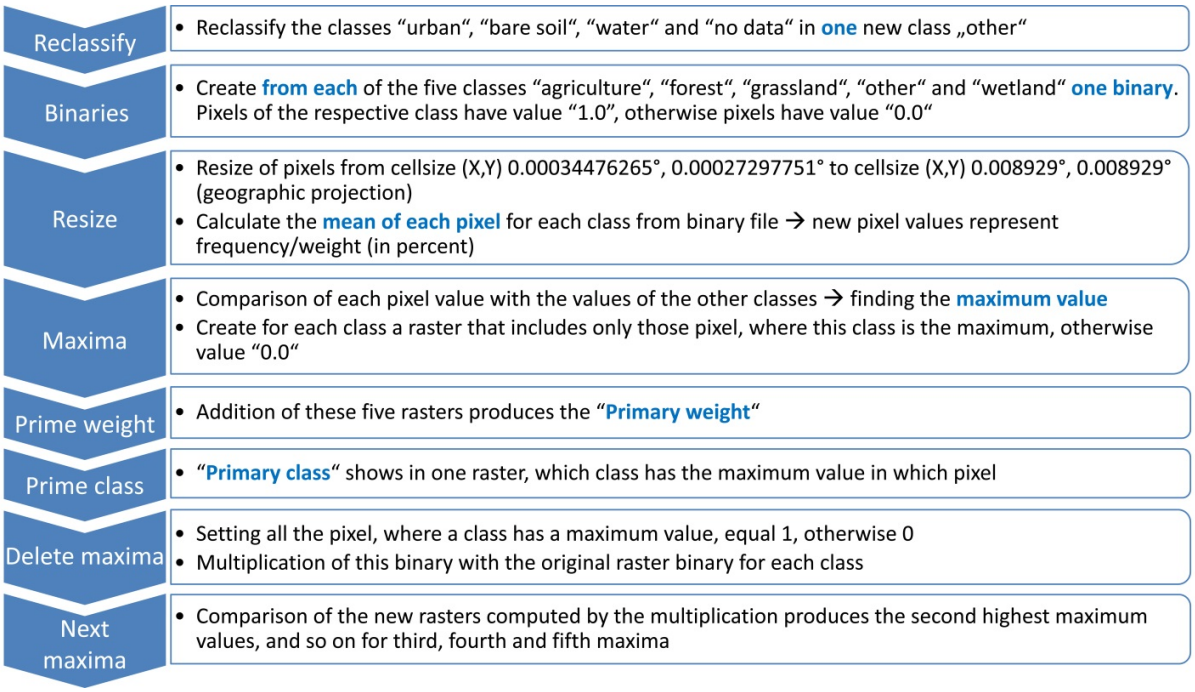


Fig. 1: Workflow for deriving coarse spatial resolution multiple land cover information and weights from high spatial resolution land cover data.

3 Results and Discussion

3.1 Multi-layer Land Cover Information

Figure 2 shows the first three land cover information data sets derived with the procedure described in section 3. The maps on the left show the land cover class assigned for the primary, secondary, and third land cover class. The maps on the right show the corresponding weight for the successive land cover classes, i.e. information about the coverage of the land cover class within the 1 km² pixel.

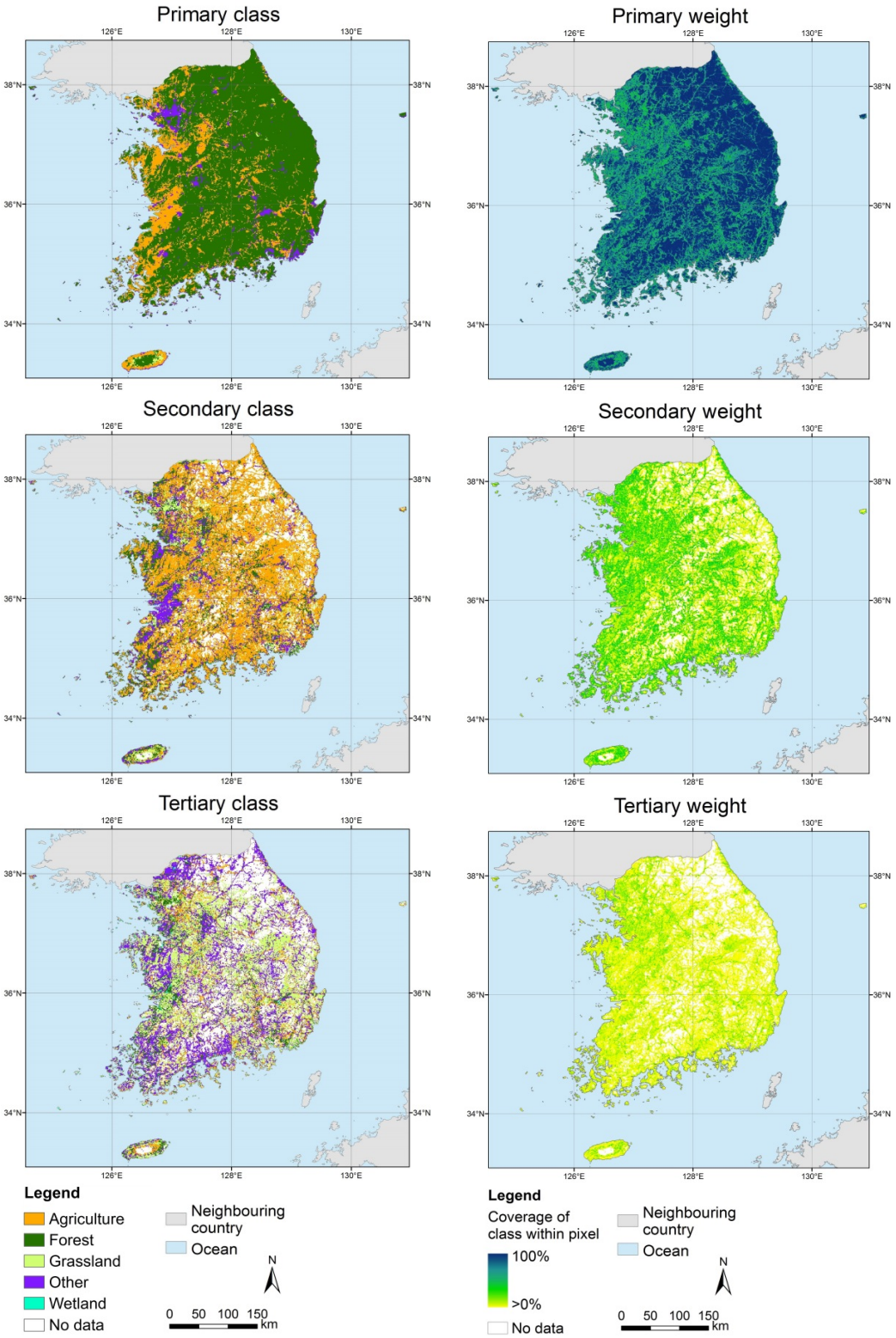


Fig 2: Left: Maps showing the primary, secondary, and tertiary (top to bottom) land cover class of each pixel for South Korea. Right: Maps showing the percentage coverage of the primary, secondary, and tertiary (top to bottom) land cover class within each pixel.

As can be seen from figure 2 (top left), the major primary land cover class within South Korea is forest, followed by agriculture. Coverage of these classes within the 1 km² pixels is up to 100% (see figure 2, top right). The most important secondary land cover class is agriculture, which spreads over large parts of South Korea but shows – in these areas – a pixel coverage of less than 50 % (figure 2, middle). The third most frequent land cover classes are "grassland" and "other"; the corresponding coverage is relatively low (figure 2, bottom).
Figure 3 shows diagrams, which illustrate the number of pixels that is classified as a certain land cover class in the sequence of primary, secondary, tertiary, quaternary, and quinary land cover class. The three main prime classes are "other", "forest", and "agriculture. Of these, forest is by far the most widespread primary class, with 74.8% of pixels within South Korea assigned to this class. The most frequent secondary class is "agriculture", followed by "forest", "other", and "grassland". Agriculture is secondary land cover class for more than half of the land area within South Korea. As third class, "grassland", and "other" are most often, assigned to 31.3% and 26.3% of pixels respectively. The class "wetland" is rare, but tends to occur more often as hierarchy decreases.

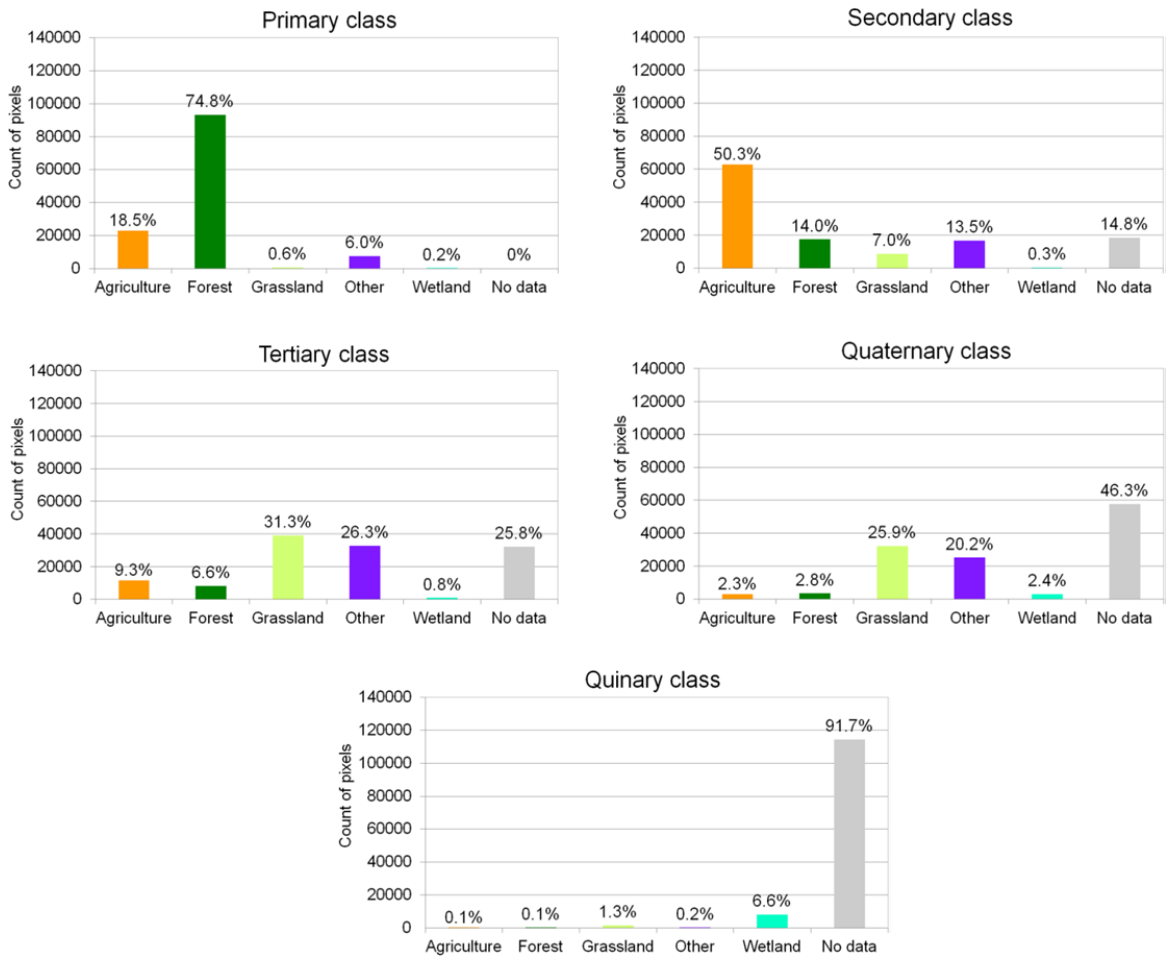


Fig. 3: Count and percentage of pixels, which have been classified as a certain land cover class (agriculture, forest, grassland, other, wetland) in the five successive land cover layers.

The diagrams from figure 3 also provide information about how many pixels have already been fully explained by the previous land cover classes. Only 14.8 % of 1 km² pixels within South Korea are fully covered by one land cover class, i.e. the assigned primary class. 25.8% of pixels are fully explained by maximum two classes, 46.3% by maximum three classes, and 91.7 % by maximum four classes.
Table 1 shows the frequency of weights assigned to the five successive land cover classes. 44.2% of pixels are covered by the primary land cover class to more than 90%. Lower weights, i.e. coverages, become less frequent for the prime class. The minimum weight for the primary class is between 30% and 40%. The secondary class has weights between >0% and 50%. The third class reaches at maximum coverage between 20% and 30%. For the majority of pixels that contain a third land cover class, its coverage is less than 10%. This is similar for the forth class. Only few 1 km² pixels within South Korea contain a fifth class, which covers not more than at maximum 10% of the 1 km² pixel.

Tab. 1: Frequency of weights assigned to the five successive land cover layers. Weights are grouped in coverage intervals of 10%. The frequency gives the percentage of pixels with a weight within the given interval.

	Frequency in %				
Coverage in %	**Primary weight**	**Secondary weight**	**Tertiary weight**	**Quaternary weight**	**Quinary weight**
>0.1 - 10	0	30.4	52.1	40.2	3.1
>10 - 20	0	18.3	10.0	1.1	0
>20 - 30	0	15.0	2.2	0	0
>30 - 40	1.3	11.1	0	0	0
>40 - 50	6.7	4.8	0	0	0
>50 - 60	10.4	0	0	0	0
>60 - 70	11.1	0	0	0	0
>70 - 80	11.9	0	0	0	0
>80 - 90	14.3	0	0	0	0
>90 - 100	44.2	0	0	0	0

3.2 Comparison to Global Land Cover Datasets

Previous comparison of global land cover datasets have revealed significant differences between products globally (GIRI et al. 2005; HEROLD et al. 2008; JUNG et al. 2006; MCCALLUM et al. 2006). In this section, we analyse how global 1 km land cover products compare to the regional land cover classification for South Korea. Three freely available global land cover datasets at coarse spatial resolution were selected for comparison: the MODIS land cover (LC) product for 2010 (FRIEDL et al. 2002), the Global Land Cover 2000 product based on SPOT-VEGETATION data (GLC 2000) (BARTHOLOMÉ & BELWARD 2005), and the ESA CCI land cover map (CCI-LC) for 2010 (UCL-GEOMATICS 2017). From the MODIS LC product, classification type 3 (LAI/fPAR) was chosen, because the land cover classification optimized for LAI/fPAR derivation is most suitable as input for NPP modelling. For comparison of area coverage, individual land cover classes were grouped to a simplified harmonized legend, as summarized in table 2.

Tab. 2: Harmonized classification for comparison and the included land cover classes from the individual legends of the high-resolution land cover map and the three global land cover datasets.

Summary class	High resolution LC	MODIS LC (Type 3)	GLC 2000	CCI-LC
Agriculture	Agriculture	Cereal Crops/Grasses (1), Broad-leaf crops (3)	Cropland (16), Mosaic of cropland and other classes (17, 18)	Rainfed cropland (10, 11, 12), Irrigated cropland (20), Mosaic cropland (>50%)/natural vegetation (30)
Forest	Forest	Forest (5, 6, 7, 8), Shrubs (2)	Tree cover (1, 2, 3, 4, 5, 6, 7, 8, 9, 10), Shrub cover (11, 12)	Tree cover (50, 60, 61, 62, 70, 71, 72, 80, 81, 82, 90), Mosaic tree and shrub/herbaceous cover (100), Shrubland, (120, 121, 122), Sparse tree (151), Sparse shrub (152)
Grassland	Grassland	Savanna (4)	Herbaceous cover (13), Sparse herbaceous or sparse shrub cover (14)	Grassland (130), Mosaic of cover types (40, 110), Lichens and mosses (140), sparse vegetation (150), Sparse herbaceous cover (153)
Wetland	Wetland	-	Regularly flooded shrub or herbaceous cover (15)	Tree cover, flooded (160, 170), Shrub or herbaceous cover, flooded (180)
Other	Urban, Bare, Water	Water (0), Non-vegetated (9), Urban (10)	Bare areas (19), Water bodies (20), Snow or Ice (21), Urban areas (22)	Urban areas (190), Bare areas (200, 201, 202), Water bodies (210), Permanent snow and ice (220)

The four maps in the top line of figure 4 show a detail of the South Korea high-resolution land cover classification with 30 m spatial resolution and the coarse resolution (~1 km) primary, secondary, and tertiary land cover classes derived in this study. Non-vegetated land cover classes from the original map were grouped in the class “other”, because a distinction of these classes is not needed for NPP modelling. The three global land cover datasets are also displayed with a simplified legend in figure 4 (see maps in bottom line).

The diagram in figure 5 shows the percent area of South Korea that is covered by the five harmonized land cover classes in the different land cover classifications. Based on the high-resolution land cover map, we observe percent coverages of 21% for agriculture, 69% for forest, 3% for grassland, and 8% for “other”; wetland is rare (0.2%). For the derived primary classes from this study, we observe a slight underestimation of agriculture (-2%), grassland (-2%), and forest (-2%), and an overestimation of forest (6%). As described in section 4.1, this can be compensated for by including the secondary and tertiary classes and applying weights.

The percent distribution of land cover classes within South Korea from the MODIS land cover product fits well to the high-resolution reference map, with a slight overestimation of agriculture (2%) and an underestimation of “other” (-2%). The ESA CCI-LC product is also relative close to the reference map, with an overestimation of agriculture (5%) and grassland (7%), and an underestimation of forest (11%). The largest discrepancy compared to the reference map and the other datasets exists for the GLC 2000. This datasets largely overestimates agriculture by 25% and underestimates other classes, mainly forest (-16%) and “other” (-7%).

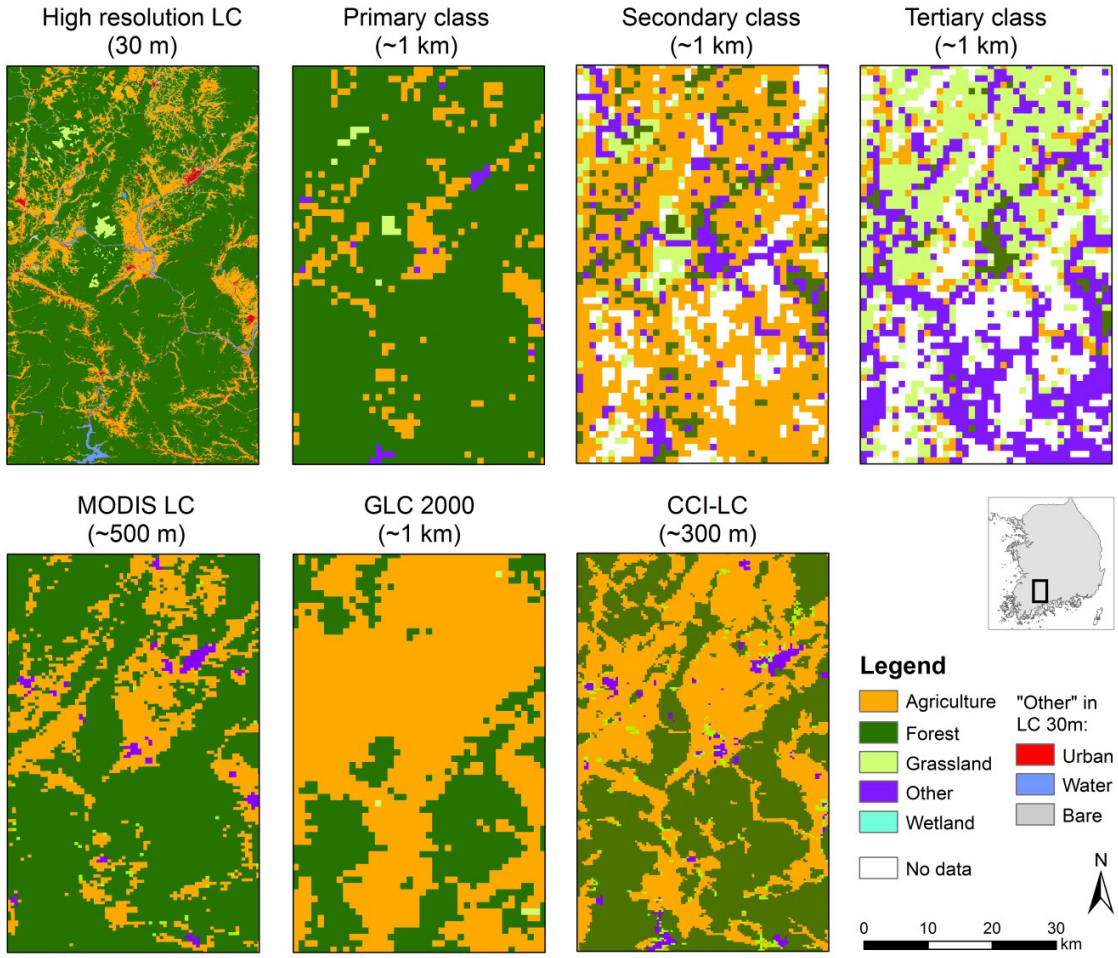


Fig. 4: Detail of the South Korea high-resolution land cover classification (top left), the primary, secondary, and tertiary land cover classes at ~1 km resolution derived in this study (top 2nd, 3rd, and 4th map), and three global land cover datasets with simplified legend for comparison bottom).

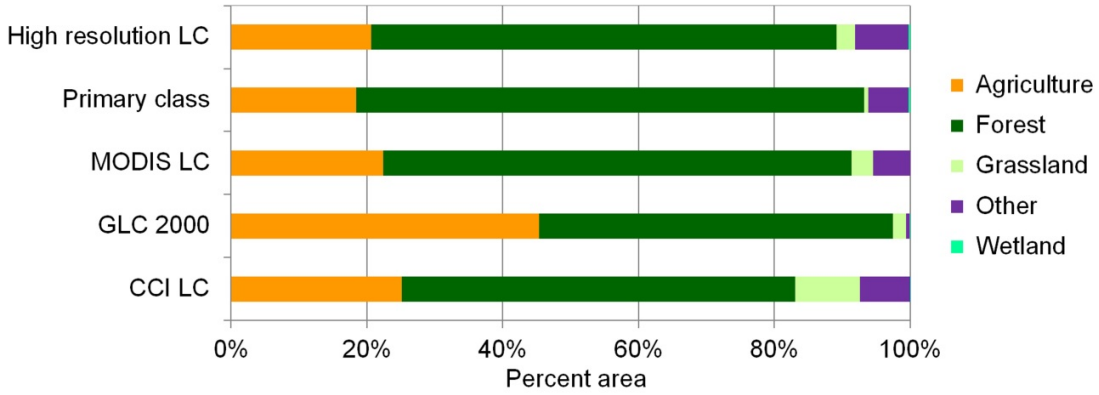


Fig. 5: Diagram showing the percent area of South Korea that is covered by the five harmonized land cover classes in the different land cover classifications.

Figure 6 shows further diagrams for comparison of the land cover classifications. The first row of diagrams shows, what percentage of the area classified as agriculture in the primary land cover class (1 km spatial resolution, derived from the Korean land cover map), is covered by which land cover class in the three global land cover products. The second and third rows show the assigned land cover classes within the three global land cover products for the area, which is classified as forest and grassland in the primary land cover class, respectively. Agriculture is given in orange, forest in dark green, grassland in light green, and "other" in violet. The class "wetland" occurs too rarely to be visible in the diagrams.

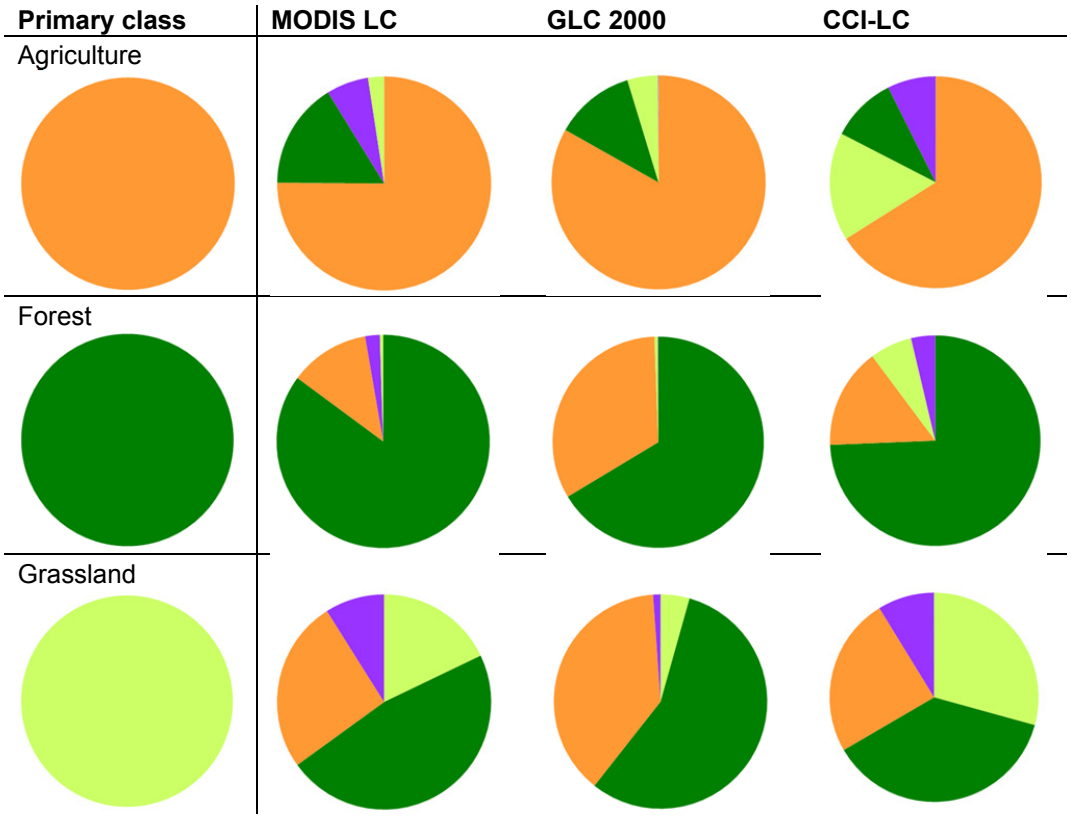


Fig. 6: Comparison of land cover products for South Korea. The diagrams show what percentage of the area classified as one of the three major land cover classes in the primary land cover class (1 km) derived from the South Korean land cover map, is covered by which land cover type (orange: agriculture, dark green: forest, light green: grassland, violet: other) in the global land cover maps.

Within this study we analysed how well a high resolution land cover classification is represented by coarse resolution land cover datasets. Of course, for NPP modelling, it is also of importance, which plant functional types are differentiated. Especially for agriculture and forest it is important to know crop types and forest types. The Landsat-based high-resolution land cover map at the coarse class level for South Korea did not differentiate crop or tree types, so we can only compare information from the global land cover maps. Different forest types assigned for South

Korea are listed in table 2 with their relative share of the forest area. The three classifications obviously vary largely with respect to forest types. None of these maps differentiates crop types.

4 Conclusions

High spatial resolution land cover data provide valuable information about the distribution of vegetation on the land surface. NPP models usually derive vegetation productivity on a coarser spatial resolution and do not take into account this wealth of information. A spatial resolution of 1 km can already be considered high resolution for NPP modelling. However, in many regions on the Earth land cover and land use patterns vary on a scale smaller than 1 km². In this study, we analyse how well a coarse resolution ~1 km land cover classification is able to represent land cover classification at higher (30 m) resolution for the study area of South Korea.

We find that the primary vegetation class at 1 km spatial resolution is only able to represent higher resolution land cover heterogeneity for 14.8% of the area. Especially widely distributed agricultural areas, which were not the majority class at 1 km² were missed. Multi-layer land cover information with secondary and tertiary classes is able to fully represent 25.8% and 46.3% of higher-resolution land cover heterogeneity. For our test case and the example classification, we thus recommend including at least two vegetation classes. With three successive layers, already a good representation of the high resolution classification can be reached, as remaining land cover proportions typically not exceed 10%. The land cover base information for NPP modelling should include both the class assignment and its sub-pixel coverage. For other study areas or land cover classifications, required classes might vary. The number of desirable classes should be defined based on both the percentage of pixels that are covered within each layer and the remaining percent of not yet captured classes.

The comparison with frequently used global land cover datasets reveals varying agreement in classes' area coverage with the regional land cover classification for South Korea. Especially the GLC 2000 shows strong discrepancy and seems not a suitable base land cover map for NPP modelling for South Korea. For different areas globally, advantages and shortcomings of various global land cover datasets may vary. For regional applications, a suitable base land cover map should be chosen carefully.

If possible, we recommend using information from regional land cover data, as demonstrated in this study. Moreover, for NPP modelling it is commonly required to know plant functional types. Different vegetation types can especially be distinguished for forest and crops. While most global land cover products distinguish different forest types, crop types are not differentiated. Here, a regional land cover classification with a more detailed legend can be of major advantage. For further improving regional NPP modelling, it would be desirable to additionally down-scale available LAI time-series based on higher resolution land cover and vegetation condition data.

We conclude that for regions with similar land cover heterogeneity to South Korea, land cover maps providing only the majority land cover class at 1 km seem not able to sufficiently represent the heterogeneity of land cover on regional scale. This deficiency becomes even the more relevant, when we consider, that some NPP models are able to include more than one vegetation type per pixel. For such models, multi-layer information about both land cover types and their pixel

coverage, i.e. relative weight, are of importance to take advantage of available higher-resolution land cover information.

5 References

BAO, G., BAO, Y., YIN, Z., XIN, X, BAO, Y., BAYARSAIKAN, S., ZHOU, Y. & CHUNTAI, B, 2016: Modeling net primary productivity of terrestrial ecosystems in the semi-arid climate of the Mongolian Plateau using LSWI-based CASA ecosystem model. International Journal of Applied Earth Observation and Geoinformation, **46**, 84-93.

BARTHOLOMÉ, E. & BELWARD, A.S., 2005: GLC2000: A new approach to global land cover mapping from Earth Observation data. International Journal of Remote Sensing, **26**, 1959-1977.

EISFELDER, C., KLEIN, I., NIKLAUS, M. & KUENZER, C., 2014: Net primary productivity in Kazakhstan, its spatio-temporal patterns and relation to meteorological variables. Journal of Arid Environments, **103**, 17-30.

EISFELDER, C., KLEIN, I., BEKKULIYEVA, A., KUENZER, C., BUCHROITHNER, M.F. & DECH, S., 2017: Above-ground biomass estimation based on NPP time-series – a novel approach for biomass estimation in semi-arid Kazakhstan. Ecological Indicators, **72**, 13-22.

FRIEDL, M.A., MCIVER, D.K., HODGES, J.C.F., ZHANG, X.Y., MUCHONEY, D., STRAHLER, A.H., WOODCOCK, C.E., GOPAL, S., SCHNEIDER, A., COOPER, A., BACCINI, A., GAO, F. & SCHAAF, C., 2002: Global land cover mapping from MODIS: Algorithms and early results. Remote Sensing of Environment, **83**, 287-302.

GESSNER, U., MACHWITZ, M., ESCH, T., TILLACK, A., NAEIMI, V., KUENZER, C. & DECH, S., 2015: Multi-sensor mapping of West African land cover using MODIS, ASAR and TanDEM-X/TerraSAR-X data. Remote Sensing of Environment, **164**, 282-297.

GIRI, C., ZHU, Z. & REED, B., 2005: A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. Remote Sensing of Environment, **94**(1), 123-132.

HEROLD, M., MAYAUX, P., WOODCOCK, C.E., BACCINI, A. & SCHMULLIUS, C., 2008: Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. Remote Sensing of Environment, **112**(5), 2538-2556.

JUNG, M., HENKEL, K., HEROLD, M. & CHURKINA, G., 2006: Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sensing of Environment, **101**(4), 534-553.

KLEIN, I., GESSNER, U. & KUENZER, C., 2012: Regional land cover mapping and change detection in Central Asia using MODIS time series. Applied Geography, **35**, 219-234.

LEINENKUGEL, P., OPPELT, N. & KUENZER, C., 2014: A new land cover map for the Mekong: Southeast Asia's largest transboundary river basin. Pacific Geographies, **41**, 10-15.

MATSUSHITA, B. & TAMURA, M. (2002) - Integrating remotely sensed data with an ecosystem model to estimate net primary productivity in East Asia. Remote Sensing of Environment, **81**, 58-66.

MCCALLUM, I., OBERSTEINER, M., NILSSON, S. & SHVIDENKO, A., 2006: A spatial comparison of four satellite derived 1 km global land cover datasets. International Journal of Applied Earth Observation and Geoinformation, **8**(4), 246-255.

NIEMEIJER, C., 2002: Developing indicators for environmental policy: data-driven and theory-driven approaches examined by example. Environ. Sci. Policy, **5**(2), 91-103.

NIKLAUS, M., EISFELDER, C., GESSNER, U. & DECH, S., 2015: Land Degradation in South Africa - A Degradation Index Derived from 10 Years of Net Primary Production Data. In: Remote Sensing Time Series Revealing Land Surface Dynamics, Editors: Kuenzer C., Dech, S., Wagner, W.; Springer, Chapter 12, 247-267.

PRIETO-BLANCO, A., NORTH, P.R.J., BARNSLEY, M.J. & FOX, N., 2009: Satellite-driven modelling of Net Primary Productivity (NPP): Theoretical analysis. Remote Sensing of Environment, **113**, 137-147.

SCHIMEL, D.S., 1995: Terrestrial biogeochemical cycles: global estimates with remote sensing. Remote Sensing of Environment, **51**(1), 49-56.

SIMIC, A., CHEN, J., LIU, J. & CSILLAG, F., 2004: Spatial scaling of net primary productivity using subpixel information. Remote Sensing of Environment, **93**(1-2), 246-258.

TUM, M., STRAUSS, F., MCCALLUM, I., GÜNTHER, K. & SCHMID, E., 2012: How sensitive are estimates of carbon fixation in agricultural models to input data? Carbon Balance and Management, **7**(3), 1019-1034.

TUM. M., ZEIDLER, J.N., GÜNTHER, K.P. & ESCH, T., 2016: Global NPP and straw bioenergy trends for 2000-2014. Biomass and Bioenergy, **90**, 230-236.

UCL-GEOMATICS, 2017: Land Cover CCI. Product User Guide Version 2.0 (CCI-LC-PUGV2). Available online: https://www.esa-landcover-cci.org/?q=webfm_send/91.

WIßKIRCHEN, K., TUM, M., GÜNTHER, K.P., NIKLAUS, M., EISFELDER, C. & KNORR, W., 2013: Quantifying the carbon uptake by vegetation for Europe on a 1 km² resolution using a remote sensing driven vegetation model. Geoscientific Model Development, **6**, 1623-1640.

Analyzing the Spatial Relationship between Building Volumes and Land Surface Temperature in Upper-Hill, Nairobi, Kenya

PATRICIA W. MWANGI[1,2]

Abstract: In urban areas, temperatures are higher than rural areas at night as the impervious surfaces absorb heat during the day and release it at night. Higher energy requirements within buildings, health problems, air pollution are just a few problems associated with urban heat island effect. Increasing populations and greater demand for urban space has led to the destruction of vast vegetative cover for impervious surfaces. Remote sensing studies on surface temperatures have mainly focused on the 2 dimensional aspect of buildings, without much emphasis on 3 dimensional analyses. The main objective of this paper is to analyse the geometry of individual buildings, height, area and volume, and determine their relationship with land surface temperature. In this paper, topographical data acquired from Survey of Kenya was extracted and used to generate a 2 meter digital elevation model (DEM) to calculate building heights. DEM accuracy was carried out using ground control points (GCP) and a value of 0.24 meters was obtained. Landsat 30 meter imagery was analysed to generate land surface temperatures (LST). Temperatures of buildings within the study area had a maximum and minimum of 27.6 7^0C and 35.40 ^{0}C within the gridded surface. Height, area and volume regression analysis results showed they each had a negative correlation with land surface temperature with height having the highest R^2 value of 0.08.

1 Introduction

Transformation of the earth's surface into urban areas has modified the temperature within these areas compared to surrounding rural environments resulting to a phenomenon known as urban heat island (UHI) (ZHAO et al. 2015). Built-up densities and population are factors that affect the development of urban heat islands (UHI) with few studies considering built-up volumes and heights in urban heat studies (WU et al. 2013). Improvements in remote sensing sensor spatial resolution have enabled research in urbanization and surface temperature possible. Satellite sensors that have thermal bands such as Landsat (100 meters) and ASTER (90 meter) and are freely available, have enabled research over vast areas where fine thermal imagery may not be accessible or expensive to acquire. Land surface temperature studies are important in determining conducive environmental conditions for people to live in (WU et al. 2013).
Previous studies have examined land surface temperature in large cities, but there is need to examine changes in land surface temperature in smaller regions within the city or medium-sized cities which develop at a faster rate than other parts of the city (WU et al., 2013). Research contribution on individual buildings indicate the importance of pattern and closeness of land features as well as density and not only the surface material (ZHAO et al. 2015).

[1] Kenyatta University, Department of Spatial Planning and Urban Management, Nairobi, Kenya
E-Mail: mwangi.patricia@ku.ac.ke
[2] Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Germany

Digital surface models derived from very high resolution data, either from stereo-satellite imagery or aerial photos provide highly detailed topographical information that provides critical data for planning or disaster management (ECKERT 2008; ALOBEID et al. 2009; EL GAROUANI et al. 2014; ALOBEID 2011).WU et al. (2013) developed a 3dimensional urban index (3DUI) in studying UHI while considering the height and vertical parameters in built-up areas where heat waves were experienced. ZHAO et al. (2015) studied the relationship between rooftop configuration from Quickbird imagery and surface temperatures from high resolution MASTER thermal data. The purpose of this research is to demonstrate the use of remote sensing in analyzing the impact the building size has with regards to land surface temperature. This research intends on analyzing land surface temperature distribution on individual buildings in the study area in relation to its physical dimensions.

2 Study Area

The study area is in Upper-Hill, Nairobi Kenya, which is 4 Km from the central business district (CBD). The study area covers approximately 4 Km^2. The area was selected due to its rapid development as businesses choose locations where there is less traffic congestion, ample parking pace and better access. Ground coverage ratio within the study area has increased to 35-60% and plot ratio of 150-300% leading to an increase in impervious surfaces in the area (MWANGI et al. 2018). Rapid urban developments have occurred with wooded grasslands replaced with concrete structures. This has led to an increase in minimum and maximum temperatures of 1.6 ^{0}C and 3.65 ^{0}C respectively between 1987 and 2017 ((MWANGI et al. 2018).

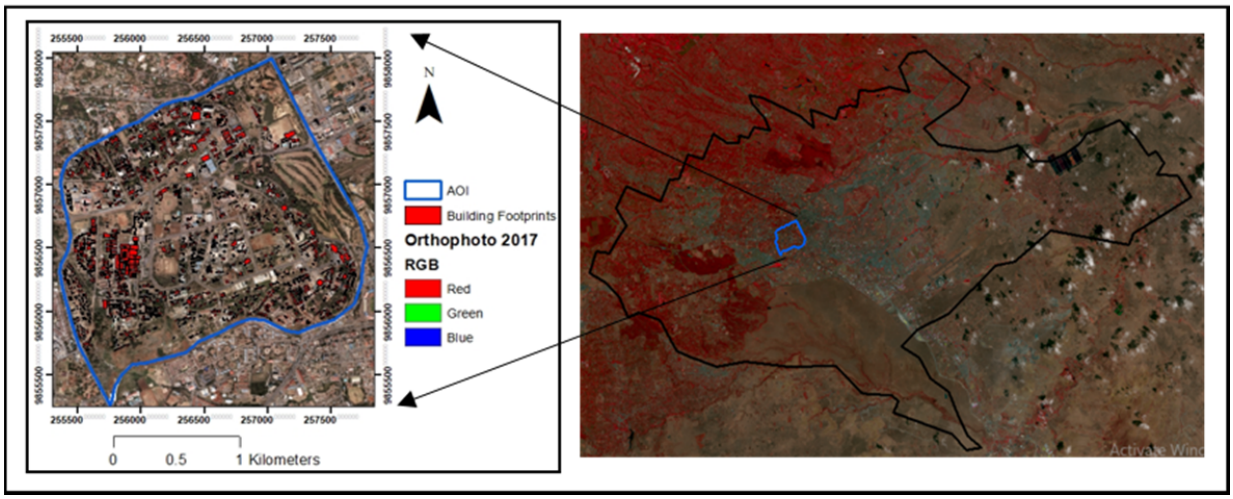


Fig. 1: Location of UpperHill in Nairobi County, Kenya

3 Data processing and methods

Digital topological datasets developed from the stereo imagery acquired in February 1998 were obtained from Survey of Kenya (SoK), which is the national mapping agency in Kenya, at a scale of 1:2500. Landsat L1T imagery, at 30 meter spatial resolution, acquired on 21st February 2000 Landsat 7 sensor was downloaded from the USGS website. There were no thermal data products available during the year which the stereo-aerial imagery was acquired. All data was

processed at UTM, Arc1960 Zone 37 South projection. Figure 2 below gives a summary of the data processes performed on the data.

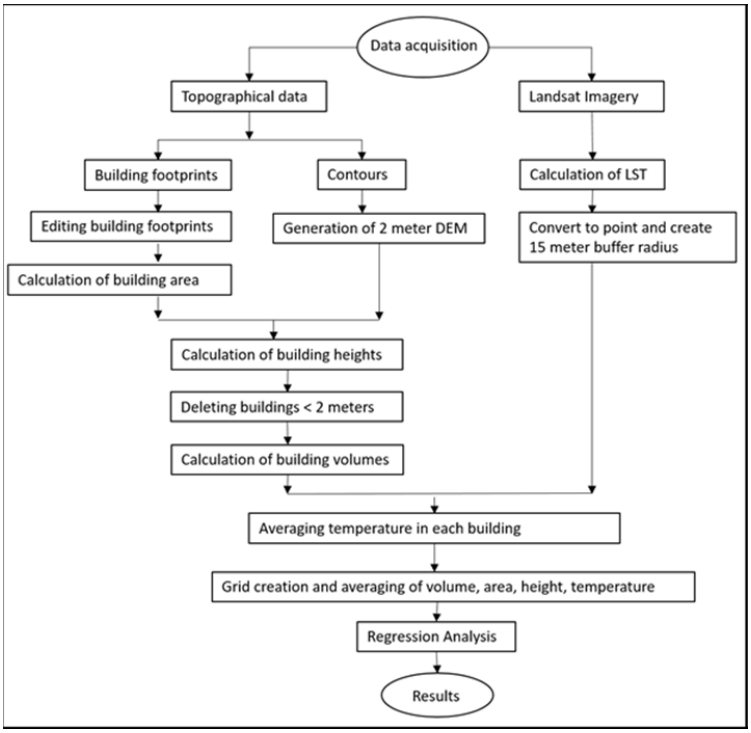


Fig. 2: Methodology framework for topographical datasets and LST

3.1 Editing and Validation of DEM and Building Footprints

Building footprints were extracted from the polygon dataset which also contained other feature classes that were not related to building footprints. Digital stereo-imagery acquired in February 2017 by Ramani Geosystems, Kenya was processed and the orthophoto used a background image in editing the buildings which were still observable from the time of acquisition of the aerial image as shown in figure 3.

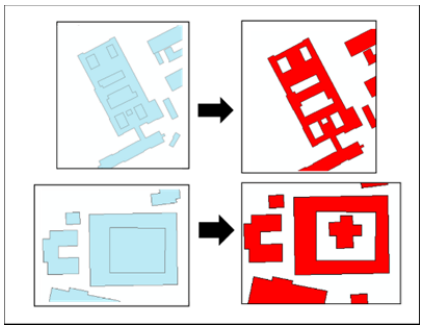


Fig. 3: Edited building footprints using orthophoto as background image

3.1.1 Accuracy of DEM

A DEM at a spatial resolution of 2 meters was generated from the contours dataset from the 1998 topographical data, which had an interval of 2 meters. A value that is calculated to determine DEM accuracy is the root mean square error (RMSE).

$$\text{RMSE} = \sqrt{\frac{\sum_i^m (X_i - X_j)^2}{m-1}} \quad (1)$$

Where X_i is the DEM cell value and X_j is the corresponding sample point elevation, m is the number of sample points.

The accuracy of the interpolated DEM was assessed using GCP points collected in 2017 in areas that had not changed from 1998. An RMSE of 0.23 meters was obtained using equation (1). This accuracy shows how accurately the generated DEM represents the ground truth (Guo-an *et al.*, 2001).

The building dataset had attribute elevation data which enabled the calculation of building heights as the difference between the building elevation and the DEM. Buildings whose height was less than 2 meters were deleted, hence having 1348 building footprints as the final dataset for analysis. Building area was automatically calculated in ArcGIS and verification done manually. Building volume for each building was calculated as the product between the area and its height in ArcGIS 10.4.

3.1.2 Accuracy of Buildings

Since we were not able to measure the building heights to verify the accuracy of the building footprints, data was obtained from internet sources from (http://www.emporis.com/en/wm/ci/bu/?id=100051, visit: 23.01.2018). This was in regards to buildings that are still existing in the study area. Floor height was also used to determine the approximate height of the building since building height information is not freely available in this study area. Only three buildings whose data was available from the internet were used to verify as shown in table 1.

Tab. 1: Comparison between actual and extracted building heights

Building name	Floors	Actual Height (m)	Extracted Height (m)	Difference
Social security House	28	103	105.19	2.19
Ministry of Transport	13	43	25.90	17.10
Ardhi House	12	36	23.37	12.63

There was an underestimation of building heights especially for the shorter buildings. Without the stereo-aerial photos or reference data it would be difficult to determine whether the error occurred during the digitization process or from processing the stereo-photos. Eckert, (2008) compared DSM extraction of high resolution stereo-satellite imagery using different software and results showed that in each of the software, buildings less than 50 meters had height differences of <10 meters.

3.2 Calculation of Land Surface Temperature

Landsat imagery was used to calculate land surface temperature for the study area using equations published by USGS for processing data acquired from Landsat 7 ETM + sensor (NASA, 2009). Computations were carried out using spatial modeler in ArcGIS 10.4.

3.2.1 DN Values to Top of Atmosphere (TOA)

Spectral information is in digital number (DN), which has to be converted to reflectance values for analysis. Landsat 7 ETM+ has two thermal bands, 6a and 6b, resampled to 30 meters. Band 6a was used for this analysis due to its low radiance variance. Equation (2) shows the conversion from DN to Top of Atmosphere (TOA) radiometric values.

$$L_\lambda = \left(\frac{L_{max} - L_{min}}{Q_{calmax} - Q_{calmin}}\right) * (Q_{cal} - Q_{calmin}) + L_{min} \quad (2)$$

Where,

L_λ spectral radiance
Q_{calmin} minimum quantized calibrated pixel value in DN
Q_{calmax} maximum quantized calibrated pixel value in DN
Q_{cal} DN value of the pixel
L_{min} minimum radiance detected by the sensor
L_{max} maximum radiance detected by the sensor

Equation (2) was used to calculate the at-satellite brightness:

$$T_B = K_2 / \ln\left(\frac{K_1}{L_\lambda} + 1\right) \quad (3)$$

Where:

T_B satellite brightness temperature in degrees Celsius
K_1 band specific thermal conversion constant (K1_CONSTANT_BAND_n, where n is band 6a)
K_2 is the band specific thermal conversion constant (K2_CONSTANT_BAND_n, where n is band 6a)

3.2.2 Emissivity

NDVI was calculated using the spectral radiance values of the red and infra-red (NIR) bands as calculated using equation (2) so as to determine the land surface emissivity (LSE).

$$R_{TOA_\lambda} = \frac{\pi L_\lambda d^2}{E_{sun\lambda} Sin\theta_{SE}} \quad (4)$$

Where:

R_{TOA} top of atmosphere (TOA) planetary reflectance for band x and is unitless
π 3.141592654
d Earth_Sun_distance in astronomical units
$E_{sun\lambda}$ Band specific mean solar exoatmospheric irradiance
θ_{SE} Sun elevation angle in degrees from the metadata

NDVI was then calculated using the resulting calculated reflectance values from equation (4) of the red and infra-red bands using equation (5).

$$NDVI = \frac{NIR-R}{NIR+R} \quad (5)$$

Equation (8) calculates the vegetation portion to obtain the LSE as shown in equation (6).

$$P_v = \left(\frac{NDVI-NDVI_{min}}{NDVI_{max}+NDVI_{min}}\right)^2 \quad (6)$$

Where:

P_v	vegetation portion
NDVI	normalized difference vegetation index
$NDVI_{min}$	minimum NDVI
$NDVI_{max}$	maximum NDVI

where the minimum NDVI is the value for pure soil normally given as 0.2 and maximum NDVI is the value of pure vegetation given as 0.5.

LSE is then computed using equation (7):

$$LSE = 0.004 * P_v + 0.986 \quad (7)$$

3.2.3 Land surface temperature

Using the at-satellite brightness temperature and the Land Surface Emissivity, LST was computed in degrees Celsius as shown in equation (8).

$$LST = [T_B/(1+(\lambda*\frac{T_B}{\rho})*ln(LSE))]-273.15 \quad (8)$$

Where:

LST	land surface temperature
T_B	at-satellite brightness temperature
λ	wavelength of emitted radiance (λ=11.5μm)
ρ	h*c/σ ($1.438*10^{-2}$ m K)
σ	Bolzmann's constant ($1.38 * 10^{-23}$ J K-1)
h	Planck's constant ($6.26 * 10^{-34}$ J s)
c	velocity of light ($2.998 * 10^{-8}$ m s-1)

Figure 4 below shows the calculated LST around the study area. Areas with lower temperatures have denser vegetation compared to other parts of the study area.

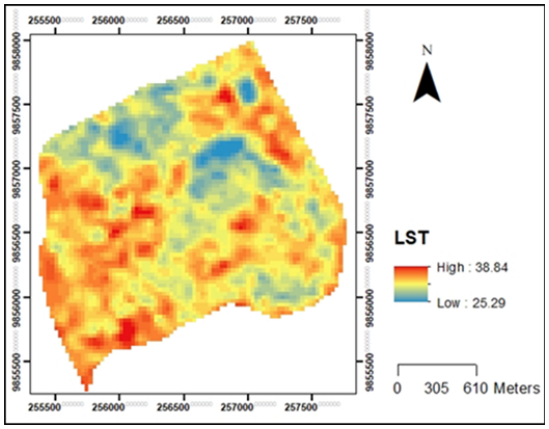


Fig. 4: Calculated LST

3.3 Calculation of Aggregate values

Land surface temperature raster dataset was converted to vector point data, where a 15 meter buffer around each point was created. Spatial analysis was performed between LST polygons and the building data, where a mean LST value was calculated for each building. Summary statistics for the dataset of buildings >2 meters' height are as shown in table 2.

Tab. 2: Summary statistics of elevation, building height, area and volume

	Minimum	**Maximum**	**Mean**
DEM Elevation (m)	1663.09	1735.95	1708.27
LST (^{0}C)	26.27	37.70	32.32
Height (m)	2.00	105.19	5.26
Area (m^2)	4.16	8570.08	261.76
Volume (m^3)	12.64	185483.42	2288.35

A polygon grid of cells 180 meters by 180 meters was created within the study area. This is because of the distribution and size of buildings especially in the center of the study area. 121 grid cells were created each having averaged values of LST, building volumes, area and height of buildings located in each cell, as shown in figure 5.

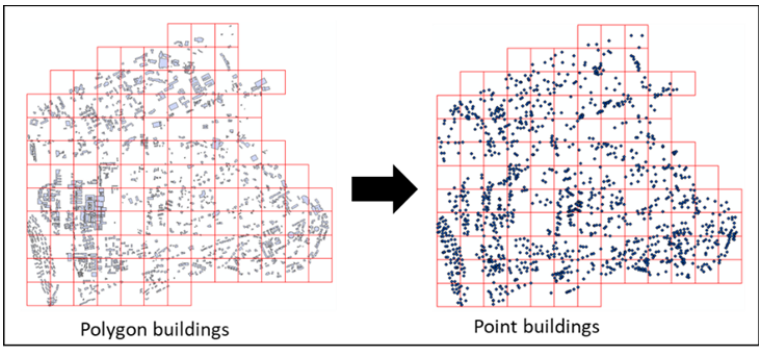


Fig. 5: Building footprints in 180 meter × 180 meter grid cells

4 Results and Discussion

A spatial distribution of the datasets within the grid is as shown in figure 6. Visual inspection in the lower left corner of the study area indicates that areas that had the highest temperatures recorded lower building volume and area. Higher volumes were towards the north western parts of the study area. Areas that have the lower heights, towards the lower right have higher temperatures compared to areas with higher buildings on the northern parts of the study area as in figure 6 (c). The minimum and maximum temperature range within the study area using the averaged values from the buildings was 27.67 ^{0}C and 35.40 ^{0}C respectively. This was comparatively lower than the minimum and maximum temperature which was 28.02 ^{0}C and 35.60 ^{0}C respectively, within the grid cells of the LST raster data without averaging the building dataset.
From figure 5, buildings tend to be more clustered towards the lower left. These buildings as well as those on the northern part of the study area have greater volumes as they are large commercial buildings. However, the northern part is highly vegetated compared to the other parts of the study area, which could explain the lower temperatures despite having higher density of buildings.

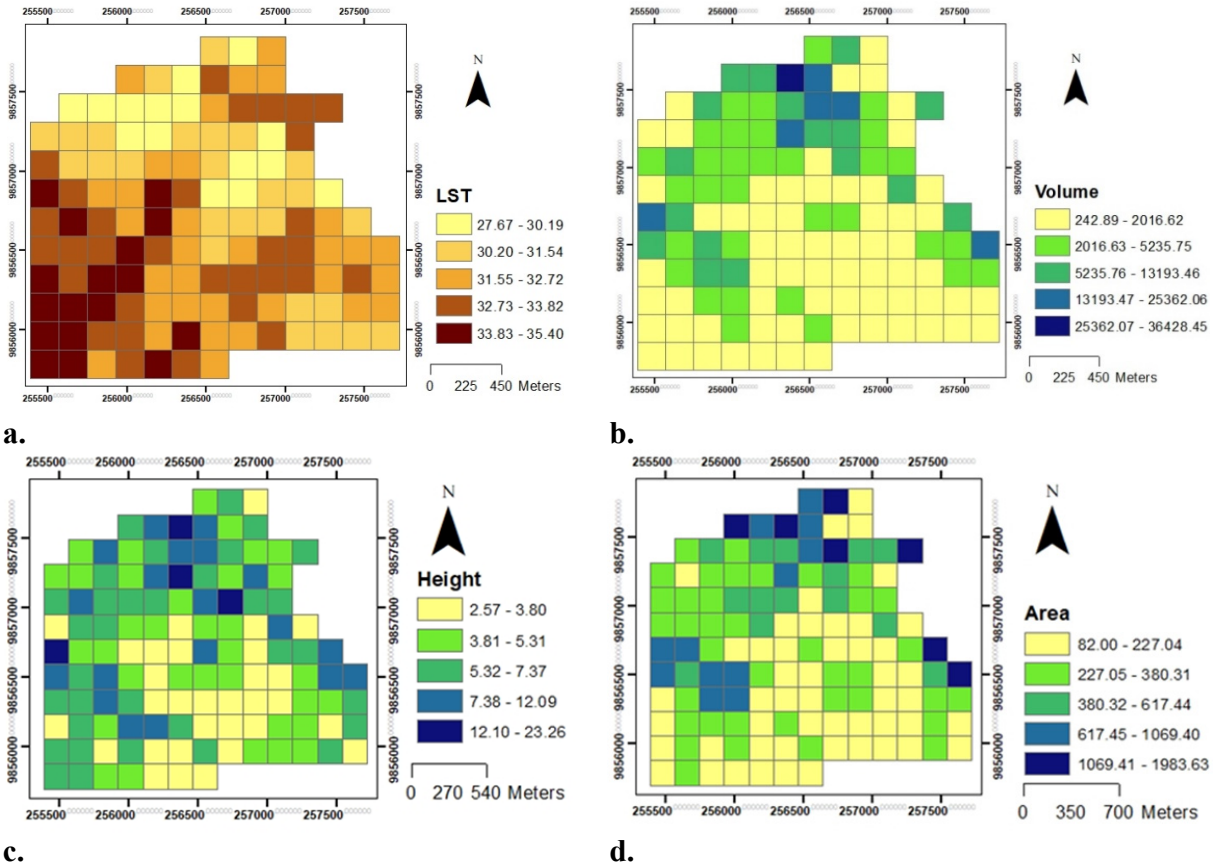


Fig. 6: Spatial distribution of LST (a), building volume (b), area (c) and height (d) in the grid

A linear regression analysis was undertaken on this dataset comprising of 121 spatially averaged building datasets from the grid. Results in figure 7 below show a negative correlation between land surface temperature with volume, area and height. The reliability of the model to predict LST using area, height and volume from the R^2 value indicate that they are not strong predictors, especially area, as also shown in figure 7 (b). This would indicate that there are other factors that affect these data sets such as vegetation density which would need to be considered.

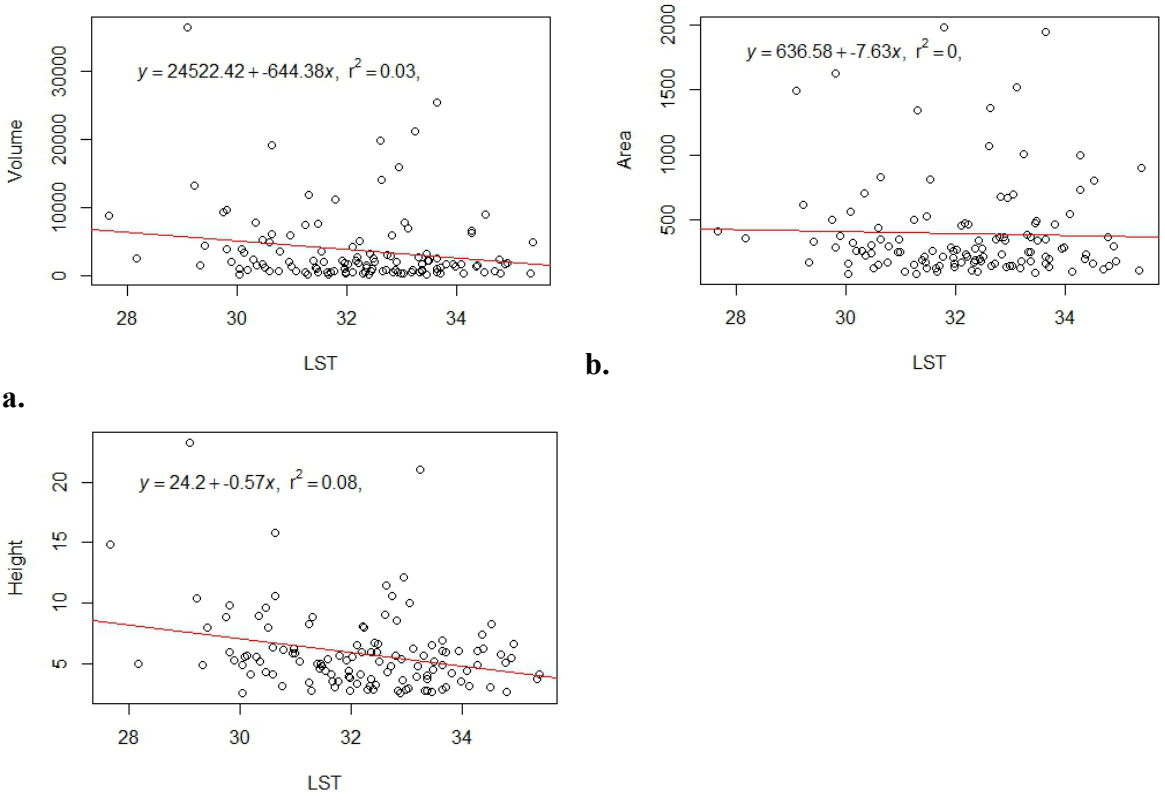


Fig. 7: Regression analysis of land surface analysis with volume (a), area (b), height (c)

5 Conclusion

Quantitative height accuracy and building shape is very important in deriving volumes of buildings for purposes of determining building energy requirements or for planning purposes. This is from the large difference in building heights between the actual building height and those obtained from the digitized buildings. This may have had influence in the results of the analysis. Results show that building height would be a stronger predictor for land surface temperature. Results show that the influence of building volumes on land surface temperature is not significant and cannot be sorely explained by buildings. Other factors would need to be taken into consideration such as vegetation density within the study area. Thermal remote sensing dataset from Landsat L1T data products have been resampled to 30 meters hence analysis on individual buildings can be undertaken in areas where higher resolution imagery is not available. Future work will involve comparison of results using time-series data.

6 References

ALOBEID, A. 2011: Assessment for Matching Algorithms for Urban DSM Generation from Very High Resolution Satellite Stereo Images. Fachrichtung Geodäsie und Geoinformatik der Leibniz University Hanover.

ALOBEID, A., JACOBSEN, K. & HEIPKE, C., 2009: Building height estimation in urban areas from very high resolution satellite stereo images. ISPRS Hannover Workshop, **5**, 2-5.

ECKERT, S., 2008: 3D-Building Height Extraction from Stereo IKONOS Data Quantitative and Qualitative Validation of Digital Surface Models Derivation of Building Height and Building Outlines.

EL GAROUANI, A., ALOBEID, A. & EL GAROUANI, S., 2014: Digital surface model based on aerial image stereo pairs for 3D building. International Journal of Sustainable Built Environment, **3**(1), 119-126.

GUO-AN, T.A.N.G., STROBL, J., JIAN-YA, G.O.N.G., MU-DAN, Z.H.A.O. & ZHEN-JIANG, C., 2001: Evaluation on the accuracy of digital elevation models. Journal of Geographical Sciences, **11**(2), 209-216.

MWANGI, P.W., KARANJA, F.N. & KAMAU, P.K., 2018: Analysis of the Relationship between Land Surface Temperature and Vegetation and Built-Up Indices in Upper-Hill, Nairobi. Journal of Geoscience and Environment Protection, **6**(1), 1-16.

NASA. 2009: Landsat 7 science data users handbook. National Aeronautics and Space Administration Landsat, 186.

WU, C.D., LUNG, S.C.C. & JAN, J.F., 2013: Development of a 3-D urbanization index using digital terrain models for surface urban heat island effects. ISPRS Journal of Photogrammetry and Remote Sensing, **81**, 1-11.

ZHAO, Q., MYINT, S.W., WENTZ, E.A. & FAN, C., 2015: Rooftop surface temperature analysis in an urban residential environment. Remote Sensing, **7**(9), 12135-12159.

Deep Learning for the Classification of Building Facades

DOMINIK LAUPHEIMER[1] & NORBERT HAALA[1]

Abstract: In recent years, 3D models containing both geometric and semantic information have become of great public interest. The geometric information of 3D models can be provided by (a combination of) photogrammetric methods, laser scanning and traditional surveying. For providing the semantic information for 3D urban models in an automated way, we established an end-to-end approach for classifying images of building facades into five different utility classes (commercial, hybrid, residential, specialUse, underConstruction) by using Convolutional Neural Networks (CNNs). We did several experiments on different data sets with various CNNs for evaluating the performance of this approach. Using Class Activation Maps (CAMs), we examined which features are learned during the training process in order to sort the facades into the considered classes.

1 Introduction

In the last few years Artificial Neural Networks (ANNs) have become state of the art in various research areas like image recognition, object recognition and speech recognition. ANNs are weighted graphs whose parameters $\boldsymbol{\theta}$ are trainable. Using ANNs is said to be Deep Learning. Particular ANNs that have been originally inspired by the human visual cortex are called Convolutional Neural Networks (CNNs). They proved to be very useful in the visual domain due to their properties like locality, parameter sharing and dimension reduction (RUSSAKOVSKY et al. 2015). The architecture of a CNN can be seen in Fig. 1.

Our aim is to establish an end-to-end approach for classifying street-level images of building facades into five different utility classes (see Tab. 1) by using CNNs. This is a highly complex real-world task struggling with various problems depending on image properties, object properties and environmental factors (variation of object scale and orientation, changing amount of buildings per image, cropped facades in images, occlusions, changing illumination, …). Moreover, the appearance of a facade is not necessarily in accordance with the actual usage of the building, e.g. there can be a medical practice in a former purely residential building without a constructional change of the facade. We assume, this discrepancy cannot be solved within the visual space.

For training the CNNs we use labeled Google Street View (GSV) images. The labels are acquired by the linking between labeled LoD3 building models and GSV images provided by TUTZAUER & HAALA (2017). The subsequent aim - using the trained CNN - is to enrich any 3D urban model where the semantic information is not already known.

We experiment on different data sets (see 2.2 Data Preparation) with various CNNs (VGG16, VGG19 (SIMONYAN & ZISSERMAN 2014), Resnet50 (HE et al. 2016), InceptionV3 (SZEGEDY et al. 2016), self-designed networks) in order to investigate the performance of our

[1] Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, E-Mail: [dominik.laupheimer, norbert.haala]@ifp.uni-stuttgart.de

end-to-end approach where no hand-crafted features are used. In order to examine which features are learned during the training process, we make use of Class Activation Maps (CAMs) visualizing areas being important for the network's decision (SELVARAJU et al. 2016). A human interpreter can derive the learned features by looking at those important areas. Furthermore, with the help of CAMs we investigate how our trained CNNs perform on data that has different properties than the used training data. Classification results and CAMs are shown in 2.4 Results.

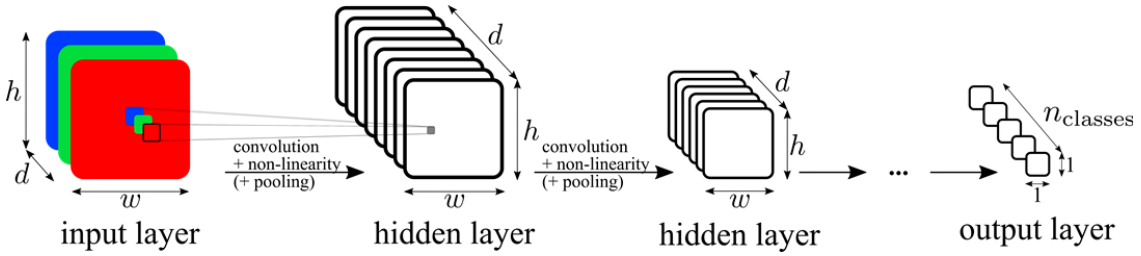


Fig. 1: Architecture of a CNN for multiclass classification with an input layer (here: RGB image), multiple hidden layers and an output layer. The amount of classes $n_{classes}$ defines the size of the output layer representing the estimated label vector $\hat{y}$. Spatial dimensions w and h depend on the input image and the used convolutions (and poolings). The depth d of hidden convolutional layers depends on the amount of used convolutional kernels applied to the previous layer. Each depth slice is a so-called activation map.

Recent years have shown a great effort in scene understanding and consequently, extending 3D urban models with semantic information. Similar to our task, MOVSHOVITZ-ATTIAS et al. (2015) perform a fine-grained multi label classification of store fronts exhausting GSV data. It can be seen as a more detailed classification of our class *commercial*. MARTINOVIC et al. (2015) do segmentation and labeling of facades in 3D space. Unlike us, they classify only components of facades whereas we classify the facades as a whole. Therefore our classification is on a higher abstraction level.

Tab. 1: Considered classes of building facades for our end-to-end approach.

name	**definition**
commercial	purely commercial use
hybrid	mixture of commercial and residential use
residential	purely residential use
specialUse	anything else not matching the other four classes, e.g. gyms and churches
underConstruction	sites being under construction independently on their actual building class

2 Setup and Training of the CNNs

The aim of our end-to-end approach is to provide an automated pipeline for extracting semantic information in images showing building facades. As a first step, we perform image classification considering only five different classes (see Tab. 1) as we want to investigate the feasibility of such a pipeline when using CNNs. In order to learn features properly, a huge amount of labeled

data for the training process has to be provided. As there doesn't exist any benchmark for our purposes, we have to provide a labeled data set on our own (see 2.2 Data Preparation). A labeled data set consists of images $\boldsymbol{x}$ labeled with a correspondent label $\boldsymbol{y}$, i.e. tuples $(\boldsymbol{x}, \boldsymbol{y})$. We provide different data sets with images for training (*train set*), for validating the network's performance during the training process (*validation set*) and for evaluating the network's performance after the training process (*test set*). Results are shown in 2.4 Results.

2.1 Pipeline

In order to use CNNs for image classification you have to train them by using ground truth data (see 2.3 Training). A labeled input image $(\boldsymbol{x}, \boldsymbol{y})$ is fed to the CNN, processed by its current parameters $\boldsymbol{\theta}$ which produce a predicted label vector $\hat{\boldsymbol{y}}$. The label vector can be interpreted as a probability density function over the considered classes. This output is compared to the ground truth $\boldsymbol{y}$ (one-hot label vector). The discrepancy between $\hat{\boldsymbol{y}}$ and $\boldsymbol{y}$ is backpropagated during the network wherefore the parameters $\boldsymbol{\theta}$ are updated (LECUN et al. 1998). The more labeled data are available for the training, the better the parameters can anneal to their ground truth values and the better is the achieved classification. Once training is done one can process any new unseen data considering the CNN as a black box accomplishing the classification (see Figure 2).

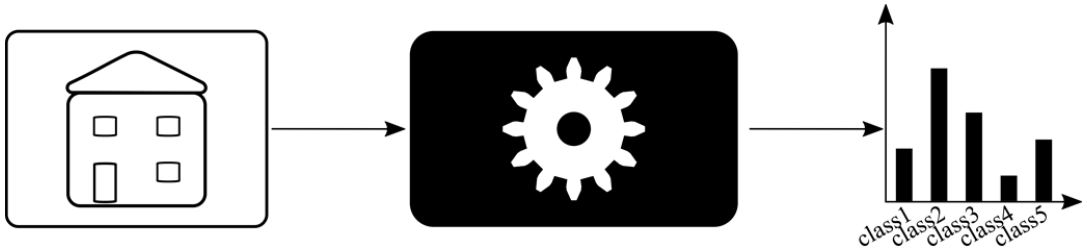


Fig. 2: Concept of the classification pipeline using CNNs. An input image showing a facade is processed by the CNN producing a probability density function over the considered classes

2.2 Data Preparation

Providing a big enough labeled data set in order to deal with the problem's complexity is a big challenge and a very tedious task. We provide different data sets for binary and multiclass classification tasks in order to investigate the impact of the separability of classes. All images are RGB images from GSV and labeled based on the fine-grained labeling given by TUTZAUER & HAALA (2017). These labels have to be sorted into our coarse classification given in Tab. 1. Obviously, in the real world there are more facades of residential buildings than facades belonging to other classes causing imbalanced data sets, initially. With the help of a priori data augmentation (horizontal flipping, warping, cropping, jittering and modification of saturation) we created equally distributed training data sets. I.e. every considered class has the same amount of samples.

2.3 Training

Generally, ANNs can be seen as function approximation machines consisting of a huge amount of parameters $\boldsymbol{\theta}$ which are learned during a training process by using a labeled *train set* (supervised learning). The schematic representation of an ANN is given in Figure 2. Equation (1)

shows the mathematical representation of the function approximation machine (*left side*) and its ground truth relation (*right side*) with perfectly known parameters $\boldsymbol{T}$ resulting in perfect labels $\boldsymbol{y}$.

$$\hat{\boldsymbol{y}} = f(\boldsymbol{x}, \boldsymbol{\theta}) \qquad \boldsymbol{y} = f(\boldsymbol{x}, \boldsymbol{T}) \qquad (1)$$

The learning process adapts the originally randomly initialized parameters $\boldsymbol{\theta}$ in an optimal sense. Mathematically speaking, the learning process is an optimization problem where the trainable parameters have to minimize a cost function $C(\boldsymbol{y}, \hat{\boldsymbol{y}})$ measuring the discrepancy between $\hat{\boldsymbol{y}}$ and $\boldsymbol{y}$. C is a proxy measure for the classification error. Due to the architecture of CNNs the learned parameters are stored in convolutional masks. Hence, they can be interpreted as feature detectors. We used RGB images of size 400px × 400px. Our *train set* consists in total of 75.000 images (15.000 images per class) due to data augmentation. *Validation set* and *test set* consist of 350 images each (70 images per class).

2.4 Results

Our work shows, that an overall accuracy of approximately 64% can be achieved with current state-of-the art models (VGG16, VGG19, ResNet50, InceptionV3) when being evaluated on data sets provided by us. As expected, classes *specialUse* (accuracy: 25.71%) and *underConstruction* (accuracy: 48.57%) perform worse than the other classes due to their definitions (high intra-class variance, bad separability) and lower the overall accuracy by approximately 18%. Furthermore, *residential* performs best (accuracy: 98.57%), which is pleasant as the majority of real-world buildings belongs to this class.

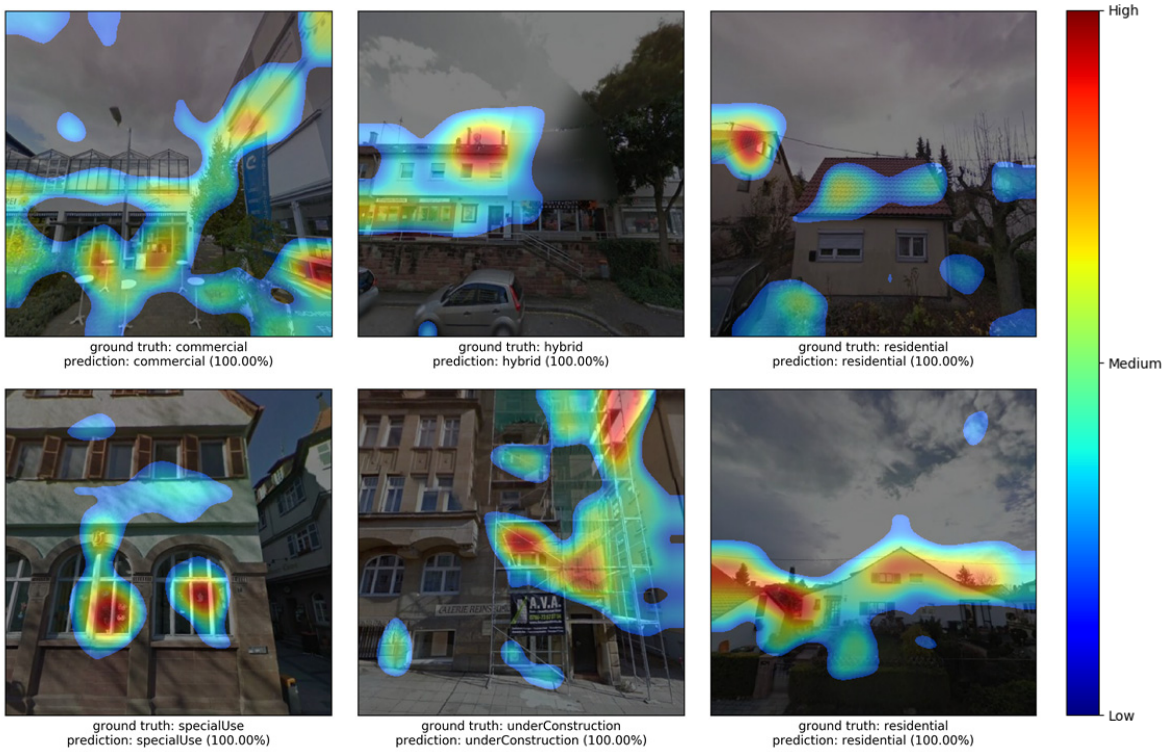


Fig. 3: Some examples of correctly classified images overlaid with CAMs showing the class-specific important areas. The image of class *specialUse* shows a kindergarten (*bottom left*) with bricolages at the windows. The respective color of the CAM represents the importance regarding the network's prediction

By looking at images of correctly classified images overlaid with CAMs (see Figure 3) we can collect a subset of learned features for each class (see Table 2). Interestingly, they are similar to features a human operator would use instinctively. CAMs base on the activation maps of the last convolutional layer of a trained CNN. These activation maps contain high-level features while conserving the spatial information (which is discarded by fully connected layers). The gradient of the predicted class with respect to the activation map shows the importance regarding the network's decision. I.e. the CAM is class-specific.

Tab. 2: Extracted learned features (interpreted by using CAMs)

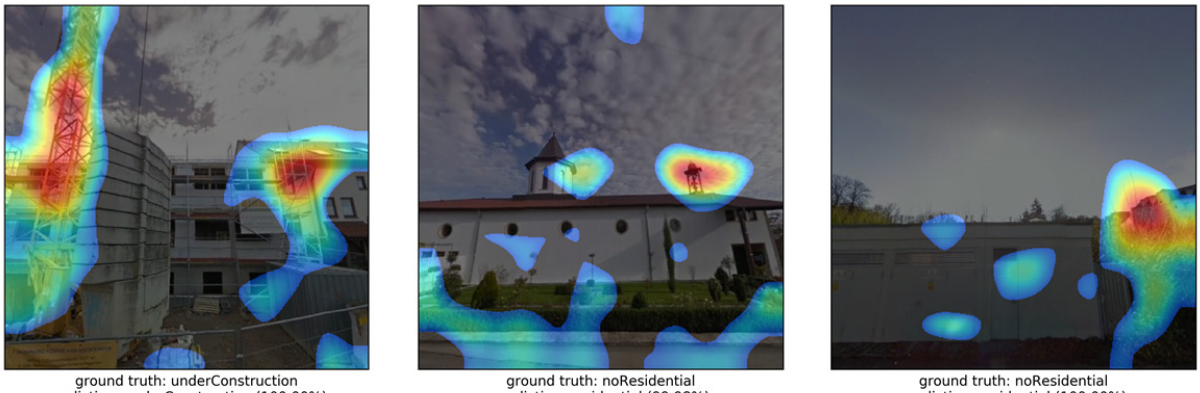
class	**features**
commercial	windows (arrangement, size, amount), advertising panels/signal colors, shop windows, big doorways, roof shape (flat roof), huge chimneys
hybrid	windows (arrangement, size, amount), advertising panels/signal colors, shop windows, roof shape (saddle roof), roof extensions (e.g. dormers, chimneys or antennas), environmental objects (sky, vegetation, cars, humans), big doorways
residential	windows (arrangement, size, amount), roof shape (saddle roof), roof extensions (e.g. dormers, chimneys or antennas), environmental objects (sky, vegetation, cars, humans), arrangement/skyline of adjacent buildings
specialUse	building shape, decorated windows, special doorways
underConstruction	scaffolds, cranes

(a shows a correctly predicted construction site overlaid with its CAM. By evaluating CAMs we could locate the problems for the image classification task. There are misclassifications due to not detected features and misclassifications due to misinterpreted features (see (b). Apart from these misclassifications, there are classifications that contribute to the classification error as predictions and ground truth labels differ due to different focusing. From a human point of view the building covering the majority is decisive for labeling, but CNNs only take care of features with the most impact on predictions - independently of location and size within the image. Basically, these predictions are not necessarily false predictions, but they still are lowering the overall accuracy if they do not match the human-given label. As a consequence, the classification of building facades, currently realized in form of image classification, underestimates the real power. (c shows exemplarily an image of such a "false" prediction overlaid by its corresponding CAM.

Results of our binary and ternary classifiers show that classification results can be improved by improving class separability, i.e. by reducing the intra-class variance (overall accuracies for the binary/ternary case: approximately 88% / approximately 73%). Our pragmatic approach was to reduce the amount of classes in order to avoid a new labeling.

In addition to that, CAMs are used to investigate how trained CNNs perform on different data representations. For that purpose, CAD models in LoD3, textured meshes (textured with images of Google Earth) and cut out buildings of GSV images are considered (see Figure 5). In the

following, we refer to them as non-image representations as the first two are 3D models and the latter are cut out images only. These representations have also been used by a user study on the human ability to perceive building classes from geometric representations (TUTZAUER & HAALA, 2017).



(a): Correct prediction (b): False prediction (c): "False" prediction

Fig. 4: Example images overlaid by corresponding CAMs of correct and false predictions for binary classification. The small bell tower in Figure 4b is mistaken as a chimney causing a false prediction. The "false" prediction in Figure 4c is a result of the discrepancy between human labeling and automatic labeling. The transformer house occupying the major area is decisive for the human-given label. The CNN focuses on the marginal residential building

Since data has to be provided to the network in the same manner like during training, trained CNNs cannot be applied instantaneously to the considered non-image representations. As we trained our CNNs on 400×400×3 sized images we have to provide RGB images of the same size to the networks. Therefore, snapshots of those non-image representations are taken.

Figure 5 shows snapshots of different non-image representations overlaid by corresponding CAMs. All representations have in common that there is no environmental information. A small exception is the LoD3 model generated by SketchUp. A human is placed automatically next to every 3D model by the program. This is worth mentioning, as *human* is an important feature for classes *hybrid* and *residential* (see Table 2). Please note, snapshots of textured meshes are taken from a bird's eye view which differs significantly from the exposure direction of our training images (street-level images).

Figure 6 exemplarily shows a correctly predicted building throughout all representations overlaid by corresponding CAMs. The first row shows snapshots of LoD3 models whereby in the left image the human is removed. It can be seen that the same areas activate for ground truth class *hybrid* (neglecting the human part in the right image). Removing the human means removing the most important feature (see the change of CAM's coloring and the worsening of the prediction confidence). This shows the importance of environmental objects for the classification.

The lower left shows the snapshot of the textured mesh representation. Its exposure direction differs from street-level images used for training the network. The lower right image shows the cut out GSV image (overlaid by its CAM) which achieves the best result.

To summarize, the more information is carried by the representation (e.g. color, environmental objects, etc.) and the more the representation converges to the representation used during

training, the better becomes the classification result. Snapshots of Figure 6 show that CNNs are a manifestation of *weak Artificial Intelligence*. CNNs can solve tasks reliably they are familiar with. The more the task's preconditions or the task itself differs from the task during training, the more limited is the classification quality.

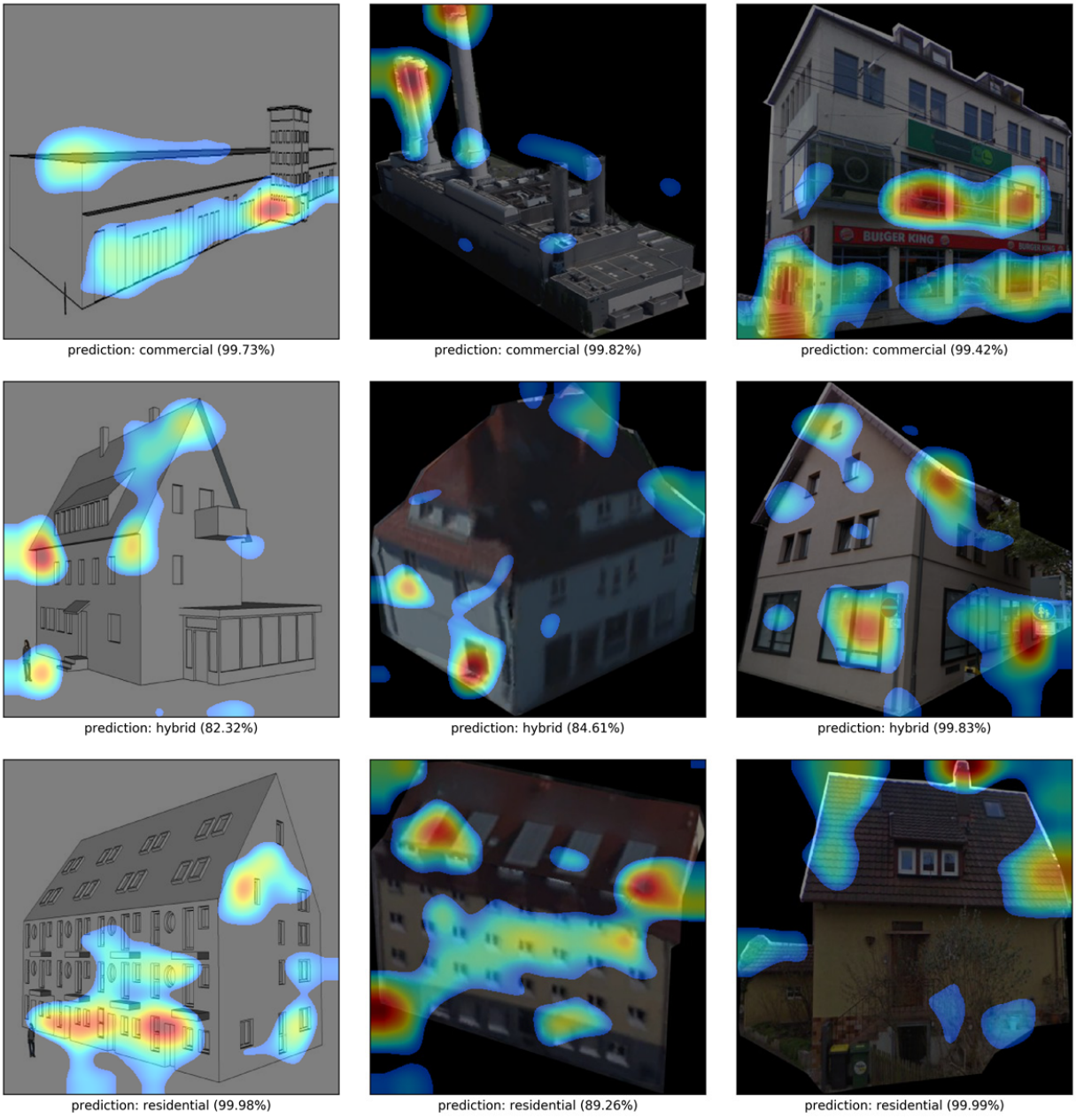


Fig. 5: Snapshots of different non-image representations that have been classified correctly overlaid by corresponding CAMs (ground truth in first row: *commercial*, ground truth in second row: *hybrid*, ground truth in last row: *residential*). The first column shows snapshots of LoD3 models. The second column shows snapshots of textured meshes (textured with Google Earth images). The last column shows cut out Google Street View images

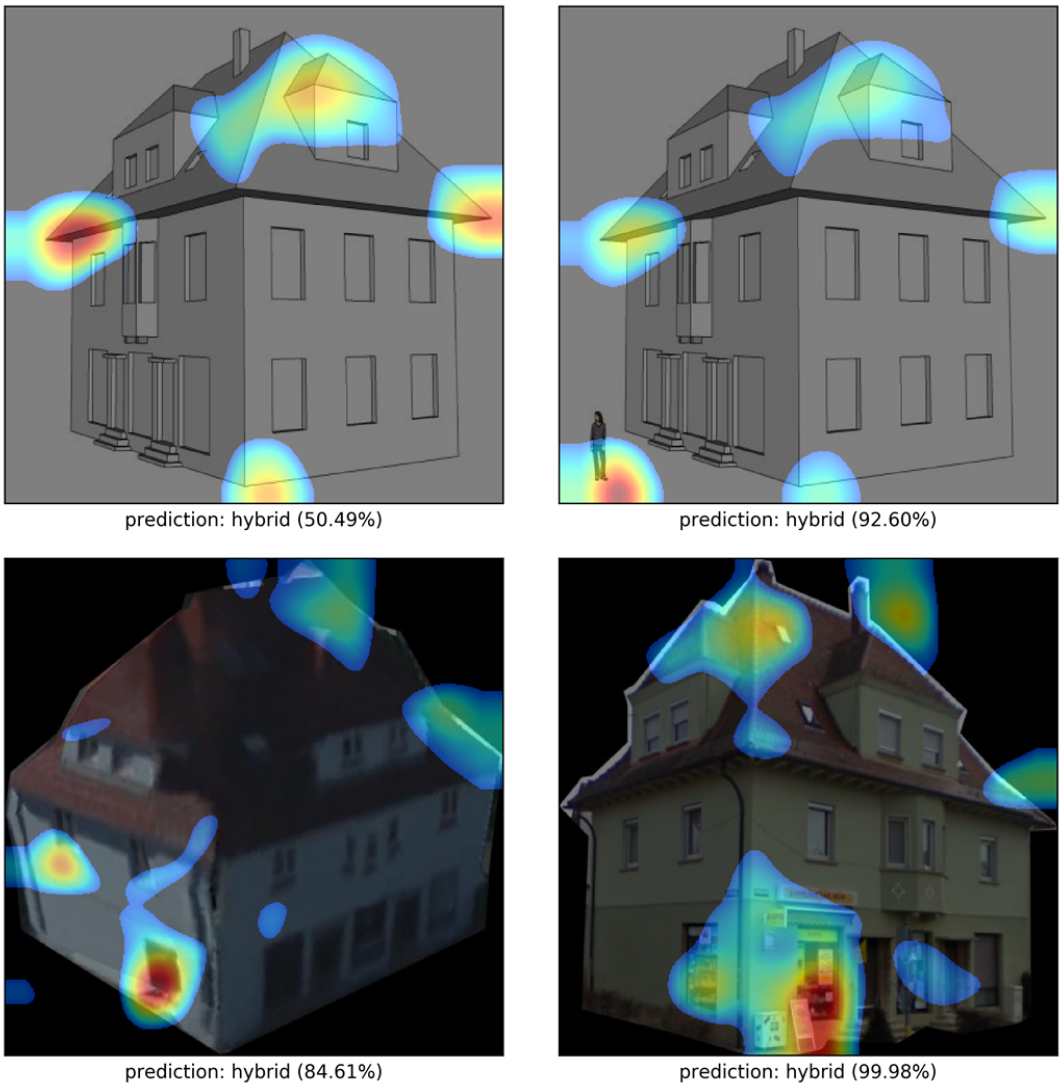


Fig. 6: Snapshots of different non-image representations of the same building and the corresponding CAMs. Please note, the building is oriented differently in the snapshots. The ground truth label is class *hybrid*. The upper row shows a snapshot of the LoD3 model without a human (*left*) and one snapshot with a human (*right*). The second row shows a snapshot of the textured mesh (textured with Google Earth images) and a cut out Google Street View image (*from left to right*).

3 Conclusions and Outlook

The presented end-to-end approach achieves promising results for enriching 3D urban models automatically with semantic information. The overall error of 36% reflects the complexity of this real-world classification task and immediately suggests to reduce the classification error in future work. This can be achieved by gathering more training data. Avoiding “false” predictions might be insoluble in the context of image classification (see (c). In future work we suggest to go from image classification to object detection and object classification where this kind of misclassifications do not apply. Alternatively, a multi-label data set should be provided. In such a

data set, every image is tagged with all classes whose instances are visible in the image - sorted by covering area in descending order.
We have to note that reality is far too complex in order to label all images properly into the considered classes. It would be interesting to see how and in how many classes images are clustered by an unsupervised learning method. Based on these results one could rethink the current classes.
In order to form a judgment of the CNNs' performance, we want to conduct a user study on our data sets and determine human performance for comparison. We know there are images where humans have problems to classify the facades correctly.

4 Acknowledgment

The author thanks PATRICK TUTZAUER for his support during the period of the master thesis. Furthermore, we acknowledge the support of NVIDIA Corporation by donation of a Titan X Pascal GPU used to generate the presented results.

5 Bibliography

HE, K., ZHANG, X., REN, S. & SUN, J., 2016: Deep Residual Learning for Image Recognition. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.

LECUN, Y., BOTTOU, L., ORR, G.B. & MÜLLER, K.-R., 1998: Efficient BackProp. Neural Networks: Tricks of the Trade, 9-50.

MARTINOVIC, A., KNOPP, J., RIEMENSCHNEIDER, H. & VAN GOOL, L., 2015: 3D All The Way: Semantic Segmentation of Urban Scenes From Start to End in 3D. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4456-4465.

MOVSHOVITZ-ATTIAS, Y, YU, Q., STUMPE, M. C., SHET, V., ARNOUD, S. & YATZIV, L., 2015: Ontological Supervision for Fine Grained Classification of Street View Storefronts. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1693-1702.

RUSSAKOVSKY, O., DENG, J., SU, H., KRAUSE, J., SATHEESH, S., MA, S., HUANG, Z., KARPATHY, A., KHOSLA, A., BERNSTEIN, M. S., BERG, A. C. & LI, F., 2015: ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision **115**, 211-252.

SELVARAJU, R. R., COGSWELL, M., DAS, A., VEDANTAM, R., PARIKH, D. & BATRA, D., 2016: Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. See https://arxiv. org/abs/1610.02391 v3.

SIMONYAN, K. & ZISSERMAN, A., 2014: Very Deep Convolutional Networks for Large-Scale Image Recognition. Computing Research Repository (CoRR abs/1409.1556).

SZEGEDY, C., VANHOUCKE, V., IOFFE, S., SHLENS, J. & WOJNA, Z., 2016: Rethinking the Inception Architecture for Computer Vision. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2818-2826.

TUTZAUER, P. & HAALA, N., 2017: Processing of Crawled Urban Imagery for Building Use. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-1/W1, 143-149.

Untersuchung zur Nutzung der Radarfernerkundung zur Klassifizierung von Feldfrüchten

LISA-MARIE KRALISCH[1], HOLGER LILIENTHAL[2] & MICHAEL BREUER[3]

Zusammenfassung: Mit Sentinel-1A Radarbilddaten wurde untersucht, ob ein reproduzierbarer Zusammenhang zwischen Feldfrüchten und Rückstreuungskoeffizienten für die Klassifikation von Feldfrüchten gefunden werden kann. Verwendet wurde das Ground Range korrigierte Datenprodukt (GRD). Mit Hilfe von Referenzdaten wurden für jede Feldfruchtart Mittelwert und Standardabweichung berechnet. Die sich so ergebenden Wertebereiche wurden für eine überwachte Klassifizierung mit einem Entscheidungsbaum verwendet. Diese lieferte eine Gesamtgenauigkeit von ca. 57 %. Im Anschluss wird die Anwendbarkeit der Wertebereiche in Hinblick auf Lage und Zeit überprüft. Das Ergebnis zeigte, dass sich der Rückstreuungskoeffizient mit dem Blickwinkel stark verändert. Außerdem ist die Verwendung von Daten, die drei Tage später aufgenommen wurden, problematisch, was die Notwendigkeit einer Korrektur der Richtungsabhängigkeit unterstreicht. Die Untersuchung dient somit als erste Grundlage für die Feldfruchtklassifizierung mit Radardaten.

1 Einleitung

Die Beobachtung und Untersuchung der Vegetation unserer Erde ist, unter anderem im Zusammenhang mit dem globalen Klimawandel, ein wichtiges und aktuelles Thema. Außerdem ist die Beobachtung des landwirtschaftlichen Fruchtartenspektrums für verschiedene Interessengruppen relevant. Dabei geht es neben der globalen Sicherung der Nahrungsmittelproduktion und einem besseren Verständnis der damit verbundenen Umweltprozesse auch um die Veränderung von Anbaukulturen und Strukturen, die sich aus Marktsituationen und politischen Entscheidungen ergeben.

Die Fernerkundung wurde hierfür schon lange als Schlüsselsystem erkannt und wird entsprechend genutzt. Die Radarfernerkundung kann durch den starken Zusammenhang zwischen Radarrückstreuung und Vegetationsgeometrie wichtige Informationen liefern. Aktive Radarsysteme besitzen den großen Vorteil, licht- und nahezu witterungsunabhängig Daten erfassen zu können, da ihr Radarsignal Wolken durchdringen kann.

Nach FERAZZOLLI (2002) sind Radarsysteme für die Feldfruchtklassifizierung geeignet, wenn sie in den L- oder C-Bändern operieren, Co- und Kreuzpolarisation anbieten und einen Blickwinkel von 30 Grad bis 40 Grad aufweisen. LILLESAND et al. (2015) geben an, dass sich X- und C-Band am besten für die Untersuchung von Feldfrüchten eignen. Die neuen Radarsatelliten Sentinel-1A und -1B des Europäischen Copernicus-Programms tragen ein C-Band SAR-System (ESA 2013)

[1] con terra GmbH, Martin-Luther-King-Weg 24,D-48155 Münster, E-Mail: kralisch.lisa@gmail.com
[2] Julius-Kühn-Institut (JKI), Bundesforschungsinstitut für Kulturpflanzen, Bundesallee 69, D-38116 Braunschweig, E-Mail: holger.lilienthal@julius-kuehn.de
[3] Beuth Hochschule für Technik Berlin, Labor für Photogrammetrie, Luxemburger Strasse 10, D-13353 Berlin, E-Mail: breuer@beuth-hochschule.de

und weisen die geforderten Merkmale auf. Das System ist als Konstellation aus zwei identischen Satelliten (Sentinel-1A und -1B) ausgelegt, deren Orbits um 180° versetzt angeordnet sind. Jeder Satellit umkreist die Erde jeweils in zwölf Tagen, so dass jeder Ort auf der Erde mindestens alle sechs Tage von einem Satelliten abgetastet wird (ESA 2013). Daraus folgt eine sehr hohe Datendichte, die bisher noch nie verfügbar war. Daraus ergeben sich völlig neue Möglichkeiten der Datenauswertung und Zeitreihenanalyse.

2 Problemstellung

Das Ziel der Arbeit ist es, auf Grundlage von Sentinel-1-Daten den Zusammenhang zwischen Feldfrüchten und Rückstreuungskoeffizient zu untersuchen, um eine Feldfruchtklassifizierung zur Unterscheidung von ausgewählten Feldfruchtarten durchzuführen. Für die Durchführungen von Nutzungsklassifizierungen gibt es eine Vielzahl von Methoden und Ansätzen: zwei häufig verwendete Methoden sind a) der dreidimensionale H/A/α Klassifizierungsraum und b) der Complex Wishart Classifier (LEE & POTTIER 2009). Der dreidimensionale H/A/α Klassifizierungsraum ist eine unüberwachte Klassifizierung und gibt lediglich Auskunft darüber, ob Vegetation existiert. Der Complex Wishart Classifier benötigt komplexe Radardaten, deren Datenvolumen bei temporal hochauflösenden Daten schwer zu handhaben ist. Entscheidungsbäume, die auf der Bildstatistik entwickelt werden, zeigten in den letzten Jahren vielversprechende Ergebnisse bei der Klassifizierung (MOGHADDAM 2009).

3 Versuch

Mit Hilfe eines Entscheidungsbaumes wurde versucht, Regeln und statistische Größen zu finden, nach denen einzelne Feldfrüchte klassifiziert werden können. Dazu wurde ein Testgebiet in Brandenburg, südöstlich von Berlin ausgewählt (siehe Abb. 1). Die Auswahlkriterien waren die geringe Reliefenergie, sodass der Einfluss der Topographie als konstant angesehen werden kann und dass alle ausgewählten Feldfrüchte in dem Gebiet angebaut worden sind. Weiterhin wurden dadurch die Rechenzeiten stark verkürzt, weil die Prozessierung nur auf den gewählten Ausschnitten (Subsets) durchgeführt wurde.

Abb. 1: Testgebiet zur Fruchtartenklassifikation in Brandenburg (rote Markierung). Kartenhintergrund: Quelle Google Earth

3.1 Datengrundlage

Als Bilddaten standen Sentinel-1A Ground-Range korrigierte Daten (GRD) aus dem Jahr 2015 zur Verfügung. Die Daten lagen in den Polarisationen VV und VH vor und wurden im IW-High-Resolution-Modus aufgenommen. Nach der radiometrischen Kalibrierung und Topographiekorrektur lagen die Bilddaten als Sigma Naught Daten in einer geometrischen Auflösung von 10m × 10m vor. Für die Auswertung wurden acht Aufnahmen vom Descending Orbit 95 und neun Aufnahmen vom Ascending Orbit 146 im Zeitraum von Mai bis August genutzt. Die Informationen liegen pro Pixel als normalisierte Rückstreuungskoeffizienten vor. Dieser Koeffizient repräsentiert die Stärke des zurückstreuten Radarsignals normalisiert über die bestrahlte Fläche (WOODHOUSE 2006:108).

Die Referenzdaten zur angebauten Feldfruchtart und den Feldgeometrien stammen aus dem Integrierten Verwaltungs- und Kontrollsystem (InVekos) des Landes Brandenburg. Für die Untersuchung werden acht Hauptkulturen ausgewählt. Dazu gehören Kartoffeln, Mais, Wintergerste, Wintertriticale, Winterraps, Winterroggen, Winterweizen und Zuckerrüben.

3.2 Ausführung

Als Erstes wurde eine unüberwachte Klassifizierung der acht Aufnahmen mit dem ISODATA-Algorithmus (BALL & HALL 1967) durchgeführt, um einen groben Überblick der Daten zu erhalten. Der Algorithmus ermöglicht es, zusammengehörige Agrarflächen allein auf Basis der Information des normierten Rückstreuungskoeffizienten verschiedenen Klassen zuzuordnen (siehe Abb. 2). Hintergrund ist die Annahme, dass jede Feldfrucht ähnliche Rückstreuungswerte aufweist, bzw. sich die Rückstreuungswerte zeitlich für jede Fruchtart spezifisch verändern. Der Koeffizient muss eindeutige Cluster im Merkmalsraum bilden.

Abb. 2: Ergebnis einer unüberwachten Klassifikation von Feldern nahe Rüdersdorf in Brandenburg aus dem Jahr 2015.

Mit einem Entscheidungsbaum sollen Regeln genutzt werden, nach denen die einzelnen Feldfrüchte klassifiziert werden können. Zunächst wurden für jedes Feld deskriptive statistische Größen wie Mittelwert, Standardabweichung und Spannweite (Range) bestimmt. Diese statistischen Werte eignen sich in der Regel besonders gut, um komplexe Verhältnisse zu beschreiben und große Datenmengen zu analysieren. Es wurde angenommen, dass sich jede Feldfrucht über den Mittelwert und eine Standardabweichung (68,28% der Daten) charakterisieren lassen.

Durch die vorhandenen Referenzdaten können alle Pixel innerhalb eines Feldes mit derselben Feldfruchtart ausgewählt und deren zugehöriger Mittelwert sowie die Standardabweichung berechnet werden.

Um eine Übersicht darüber zu erhalten, inwieweit sich die Spannweiten der einzelnen Feldfürchte überlagern, wurden für jeden Zeitpunkt zwei Diagramme (für jede Polarisation) erstellt. Abbildung 3 zeigt zwei Beispiele vom 4. Mai 2015. Die Ordinate bildet den normalisierten Rückstreuungskoeffizienten ab. Auf der Abszisse sind die einzelnen Feldfrüchte eingetragen. Die großen mittleren Punkte zeigen den Mittelwert. Die kleineren Punkte zeigen die Standardabweichung und somit die jeweiligen Spannweiten auf. Anhand dieser Graphiken ist erkennbar, ob und wie gut sich bestimmte Feldfrüchte zum jeweiligen Termin unterscheiden lassen.

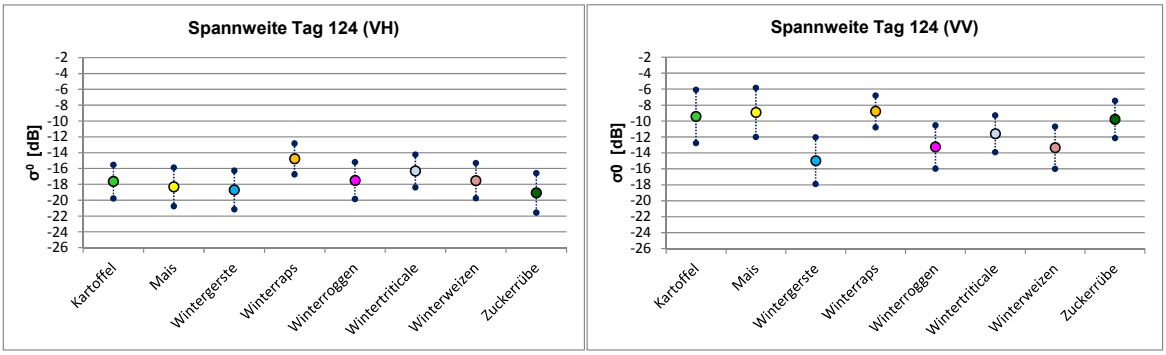


Abb. 3: Unterscheidbarkeit von Feldfrüchten am 4. Mai 2015 (Tag 124) in den Rückstreuungskoeffizienten der VH- und VV-Polarisation.

Nachdem alle Spannweiten vorliegen, wurde ein Entscheidungsbaum erstellt (siehe Abb.4). Er ist einstufig aufgebaut und jede Hypothese besitzt dieselben Äste und Variablen. Der Unterschied besteht in den Werten der Bedingungen. Ein Pixel wird dann klassifiziert, wenn dessen Wert alle Bedingungen einer Hypothese erfüllt.

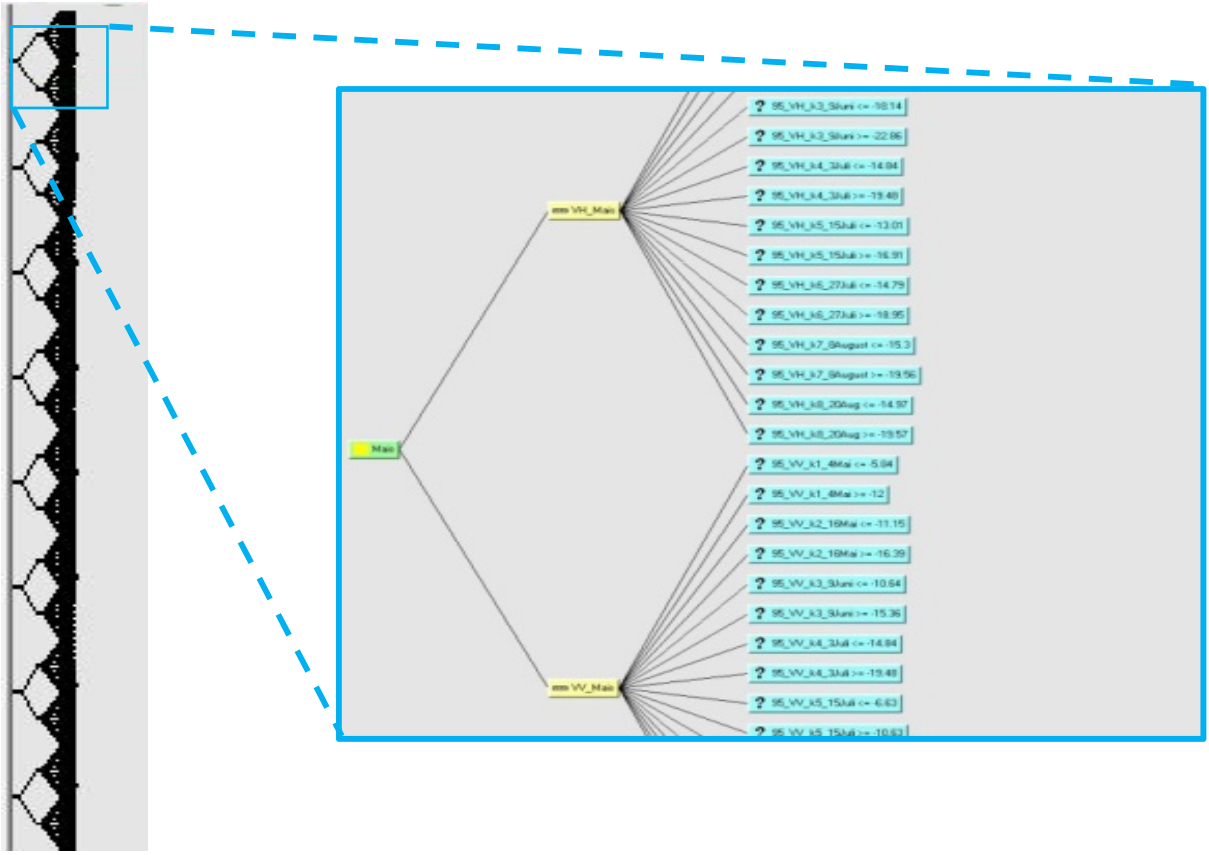


Abb. 4: Entscheidungsbaum mit allen eingetragenen Bedingungen

Radardaten weisen aufgrund von Interferenzphänomenen sogenannte Speckle Effekte auf, die wie ein Rauschen die Daten überlagern. Um diesen Effekt etwas abzumildern wurden die Bilddaten mit einem Tiefpassfilter (9×9 Pixel) bearbeitet.

3.3 Ergebnisdiskussion

Die Fruchtartenklassifikation mit einem Entscheidungsbaum auf der Basis von gefilterten Radardaten konnte eine Gesamtgenauigkeit von ca. 57% erreichen. Die Grenze einer akzeptablen Klassifikation liegt nach CONGALTON & GREEN (2009) bei 85%. Somit wäre das Ergebnis an diesem Maß gescheitert. Abbildung 5 zeigt die Konfusionsmatrix der Klassifikation.

Subset 1 (D95) Original		Referenz									
		Undefiniert	Mais	Wintergerste	Winterweizen	Wintertriticale	Winterraps	Winterroggen	Zuckerrübe	Kartoffeln	Summe
Klassifizierung	Undefiniert	16875508	1152430	695298	1223737	313696	1444659	1341731	93709	10022	23150790
	Mais	42170	96312	1118	3837	883	221	3820	719	173	149253
	Wintergerste	23682	823	55330	5114	1039	329	2941	23	15	89296
	Winterweizen	88421	4219	5755	105266	13182	395	29013	101	24	246376
	Wintertriticale	63507	4690	4882	36824	22350	577	53695	49	33	186607
	Winterraps	30445	1392	610	614	189	109537	594	42	5	143428
	Winterroggen	51944	5409	6686	19199	12605	539	79947	48	18	176395
	Zuckerrübe	4547	6552	135	97	13	55	152	7355	158	19064
	Kartoffeln	25730	27478	1436	1287	598	207	2450	739	590	60515
	Summe	17205954	1299305	771250	1395975	364555	1556519	1514343	102785	11038	24221724

Abb. 5: Konfusionsmatrix der Klassifikation

Die Arbeit versuchte in einem ersten einfachen Ansatz, eine Fruchtartenklassifikation anhand von Entscheidungsbäumen zu automatisieren. Die fruchtartenspezifische Phänologie konnte noch nicht ausreichend berücksichtigt werden, zudem war die Datenlage 2015 noch recht fehlerbehaftet und es gab lediglich zwei Aufnahmen pro Monat. Seit 2016 arbeiten beide Satelliten operationell, so dass jetzt fünf Aufnahmen pro Monat verfügbar sind. Durch diesen Umstand sollten sich jetzt bessere Ergebnisse erzielen lassen.
Klassifizierungen auf Basis der normalisierten Rückstreuungskoeffizienten scheinen grundsätzlich möglich zu sein. Da die Rückstreuungssignale mit dem Einfallswinkel variieren, muss eine Auswertung separat pro Orbit erfolgen. Eine mögliche Lösung wäre es, die Orbits in einheitliche Einfallswinkelabschnitte einzuteilen und diese mit entsprechenden Entscheidungsbäumen zu klassifizieren. Eine andere Lösung wäre die Nutzung richtungskorrigierter Koeffizienten.
Die Nutzung von Entscheidungsbäumen stellt einen Baustein in einer ganzen Reihe von Klassifikationsmethoden dar. Die geschickte Kombination mit landwirtschaftlichem Hintergrundwissen, Plausibilitätsannahmen und weiteren Machine-Learning-Verfahren bieten eine Vielzahl neuer Möglichkeiten, die separat getestet werden müssen.

4 Fazit und Ausblick

Die Ausgangsfrage der Untersuchung lautet, ob sich Radardaten zur Feldfruchtklassifizierung eignen. Diese Frage kann definitiv positiv beantwortet werden. Studien von LEE & PORTIER (2009) und FERRAZZOLI (2002) zeigen, dass sich Radardaten zur Klassifizierung von Vegetation grundsätzlich eignen.
Die Sentinel-1 Satelliten bieten eine hohe zeitliche Auflösung, die Zeitreihenanalysen überhaupt erst möglich machen. Leider nehmen die C-SAR-Systeme der Sentinel-1 Satelliten nicht vollpolarimetrisch auf, was die Auswertungsmöglichkeiten limitiert. Die besten Ergebnisse werden mit multifrequenten Daten erreicht. Derzeit sind kostengünstige multifrequente Satellitensysteme noch nicht verfügbar.
Radardaten können trotz des enormen Potenzials den Informationsgehalt optischer Daten momentan noch nicht ersetzen, sie können diese aber unterstützen. Die Kombination und Synergien beider Technologien werden zu besseren Ergebnissen führen.
Die Ergebnisse dieser Arbeit zeigen, dass gleiche Feldfrüchte ähnliche Rückstreuungen aufweisen, dass die normierten Rückstreuungskoeffizienten allein aber keine zuverlässige Basis für eine Klassifizierung sein können. Das Verfahren kann jedoch noch verbessert werden. Es

könnte für jede Feldfruchtart ein heuristisches Streuungsmodell erstellt werden, um das Verhalten der Welle in den verschiedenen Geometrien berechenbar zu machen (ULABY et al. 2014). Zusätzlich zu diesen Modellen könnte auf das Verhalten der einzelnen Feldfruchtarten im Zusammenhang mit verschiedenen Parametern des Aufnahmesystems der Sentinel-Satelliten eingegangen werden.
Außerdem sollte der Richtungsabhängigkeit Aufmerksamkeit geschenkt werden. Die Frage ist, ob die Klassifikationsgenauigkeit mit einer richtungskorrigierten Größe (z.B. Gamma) anstelle des normierten Rückstreuungskoeffizienten (Sigma Naught) noch gesteigert werden kann.
Die Nutzung von Wetterdaten (räumlichen Niederschlagsverteilungen) hilft einige Phänomene innerhalb der Radardaten zu erklären. Bestandesfeuchtigkeit, Tau und Niederschläge beeinflussen die Radarrückstreuung. Bei der Erstellung der Entscheidungsbäume und bei der Klassifikation der Radardaten sollten derartige Datensätze ausgenommen werden. Als Hintergrundwissen sind Informationen zur Aussaat, Phänologie und Ernte der einzelnen Feldfrüchte entscheidend, um sie sicher trennen zu können. Das Wachstum der Feldfrüchte ist unter anderem von Temperatur und Niederschlag abhängig. Durch einfache Modelle könnte die Phänologie der einzelnen Kulturen besser geschätzt werden, um dann die geeigneten Entscheidungsbäume auszuwählen. Die Radarfernerkundung bietet hier noch viel Forschungspotenzial. Die Kombinationsmöglichkeiten der vielen verfügbaren Parameter und der Eigenschaften der Radardaten sollten in Zukunft weiter erforscht werden.

5 Literaturverzeichnis

BALL, G. & HALL, D., 1967: A clustering technique for summarizing multivariate data. Behavioral Science 12, 153-155.

CONGALTON, R.G. & GREEN, K., 2009: Assessing the Accuracy of Remotely Sensed Data. 2. Auflage, CRC Press, Taylor & Francis Group, Boca Raton.

ESA, 2013: Sentinel-1 User Handbook. https://sentinel.esa.int/documents/247904/685163/Sentinel-1_User_Handbook, letzter Zugriff 10.08.2016.

FERRAZZOLI, P., 2002: SAR for agriculture: Advances, problems and prospects. Proceedings of the Third International Symposium, 'Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications', 47-56.

LEE, J.-S. & POTTIER, E., 2009: Polarimetric Radar Imaging From Basics To Applications. CRC Press, Taylor & Francis Group, Boca Raton.

LILLESAND, T.M., KIEFER, R.W. & CHIPMAN, J.W., 2015: Remote Sensing and Image Interpretation, 7. Ausgabe, Wiley, Hoboken.

MOGHADDAM, M., 2009: Polarimetric SAR Phenomenology and Inversion Techniques for Vegetated Terrain. In: The SAGE Handbook of Remote Sensing 2009, Warner, T. A.; Nellis, M. D. & Foody, G. M. (Hrsg.), SAGE Publications Ltd., London, 79-92.

ULABY, F.T., LONG, D.G., BLACKWELL, W., ELACHI, C., FUNG, A., RUF, C., SARABANDI, K., ZEBKER, H. & VAN ZYL, J., 2014: Microwave Radar and Radiometric Remote Sensing. Ann Harbor.

WOODHOUSE, I.H., 2006: Introduction to Microwave Remote Sensing. CRC Press, Taylor & Francis Group. Boca Raton.

Detaillierte Repräsentation des Straßenraums in 3D Stadtmodellen

CHRISTOF BEIL[1] & THOMAS H. KOLBE[1]

Zusammenfassung: Dreidimensionale semantische Stadtmodelle werden immer häufiger für Simulationen und Analysen urbaner Räume verwendet. Bislang liegt der Fokus dabei jedoch stark auf Gebäudemodellen. Für zahlreiche Anwendungsfälle wären allerdings auch 3D Modelle des Straßenraums von großem Nutzen, jedoch gibt es hierfür nur wenige Richtlinien. Relevante Standards wie GDF oder OpenDRIVE werden daher analysiert und dabei wird festgestellt, dass sich diese in erster Linie auf linienhafte oder parametrische Repräsentationen konzentrieren. Daraufhin werden 16 potentielle Anwendungen vorgestellt und der Nutzen, den diese aus einer flächenhaften 3D-Straßenraummodellierung ziehen würden, diskutiert. Auf Grundlage dieser Erkenntnisse wird das Transportation Datenmodell des OGC Standards CityGML2.0 hinsichtlich Verbesserungs- und Erweiterungs-möglichkeiten untersucht. Im Zuge dessen wird ein CityGML konformes Modellierungskonzept für mehrere Detailstufen (LoD) vorgestellt. Abschließend wird dieses Konzept auf Grundlage frei verfügbarer Daten (Open Data) für ein semantisches 3D-Stadtmodell von New York City implementiert. Als Ergebnis werden so über 500.000 Straßenraumobjekte vollautomatisch erzeugt. Diese umfassen u.a. Straßenflächen, erhöhte Gehwege, Parkplatzflächen und Verkehrsinseln, welche allesamt mit einer Vielzahl thematischer Attribute angereichert sowie texturiert sind. Das Modell wird zudem mit Hilfe des Open Source „WebGL Virtual Globe Cesium" visualisiert und so für interaktive Attributabfragen und Analysen zugänglich gemacht.

1 Motivation und Problemstellung

Im Rahmen mehrerer Projekte und Masterarbeiten am Lehrstuhl für Geoinformatik der Technischen Universität München wurde ein semantisches Stadtmodell für New York City im Level of Detail 1 (LoD1) erstellt. Hauptaugenmerk lag dabei im ersten Schritt auf der semantischen und weniger auf der geometrischen Genauigkeit des Modells. Dieser Artikel beschreibt die Untersuchungen und erzielten Ergebnisse der Masterarbeit des ersten Autors, die sich aus einem konzeptionellen und einem praktischen Teil zusammensetzten. Zunächst sollen Möglichkeiten zur detaillierten Repräsentation und Modellierung des Straßenraums erarbeitet werden. Hierzu sind verwandte Arbeiten und Standards zu untersuchen und zu diskutieren. Anschließend sollen Anwendungsfelder eruiert werden, die von detaillierten Straßenraummodellen profitieren würden. Dabei soll auch auf Anforderungen an die Art der Modellierung eigegangen werden. Konzeptionelle Überlegungen, z.B. wie der Straßenraum sinnvoll in Einzelobjekte zerteilt oder wie mit Kreuzungsbereichen umgegangen werden könnte, sollen daraufhin zu Verbesserungsvorschlägen des für 3D-Stadtmodelle verwendeten Standards CityGML führen. Abschließend sollen ausgewählte Teile des Konzepts anhand des erwähnten Stadtmodells von New York City auf Grundla-

[1] Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstraße 21, D-80333 München, E-Mail: christof.beil@tum.de

ge frei verfügbarer Daten möglichst automatisiert umgesetzt werden. Ergebnis soll ein semantisch sowie geometrisch stark verbessertes Straßen- und Stadtmodell sein.

2 Methodik

Zunächst wird der OGC Standard CityGML sowie verwandte Standards, welche im Bereich der Straßenmodellierung relevant sind, vorgestellt und untersucht. Im Anschluss daran werden insgesamt 16 potentielle Anwendungsfelder, die von detaillierten Straßenraummodellen profitieren würden, erarbeitet. Im Zuge dessen werden auch die jeweiligen Anforderungen der einzelnen Anwendungen an solche Modelle analysiert und diskutiert sowie hierfür benötigte Objektklassen und Attribute benannt. Auf diesen Erkenntnissen aufbauend wird dann ein konzeptionelles Modell zur Straßenmodellierung erstellt. Dabei wird das aktuell gültige Transportation Datenmodell des CityGML2.0-Standards hinsichtlich möglicher Defizite untersucht sowie Vorschläge für dessen Erweiterung und Verbesserung gemacht. Dieses Konzept wird abschließend auf Grundlage frei verfügbarer Daten für das gesamte Stadtgebiet von New York City automatisiert umgesetzt.

2.1 Untersuchung verwandter Arbeiten, Standards und Richtlinien

Zu Beginn werden einige Standards und Richtlinien, die sich mit der Straßenraummodellierung befassen, vorgestellt und analysiert. Es zeigt sich dabei, dass Straßen in den meisten Standards durch linienhafte oder parametrische Repräsentationen dargestellt werden. Zudem wird das Transportation Modell von CityGML vorgestellt. Die teilweise sehr umfangreichen Vorgaben der verwandten Standards enthalten zahlreiche Konzepte und Ideen, die in der Folge als Grundlage für eigene Weiterentwicklungen des CityGML Transportation Modells dienen. Folgende Standards und Richtlinien werden ausführlich untersucht: **City Geography Markup Language (CityGML)** ist ein XML-basiertes Format zur Repräsentation, Speicherung und zum Austausch semantischer 3D-Stadt- und Landschaftsmodelle. Der derzeit gültige Standard CityGML2.0 wurde im Jahr 2012 vom Open Geospatial Consortium (OGC) verabschiedet. Darin werden zahlreiche Klassen und Beziehungen unterschiedlicher thematischer Objekte im Hinblick auf deren räumlichen, semantischen und visuellen Eigenschaften in insgesamt 5 Detailgraden (LoD) definiert (GRÖGER et al. 2012). Zur Modellierung des Straßenraums steht das sogenannte Transportation Datenmodell zur Verfügung, welches den Straßenraum in LoD0 durch Liniennetze und ab LoD1 durch flächenhafte Repräsentationen modelliert. Es stehen dabei vier thematische Oberklassen (Road, Square, Track und Railway) zur Verfügung. Ab LoD2 kann der Straßenraum zudem in thematische Unterklassen (TrafficArea und AuxiliaryTrafficArea) feiner untergliedert werden. **Geographic Data Files (GDF)** ist ein Standard, der vor allem im Bereich der Fahrzeugnavigation weit verbreitet ist. Objekte werden darin konzeptionell in drei Detailgrade (Level-0, Level-1 und Level-2) unterteilt. Der Straßenraum kann dabei durch ein sogenanntes „Road Network“ repräsentiert werden (ISO 14825, 2011). Das Bundesministerium für Verkehr, Bau- und Wohnungswesen stellt Richtlinien zur einheitlichen Aufnahme, Bearbeitung und Vorhaltung der Merkmale von Straßen bereit. Hierin wird der Straßenraum durch die **Anweisung Straßeninformationsbank (ASB)** aus technischer und durch den **Objektkatalog für das Straßen- und Verkehrswesen (OKSTRA)** aus fachlicher Sicht beschrieben (BAST 2016; BMVI 2014). **OpenDRIVE** wird als offenes Datenformat häufig für Anwendungen im Bereich von Fahrsimula-

tionen verwendet. Straßen werden darin durch eine Referenzlinie repräsentiert, welche sich aus mehreren Abschnitten geometrischer Primitive, wie Linien, Kurven oder Klothoiden, zusammensetzt. Relativ zu dieser Referenzlinie werden Straßeneigenschaften, wie z.B. Fahrspuranzahl parametrisch gespeichert (DUPUIS et al. 2015). Eine ausführliche Diskussion kann in BEIL (2017) nachgelesen werden.

2.2 Anwendungsperspektiven für Straßenraummodelle

Im folgenden Abschnitt wird untersucht, welchen Nutzen eine detaillierte Repräsentation des Straßenraums für unterschiedliche Anwendungen mit sich bringen würde und welche Anforderungen diese an die Art der Modellierung stellen. Insgesamt werden 16 Anwendungsfelder betrachtet, welche von detaillierten Modellen des Straßenraums und insbesondere flächenhaften 3D-Repräsentationen profitieren würden. Jede Anwendung und der jeweilige Mehrwert, den diese aus Straßenmodellen ziehen würden, wird unter Angabe einschlägiger Quellen vorgestellt. Die beschriebenen Anwendungen stellen Anforderungen an die Art der Modellierung, welche sich in einzelne Kategorien zusammenfassen lassen. Als Anforderungen werden Lagegenauigkeit, thematische Genauigkeit, Aktualität, zusätzliche Attribute, Zeit als 4. Dimension sowie eine ansprechende Visualisierung definiert und deren jeweilige Wichtigkeit mit Hilfe der in Tabelle 1 dargestellten Punktetabelle zusammengefasst. Weiterhin werden benötigte Objektklassen, Attribute und Geometrien recherchiert und diskutiert, indem diese benannt werden, untersucht wird, ob sie bereits im aktuellen CiytGML2.0 Standard enthalten sind und falls dies nicht der Fall ist, aufgezeigt wird, wie diese integriert werden könnten.

Tab. 1: Anforderungen an Straßenraummodelle für die jeweilige Anwendung (BEIL 2017)

Anwendung	Lagegenauigkeit	Thematische Genauigkeit	Aktualität	Zusätzl. Attribute	Zeit	Visualisierung
1) Flächennutzung	●●●●	○●●●	●●●●	○○●●	●●●●	○●●●
2) Hitzeeinwirkung	○●●●	○○○●	○●●●	●●●●	○○●●	○○○●
3) Kommunale Aufgaben	○●●●	●●●●	○●●●	●●●●	○○○●	○○○●
4) Lärmkartierung	○●●●	○○●●	○●●●	●●●●	○●●●	○○○●
5) Verkehrssimulationen	○●●●	○●●●	○●●●	○●●●	●●●●	○○●●
6) Fahrdynamik	○○○●	○●●●	○○●●	●●●●	○○●●	○○●●
7) Autonomes Fahren	●●●●	●●●●	●●●●	○●●●	○○●●	○○○●
8) Fahrtraining	○○○●	●●●●	○○○●	○●●●	○○●●	●●●●
9) Beschilderung	○●●●	○●●●	●●●●	○○●●	○○●●	○●●●
10) Instandhaltung	○●●●	○●●●	●●●●	●●●●	○○●●	○○●●
11) Belastbarkeit	●●●●	○●●●	●●●●	●●●●	○○●●	○○●●
12) Navigation	●●●●	●●●●	●●●●	●●●●	○○●●	○○●●
13) Notfallplanung	●●●●	○●●●	●●●●	●●●●	○○●●	○○●●
14) Feinstaubbelastung	○●●●	○○●●	○●●●	●●●●	●●●●	○○●●
15) Aquaplaning	○●●●	○●●●	○●●●	●●●●	○●●●	○●●●
16) Lichtraumprofile	●●●●	○○●●	●●●●	○●●●	○○●●	○○●●

Legende: ●●●● sehr wichtig, ○●●● wichtig, ○○●● neutral, ○○○● unwichtig

Digitale 3D-Stadtmodelle können als Grundlage für Stadtplanungen und Landmanagement dienen (ROSS 2010). Neben Gebäuden wird der urbane Raum hauptsächlich durch Straßen und Verkehrswege geprägt. Digitale Simulationen unterschiedlicher Szenarien können dabei vor allem im Bereich von Straßenplanungen zur Aufwands-, Kosten- und Nutzenanalyse verwendet werden. In diesem Zusammenhang können auch einzelne Großbauprojekte wie Brücken oder Autobahnen digital konzipiert werden. Die 3D-Visualisierung solcher Projekte kann zur frühzeitigen

Bürgerbeteiligung genutzt und so helfen, sachlichere Auseinandersetzungen zu führen. Für Kommunen besteht die gemeindliche Beleuchtungs-, Reinigungs-, Räum- und Streupflicht. Detaillierte Straßenraummodelle können Gemeinden bei der kostengünstigen und effektiven Wahrnehmung dieser Pflichten unterstützen. Dabei stellt auch die Schadenskartierung und Instandhaltung von Straßen einen beachtlichen Kostenpunkt dar. Straßenflächen eines digitalen Modells können mit weiteren Informationen, zum Beispiel über den Straßenzustand, kombiniert werden und so Aussagen über zu erwartende Reparaturkosten getroffen werden. Weiterhin können vielfältige räumliche Analysen wie etwa Solarpotential- oder Sichtbarkeitsanalysen durchgeführt werden. So können zum Beispiel lokale Hitzeinseln oder optimale Schildstandorte ermittelt werden. Weitere Anwendungsgebiete kommen aus der Automobilindustrie. Die Kenntnis über den exakten Verlauf von Straßen sowie deren Fahrbahnränder ist interessant für den wachsenden Bereich des autonomen Fahrens. Informationen über Straßenlängen und Anzahl an Fahrbahnen kann zudem von Fahrassistenzsystemen bei der Unterstützung des Fahrers bei Überhohlvorgängen verwendet werden. RANDT et al. (2007) beschreiben wie virtuelle 3D-Landschaftsmodelle für Fahrsimulatoren und Training von Notfallsituationen für Einsatzkräfte dienen können. Ein weiteres Anwendungsfeld sind Fahrdynamiksimulationen bei der Fahrzeugentwicklung (Butz et al. 2004). Auf alle weiteren in Tabelle 1 aufgelisteten potentiellen Anwendungsfelder sowie deren jeweilige Anforderungen an die Art der Straßenraummodellierung wird in BEIL (2017) ausführlich eingegangen.

2.3 Konzeptionelle Modellierung des Straßenraums mit CityGML

Ein Konzept zur CityGML konformen Repräsentation des Straßenraums wird auf der Basis von vier Säulen erarbeitet. Diese sind der momentan gültige OGC Standard CityGML2.0, der vor allem im Bereich des Transportation Datenmodells eingehend untersucht wird, relevante Standards im Bereich der Straßenmodellierung, Anforderungen an die Art der Modellierung aus Anwendersicht sowie die zur späteren Umsetzung verfügbaren Daten. Im Zuge der Konzepterarbeitung werden u.a. zahlreiche realitätsnahe Szenarien, wie Einfahrten, Kreisverkehre, Kreuzungen, Überführungen oder Straßengabelungen schematisch modelliert und beispielsweise festgelegt, wie diese in CityGML LoD0, LoD1 oder LoD2-4 modelliert werden sollten. Hierzu macht der momentane Standard nur vage oder unzureichende Angaben. Diese Defizite werden aufgezeigt und Lösungen präsentiert.

2.3.1 Defizite des aktuellen Transportation Datenmodells in CityGML2.0

Das Transportation-Modell von CityGML stellt in erster Linie einen einheitlichen Rahmen zur konzeptionellen Modellierung des Straßenraums dar und soll bestehende Standards wie beispielsweise GDF ergänzen aber nicht ablösen. Dennoch lassen sich einige Schwächen identifizieren. Es ist unklar welche Achse einer Straße (Fahrbahn oder Fahrspur) zur Modellierung in LoD0 verwendet werden soll. Weiterhin werden linienhafte Gehwegs-Repräsentationen nicht berücksichtigt. Road Objekte können ab LoD1 flächenhaft durch *MultiSurface* Geometrien modelliert werden. Es ist jedoch nicht klar definiert, welche Objekte des Straßenraums dabei zu berücksichtigen sind. Ob beispielsweise Gehwege hierbei mit einzuschließen sind, wird nicht festgelegt. Ein gravierendes Problem des gültigen Standards ist die fehlende Definition von Segmenten oder Abschnitten, in die ein Straßennetzwerk unterteilt sein könnte. Bislang ist es

möglich alle Straßenraumobjekte einer ganzen Stadt lediglich einem einzigen *Road* Objekt zuzuweisen. Zudem werden bislang Kreuzungsbereiche oder Kreisverkehre nicht explizit repräsentiert. Dies führt zu Unklarheiten in Bereichen in denen sich mehrere Straßen denselben Kreuzungsbereich teilen. Geometrische Feinheiten wie Gullydeckel oder Schlaglöcher werden bislang nicht berücksichtigt. Weiterhin ist die Definition der Klasse *Square* ungenau.

2.3.2 Datenmodell des weiterentwickelten Transportation Datenmodells

Abbildung 1 zeigt das UML-Diagramm eines weiterentwickelten CityGML Transportation Modells. Die darin enthaltenen Neuerungen und Weiterentwicklungen werden in den folgenden Abschnitten erläutert.

2.3.3 Übergang zu vier Levels of Detail

LÖWNER et al. (2016) schlagen ein neues LoD Konzept für CityGML mit lediglich vier Detailgraden vor. Dieser Idee folgend wird vorgeschlagen auch das Transportation Modell auf LoD0 bis LoD3 zu beschränken. Der linke Teil von Abbildung 2 zeigt ein „reales“ Straßenstück mit unterschiedlichen thematischen Bestandteilen wie Fahrbahnen oder Gehwegen. In LoD1 sollte die gesamte Breite der Straße, inklusive der Gehwege, als eine Fläche repräsentiert werden. Eine detailliertere Aufteilung in *TrafficSpaces* und *AuxiliaryTrafficSpaces* wird ab LoD2 möglich. In LoD3 sollen zusätzlich geometrische Feinheiten wie Gullydeckel oder Schlaglöcher modelliert werden können. Diese Art der Repräsentation, mit feiner werdender thematischer Einteilung von Objekten in höheren LoD Stufen, ist sehr ähnlich zur Gebäudemodellierung in CityGML. Zusätzlich zu flächenhaften Repräsentationen zeigt der rechte Teil von Abbildung 2 welche Achsen für linienhafte Modellierungen verwendet werden sollten. Im Gegensatz zum bisherigen Standard sollen Linienrepräsentationen in allen LoDs möglich werden. Solange Fahrbahnen nicht mehrere Meter voneinander getrennt verlaufen, sollen diese in LoD0 und LoD1 durch eine Mittelachsenlinie modelliert werden. Ab LoD1 ist sowohl eine linien- als auch eine flächenhafte Repräsentation des Straßenraums möglich. Zusätzlich zur Straßenmittelachse können ab LoD2 auch Gehwege durch Linien modelliert werden. Dies ist speziell für Fußgängernavigationsanwendungen relevant. In LoD3 kann schlussendlich jede einzelne Fahrspur durch eine eigene Linie repräsentiert werden.

Neu ist im vorgestellten Datenmodell zudem die Möglichkeit auch Löcher in der Straßenoberfläche, also zum Beispiel Schlaglöcher (RoadwayDamage), Abflüsse (Drain) oder auch Gullydeckel (Manhole), explizit modellieren zu können. Weiterhin wird eine eigene Klasse für Lichtraumprofile (ClearanceSpace) ab LoD2 vorgeschlagen, welche durch Volumenköper modelliert werden sollen.

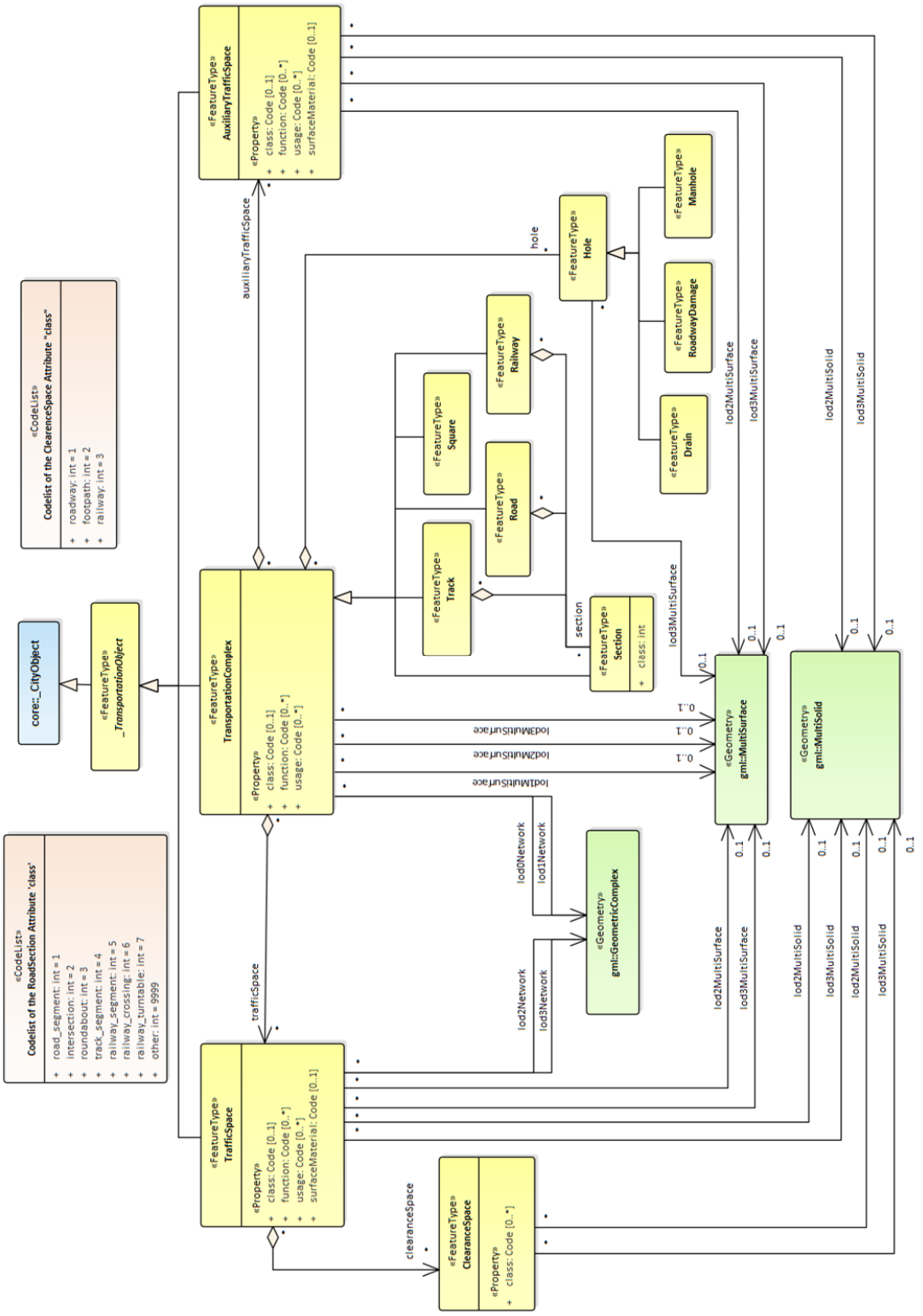


Abb. 1: UML-Diagramm des weiterentwickelten CityGML Transportation Modells (BEIL & KOLBE 2017)

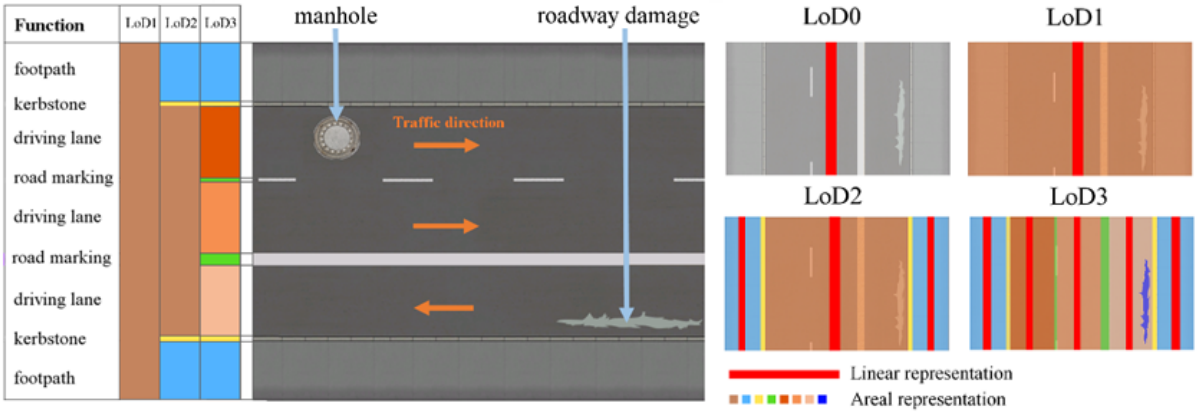


Abb. 2: Straßenabschnitt in LoD1-3 (links), Linien- und flächenhafte Repräsentation in LoD0-LoD3 (rechts) (BEIL & KOLBE 2017)

2.3.4 Einführung von Abschnitten und XLink-Konzept

Der linke Teil von Abbildung 3 zeigt wie ein Straßennetzwerk in einzelne Abschnitte unterteilt werden könnte. Eine Straße würde dabei durch ein *Road* Objekt repräsentiert werden, welches aus mehreren Abschnitten (Sections) besteht. Um welchen Typ es sich bei den jeweiligen Abschnitten handelt, kann durch passende Attributwerte festgelegt werden. Diese sind „road_segment“ (Nr. 1, 2, 4, 5, 6, 8 und 9), „intersections“ (Nr. 3) sowie „roundabout“ (Nr. 7). Im gegebenen Beispiel besteht Road A aus den Abschnitten 1, 3, 5, 7 und 9 während Road B aus den Abschnitten 2, 4 und ebenfalls 3 aufgebaut ist. Um für diesen Fall eine redundante Repräsentation von Abschnitt 3 zu vermeiden, wird das bereits im CityGML-Standard erläuterte und im rechten Teil von Abbildung 3 dargestellte XLink-Konzept vorgeschlagen. Mit Hilfe von XLinks kann die redundante Repräsentation von Geometrien und Geoobjekten, welche gleichzeitig mehreren Objekten zugehörig sind vermieden werden. Die betreffende Geometrie bzw. das Teilobjekt wird dabei nur für ein Geoobjekt explizit repräsentiert, alle weiteren Geoobjekte besitzen lediglich Verweise darauf. Dies wird im dargestellten Instanzdiagramm in Abbildung 3 verdeutlicht.

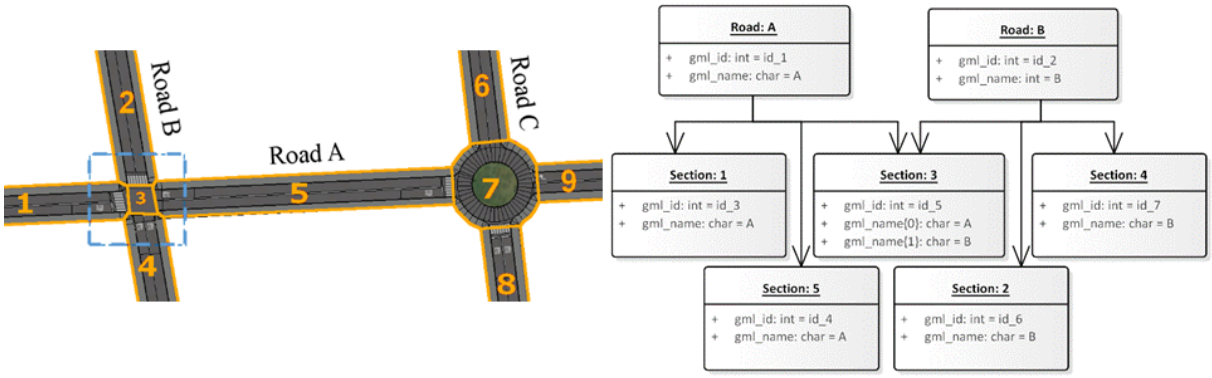


Abb. 3: Straßenabschnitte (links), Instanzdiagramm des XLink-Konzepts für zwei Straßen, die sich einen Kreuzungsbereich teilen (rechts) (BEIL & KOLBE 2017)

2.4 Implementierung des Konzepts für New York City

Für die Umsetzung von ausgewählten Teilen des Konzepts werden drei Hauptdatenquellen herangezogen. Dies sind zum einen die in vorangegangenen Arbeiten (vgl. BURGER 2015, CANTZLER 2015) erstellten und mit Attributen angereicherten Straßenmittelachsen im CityGML-Format, zum anderen umfangreiche flächenhafte Shapefiles, welche in Form einer sogenannten „Planimetric Database" auf dem New York City Open Data Portal[2] zum Download bereitstehen und z.B. Straßenflächen, Gehsteige oder Parkplätze in 2D enthalten sowie weitere Datensätze, in denen zusätzliche Informationen wie beispielsweise Geschwindigkeitsbegrenzungen oder Straßenzustandsdaten enthalten sind. Weiterhin macht das „NYC Street Design Manual" Angaben zu Ausmaßen einzelner Straßenraumobjekte sowie zu verwendeten Materialien.

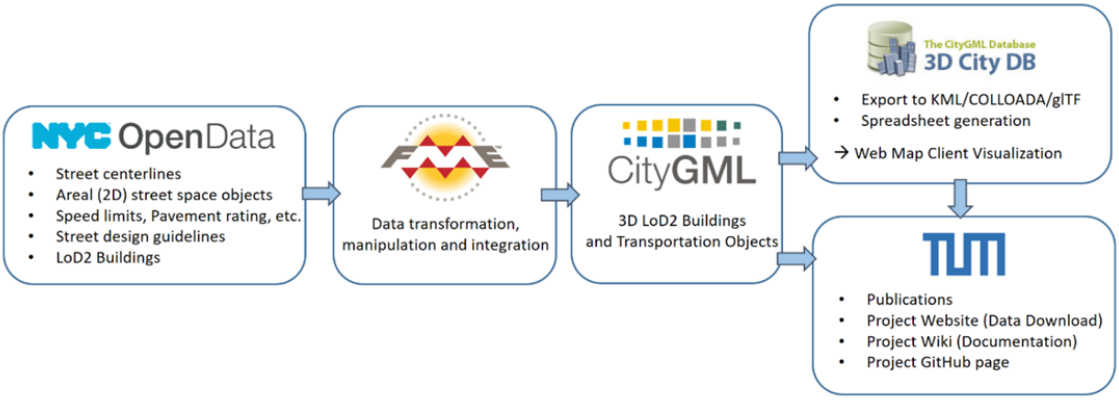


Abb. 4: Workflow zur automatischen Erzeugung eines semantischen Straßenraummodells für NYC

Abbildung 4 zeigt den Ablauf der automatischen Erzeugung eines semantischen Straßenraummodells für New York City. Die vielfältigen Informationsquellen müssen integriert und dabei erarbeitete Vorgaben aus konzeptioneller Sicht berücksichtigt werden. Ziel ist die möglichst automatische Ableitung eines CityGML-konformen Straßenraummodells für das gesamte Stadtgebiet von New York City. Alle Datentransformationen und -manipulationen werden mit der Software „Feature Manipulation Engine 2016.1" (FME) durchgeführt. Zur späteren Speicherung und Integration der teilweise sehr großen Datenmengen wird die Open Source Geodatenbank „3DCityDB Version 3.3" verwendet. Die tatsächliche Umsetzung des Konzepts muss im Rahmen des aktuell gültigen Standards CityGML2.0 umgesetzt werden, Ideen des Konzeptkapitels werden jedoch soweit möglich bereits berücksichtigt. Zunächst werden die Daten der „Planimetric Database" in das verwendete Koordinatensystem (EPSG:32118) transformiert sowie geeignete Testgebiete ausgewählt, mit denen in überschaubarer Datenmenge experimentiert werden kann. Zu komplexe Szenarien wie Autobahnkreuze werden zudem aus den Straßenmittelachsendaten herausgefiltert und gesondert bearbeitet. Die zur Verfügung stehenden Straßenmittelachsen werden zunächst um zahlreiche weitere semantische Informationen angereichert, indem Attribute aus unterschiedlichen Datensätzen mit Hilfe gemeinsamer Identifikatoren wie Straßennamen oder Objekt-ID übertragen werden. Dies umfasst u.a. Angaben zu Geschwindigkeitsbegrenzungen, dem Straßenzustand oder der Anzahl an Fahrspuren. Die so erzeugten „neuen" Straßenmit-

[2] https://opendata.cityofnewyork.us/

telachsendaten werden anschließend mit einer räumlichen Korrelationsmethode den zugehörigen Straßenflächen aus den Daten der „Planimetric Database“ zugeordnet und so semantische Informationen übertragen. Mit Hilfe unterschiedlicher Verfahren wie räumlicher Überlagerung, Nachbarschaftsanalysen oder Abstandsmessungen werden dann sämtliche Objektklassen ebenfalls geometrisch bearbeitet und mit semantischen Informationen angereichert. Gehsteige liegen in den Originaldaten beispielsweise als zweidimensionale Flächen mit Loch vor, die z.B. einen Häuserblock umschließen, aber keinerlei Attribute besitzen. Diese Flächen werden daher zunächst so zerlegt, dass kleinere Einzelflächen entstehen, welche dann eindeutig einer benachbarten Straßenfläche zugeordnet werden können. Attributwerte wie Straßennamen werden daraufhin übernommen und neue Kenngrößen, wie der Flächeninhalt des jeweiligen Objekts, berechnet. Abschließend werden Gehsteigflächen auf Grundlage der Angabe im „NYC Street Design Manual“ um 0,1528m (= 6 Inch) extrudiert und mit passenden Betontexturen versehen. Auf ähnliche Weise werden sämtliche Straßenraumobjekte wie z.B. Parkplätze oder Straßenmarkierungen bearbeitet. Zusätzlich werden weitere Objektklassen, welche in den ursprünglichen Ausgangsdaten nicht enthalten sind, abgeleitet. So überlagern sich Parkplatzflächen in den Zufahrtsbereichen mit den Straßenflächen. Diese Bereiche werden als semantisch eigenständige Objekte (Zufahrten) gespeichert. Bei der geometrischen und semantischen Bearbeitung werden die zuvor erarbeiteten konzeptionellen Überlegungen umgesetzt.

3 Ergebnisse

Insgesamt entstehen so elf thematische Objektklassen mit 508.660 Einzelobjekten, die jeweils der entsprechenden CityGML-Oberklasse des Transportation-Modells (also *Road, Square* oder *Track*) zugeordnet werden und im CityGML-Format gespeichert werden. Tabelle 2 zeigt eine Auflistung aller erstellen Objekte mit den jeweiligen Datenmengen (*komprimierte CityGML zip-files).
Alle Objekte werden zudem mit passenden Texturen versehen um eine möglichst realistische Visualisierung zu gewährleisten.

Tab. 2: Zusammenfassung aller erzeugten Straßenraumobjekte mit zugehörigen Datenmengen

CityGML Klasse	Anzahl Objekte	Datenmenge*
Curb (Randstein)	126‘626	2,02 GB
Parking Lot Entrance (Einfahrt)	24‘185	5,5 MB
Intersection (Kreuzung)	22‘854	7,9 MB
Grass (Grasfläche)	258	0,3 MB
Road Marking (Markierung)	7‘826	3,9 MB
Dividing Strips (Verkehrsinsel)	8‘841	74,8 MB
Roadbed (Straßenfläche)	72‘580	134,9 MB
Sidewalk (Gehsteig)	169‘056	1,3 GB
Parking Lot (Parkplatz)	19‘951	32,2 MB
Plaza (öffentl. Platz)	1‘360	5,5 MB
Interior Sidewalk (Gehweg)	6‘205	15,8 MB
Building (Gebäude)	> 1‘000‘000	2,4 GB

Mit Hilfe der Software 3DCityDB werden daraus KML-/COLLADA-Daten sowie korrespondierende Spreadsheets erzeugt, welche verwendet werden können, um geometrische und semantische Informationen in Form einer gekachelten Visualisierung darzustellen. Hierzu wird der Open Source „WebGL Virtual Globe Cesium“ und der 3DCityDB-Web-Map-Client verwendet. Alle Objekte des Straßenraummodells können hierin, wie in Abbildung 5 dargestellt, individuell ausgewählt und deren hinterlegte Attributinformationen abgerufen werden. In der linken Abbildung ist ein Teil des Stadtmodells im Bereich des Columbus Circle in der Nähe des Central Parks abgebildet. Die rechte Graphik zeigt ein Straßenstück der 5th Avenue vor dem berühmten Flatiron Building in Manhattan. Durch Anklicken können Informationen über Straßennamen, Fläche, Straßenzustand usw. aufgerufen werden. Der Web Client erlaubt interaktive SQL-Abfragen, die zudem kombiniert werden können. So können beispielsweise alle Straßenstücke, welche zur 5th Avenue gehören und sich in schlechtem Zustand befinden, selektiert und deren Gesamtfläche aufsummiert werden. Kreuzungsflächen sind explizit modelliert und besitzen Informationen darüber, welche Straßen sich jeweils kreuzen.

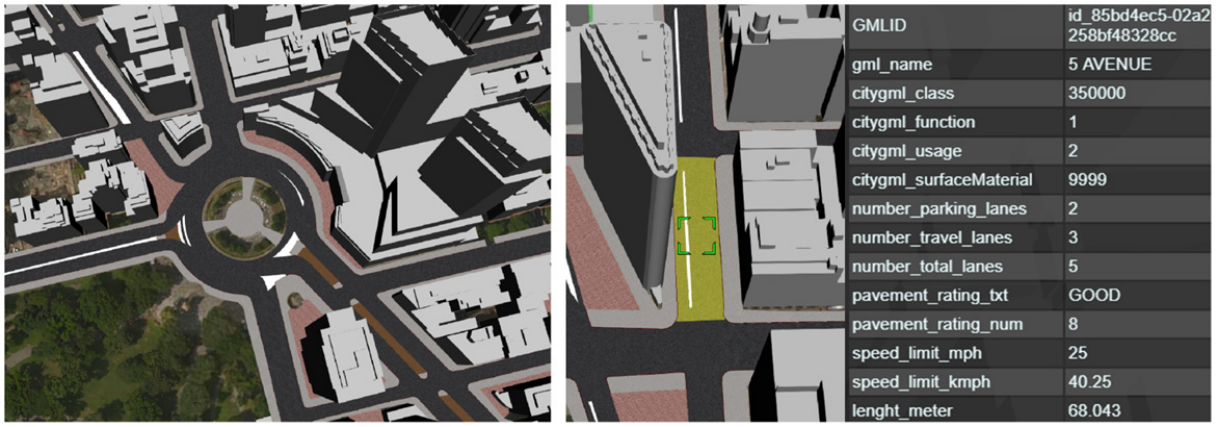


Abb. 5: Columbus Circle (links) und Straßenstück vor Flatiron Building mit Attributabfrage (rechts)

4 Fazit & Ausblick

Die Arbeit befasst sich zunächst mit relevanten Standards im Bereich der Straßenmodellierung und zeigt, dass Straßen zumeist durch linienhafte oder parametrische Repräsentationen dargestellt werden. Eine Analyse potentieller Anwendungsfelder verdeutlicht jedoch den Nutzen, den diese aus flächenhaften Repräsentationen ziehen würden. Eine anschließende Analyse des CityGML2.0 Transportation Datenmodells führt zu zahlreichen Verbesserungs- und Erweiterungsvorschlägen sowie konzeptionellen Überlegungen hinsichtlich möglicher Zerlegungen des Straßenraums in Einzelobjekte. Diese Erkenntnisse werden schlussendlich am Beispiel eines semantischen 3D- Stadtmodells für das gesamte Stadtgebiet von New York City auf Grundlage frei verfügbarer Daten umgesetzt. Die Ergebnisse der Arbeit wurden im Rahmen der 3DGeoInfo Konferenz in Melbourne 2017 vorgestellt. Das hierzu eingereichte Paper (BEIL & KOLBE 2017)

wurde mit dem „Runner-up for the Best Paper Award“ ausgezeichnet[3]. Weitere Informationen können auf der Projektwebsite[4] sowie der Projekt-Wiki-Seite[5] eingesehen werden. Hier finden sich auch mehrere Videos, die im Rahmen des Projekts entstanden sind. Zudem stehen auf der Projektwebsite sämtliche erzeugte Datensätze für das gesamte Stadtgebiet zum Download zur Verfügung. In der Folge wurden und werden die Ergebnisse der Arbeit weitergeführt; so etwa für Verkehrssimulationen oder Solarpotentialanalysen. Die zunehmende Verfügbarkeit von Straßendaten z.B. aus Mobile Mapping Systemen bietet zahlreiche Möglichkeiten, die in dieser Arbeit gesetzten Grundlagen der Straßenraummodellierung aufzugreifen und weiterzuführen. Die Erkenntnisse der Arbeit spielen bei der aktuellen Weiterentwicklung des CityGML Standards eine Rolle.

5 Literaturverzeichnis

BAST, 2016: Objektkatalog für das Straßenwesen, Version 2.017. Bundesanstalt für Straßenwesen.

BEIL, C., 2017: Detaillierte Repräsentation des Straßenraums in 3D Stadtmodellen. Lehrstuhl für Geoinformatik, Technische Universität München.

BEIL, C. & KOLBE, T.H., 2017: CityGML and the Streets of New York – A Proposal for detailed Street Space Modelling, ISPRS Annals of the Photogrammetry, Remote Sensing ans Spatial Information Sciences, **IV**(4/W5), 9-16.

BMVI, 2014: Anweisung Straßeninformationsbank Kernsystem Version 2.03. Bundesministerium für Verkehr und digitale Infrastruktur.

BURGER, B., 2015: Erzeugung eines semantischen 3D-Stadtmodells der Stadt New York auf der Basis von Open Data - DGM, Straßen und Zonierung. Lehrstuhl für Geoinformatik, Technische Universität München.

CANTZLER, B., 2015: Erzeugung eines semantischen 3D-Stadtmodells der Stadt New York auf der Basis von Open Data - DGM, Gebäude, Flurstücke, Gewässer, Parks und Bäume. Lehrstuhl für Geoinformatik, Technische Universität München.

DUPUIS, M., BAHRAM, M., GREZLIKOWSKI, H., RICHTER, A., GOLDBERG, A., KLAERNER, E., BOCK, B. & KAELLGREN, L., 2015: OpenDRIVE – Format Specification, Rrev. 1.4. VIRES Simulationstechnologies GmbH.

GRÖGER, G., KOLBE, T.H., NAGEL, C. & HÄFELE, K.-H., 2012: OGC City Geography Markup Language (CityGML) Encoding Standard, OGC 12-019.

ISO 14825, 2011: Intelligent transport systems – Geographic Data Files (GDF) – GDF5.0 International.

LÖWNER, M.-O., GRÖGER, G., BENNER, J., BILJECKI, F. & NAGEL, C., 2016: Proposal for a new LoD and multi-representation concept for CityGML. ISPRS Annals of the Photogramme-

[3] https://www.gis.bgu.tum.de/aktuelles/news-single-view/?tx_ttnews%5Btt_news%5D=97&cHash=b7bc6358ec84a142dced5e2c7143f3cc

[4] https://www.gis.bgu.tum.de/projekte/new-york-city-3d/

[5] https://wiki.tum.de/display/gisproject/3D+City+Model+of+New+York+City

try, Remote Sensing and Spatial Information Sciences, **IV**(2/W1), 11th 3DGeoinfo Conference in Athens, Greece.

RANDT, B., BILDSTEIN, F. & KOLBE, T.H., 2007: Use of Virtual 3D Landscapes for Emergency Driver Training. Proc. of the Int. Conference on Visual Simulation IMAGE in Scottsdale, Arizona.

ROSS, L., 2010: Virtual 3D City Models in Urban Land Management - Technologies and Applications. Ph.D. Thesis, Technische Universität Berlin, Germany.

Crowdsourced Vertical Indoor Mapping

OMIDREZA MOSLEHIRAD[1], GEORGIOS PIPELIDIS[2], DOROTA IWASZCZUK[1], CHRISTIAN PREHOFER[2] & URS HUGENTOBLER[3]

Abstract: In this paper, we introduce a methodology for dynamic estimation of vertical features of buildings. For this, we obtain data from built-in smart-phone sensors and use this data for altitude estimation via the barometric formula. We propose a novel approach for the identification of the reference pressure during the outdoor-to-indoor-transition of the users. We use a set of machine learning algorithms for the determination of the number of floors as well as an unsupervised clustering of each floor altitude. This is the first system capable of mapping vertical features inside buildings automatically, which is independent of any indoor infrastructure. Finally, we apply the results for enhancing a CityGML standard for generating 3D building models.

1 Introduction

Indoor maps have many applications in e.g. robotics, augmented reality and location based services. Although there are proposed models for dynamic generation of indoor maps from crowdsourced data (GRZONKA et al. 2010; ALZANTOT & YOUSSEF 2012), no methodologies for vertical mapping using crowdsourced data exist. In this paper, we enhance existing models to include information regarding the number of floors and floor height. We also propose a novel approach for identifying reference locations via human transitions from outdoor to indoor, which can be used for a precise indoor altitude estimation. First, we use the crowdsourced smart-phone data to identify the user transition from outdoor to indoor. This serves as a landmark for the extraction of a reference pressure that can be used to estimate altitude differences via the barometric formula. Vertical transitions (e.g. stairs or elevators) are filtered out, since they do not belong to floors. The remaining altitude values from multiple users are aggregated for the identification of the number of floors and each floor height. Studies have attempted to vertically localize humans or objects via pressure sensors (BOLLMEYER et al. 2013; LI et al. 2013; XIA et al. 2015). However, they all assume that reference sensor stations are permanently installed in the buildings. These are thus highly infrastructure-dependent approaches. Approaches for the dynamic generation of maps have also been proposed (KIM et al. 2014; BOETERS et al. 2015; LOCH-DEHBI et al. 2017). However, these approaches use outdoor characteristics (e.g. windows, building footprints) to map indoor areas, which in most cases is not practical since most buildings contain indoor structures that are invisible from outside, including underground structures. This is also a problem for optical and microwave imagery techniques from airborne and spaceborne platforms,

[1] Technische Universität München, Lehrstuhl für Astronomische und Physikalische Geodäsie, Arcisstr. 21 D-80333 München, E-Mail: [omidreza.moslehi, dorota.iwaszczuk]@tum.de
[2] Technische Universität München, Fakultät für Informatik, Boltzmannstr. 3, D-85748 Garching, E-Mail: [georgios.pipelidis, christian.prehofer]@tum.de
[3] Technische Universität München, Forschungseinrichtung Satellitengeodäsie, Arcisstr. 21, D-80333 München, E-Mail: urs.hugentobler@tum.de

which are not able to easily sense deep internal geometries of buildings such as floors and their vertical locations.
In this paper, we present an approach for dynamic vertical mapping of building interiors using only crowdsourced smartphone data. In our method, we perform a self-calibration of the pressure sensor for every acquisition, which makes our method independent of permanently installed reference sensors. Presented work is a continuation of our previous research (PIPELIDIS et al. 2017). We extended the previous work for outdoor-indoor transition identification with image processing techniques such as Otsu's method, where we first convert the signals into one-dimensional gray images, then we perform image binarization via the Otsu's method to separate indoors from outdoors. We evaluated the results for not only three buildings, but seven different buildings. We also solved the open question of what if the user collects data only from indoor or only from outdoor, so we can determine if the collected dataset is from only indoor or outdoor.

2 Methodology

We collect data from smart-phone users via an application that is available here (PIPELIDIS 2017). We filter out the noise and outliers from the collected data. The next step is to filter out data from outdoors, leading to identification of areas where we extract indoor pressure readings. Following this, we filter features that belong to intermediate heights (i.e. stairs or elevators). The barometric formula is used for altitude estimation from pressure readings. We aggregate all altitude data from multiple visits of the targeted building, because the users often do not visit all the floors every time. The number of floors and every floor altitude are estimated from the aggregated datasets via clustering methods. We use this information to enhance indoor characteristics of CityGML model of buildings at lower levels of detail (LoD), presented as LoD2 Plus models.

2.1 Sensor Data Collection and Signal Filtering

We aim at fusion of all available smartphone data in order to get the best vertical mapping result. For this, we collect data from pressure, light, proximity, GPS and magnetic sensors. From the pressure sensor, we collect atmospheric pressure, which will be used for altitude determination via the barometric formula. The light sensor is sensitive to any light sources when the phone is outside of the pocket. It collects light values in the unit of lx, which measures the luminous flux per unit area. The proximity sensor produces a Boolean signal. If there is an object very close to it, the signal pattern will be changed to True. This is useful if during the data collection we want to know whether the phone is in a pocket or not, so if the phone is in the pocket, then the light sensor will not sense any light sources and the light signal will not be reliable at that period. From the GPS sensor, we calculate the Geolocation uncertainty of the user. The Geolocation uncertainty refers to the accuracy of user positioning via the GPS sensor in the unit of meter. Since the visibility of GPS sensors to GPS satellites is reduced inside buildings, the positioning accuracy will be reduced by several meters of error. From the magnetic sensor, we observe magnetic disturbances of the environment. In our study, we used data collected on different dates and at different day times, temperatures and humidity conditions. All the collected datasets are available in our public GitHub repository (MOSLEHIRAD 2017). For filtering out the outliers from

pressure signal, we used a third-order one-dimensional median filter (PRATT 2007). An exemplary pressure dataset is plotted in Fig. 1.

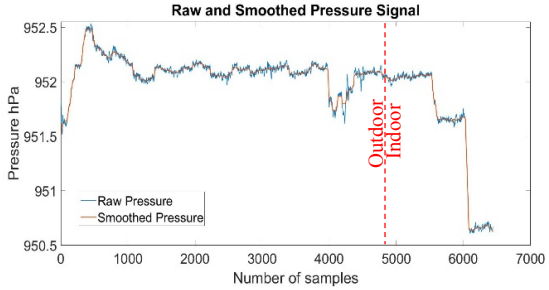


Fig. 1: Pressure sensor from outdoor and indoor. After the sample number 5000, the user entered into the building and visited three different floors: Ground, first and third floors, respectively.

2.2 Identification of Reference Pressure and Outdoor-Indoor Transition Point

The reference pressure is a big challenge for altitude estimation via the reformulation of the barometric formula $h = \left[\frac{P_0^{\frac{1}{5.25}}}{P_i} - 1\right] * \frac{T_b + 273.15}{0.0069}$, where, h is height, P_0 is the reference pressure extracted immediately after the transition point, P_i is the current pressure value and T_b is the temperature value obtained from the online weather stations. In our novel approach, we estimated P_0 from the outdoor-indoor transition of users instead of permanent installation of reference pressure sensors in buildings. For instance, after the stair-removal phase (Section 2.3), the first cluster of pressure samples are considered as the ground floor pressure samples, which we compute its mean as the reference pressure P_0. Another challenge is identification of the outdoor-indoor transition point itself. We fused built-in smart-phone sensors such as light, proximity, GPS and magnetic sensors since their signal patterns, in most cases, change significantly at transition points. In our data collection application, we modified the sampling frequency of these sensors for battery power management. In Fig. 2a, during the day, light intensity drops significantly at transition point and increases during the night, while the proximity sensor indicates whether to trust the light sensor, due to various phone poses (e.g. phone inside of pocket). In Fig. 2b, the Geolocation uncertainty pattern remained the same and is independent of day or night (always increases at transition point).

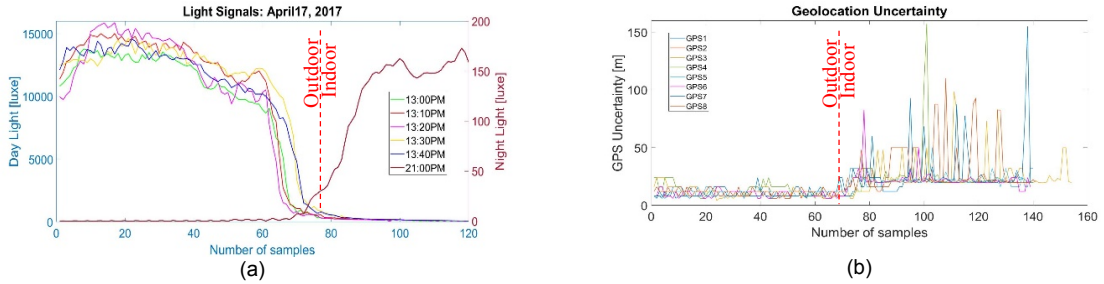
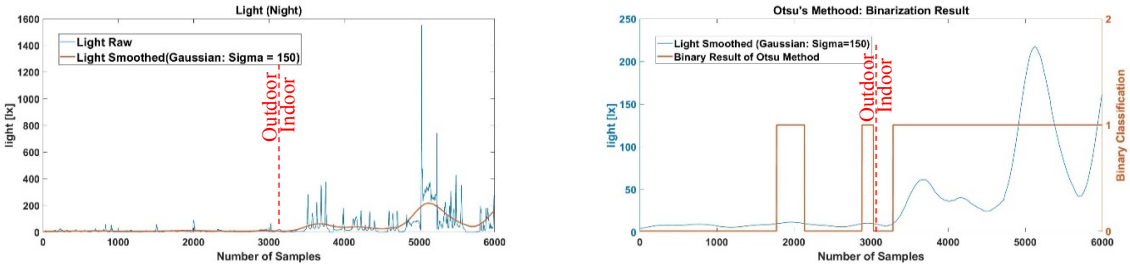


Fig. 2: Light and GPS Uncertainty data from several outdoor-indoor transitions

First, we smooth the light signal by the Gaussian filter (Fig. 3a) and convert it to a one-dimensional gray image. Afterwards, we convert it to a binary image via the Otsu's method (OTSU 1979), (Fig. 3b). If the data is from day time, an automatic gray threshold estimation for binarization is estimated, while for the night datasets, we consider a fixed empirical threshold of 10 lx. In order to have a uniform gray-scale for all the light datasets, we reduce all the samples

over the threshold of 2000 lx to 2000 lx. We selected this empirical threshold equivalent to the maximum gray-level of 255. In Fig. 3, an exemplary light dataset collected during a night is plotted.



(a) (b)

Fig. 3: (a) A light data collected during a night and its smoothed signal. The user entered the building after the sample 3000. During this data collection, there were not sufficient light sources in the ground floor (samples from 3000 to 3500). (b) Binarization result via the Otsu's method

The reliability of light datasets during nights is lower than during the days, because during days, we always guarantee the Sun to be available outdoors, but we do not guarantee if there will be all necessary light sources available indoors. Besides, if the data is collected during the day, the light intensity contrast between indoors and outdoors is higher than during the night, which gives a better contrast for classification of indoors and outdoors. In above example, we considered the worst case: light dataset from night and when we entered the building, there were not sufficient light sources in the ground floor. However, we solved this problem by fusing multiple sensors. One of the main reasons why light datasets from night time are still useful for outdoor-indoor transition identification is the difference between received light intensity from light sources at different distances from the smart-phones. During the night, outdoors, artificial light sources have much higher distance from smart-phones than indoors (e.g. light bulbs at ceilings vs. light bulbs over streets). For instance, in Fig. 3a, the light samples from 0 to 3000 are from outdoor, which have very low intensity values than indoors during the night. Unlike artificial light sources, the Sun light is independent of distance from smart-phones and produces very high intensity values. Due to smoothing via the Gaussian filter, the identified transition points via the Otsu's method might be slightly displaced from the true locations. The other problem is the areas that wrongly identified as indoors (Fig. 3b, area around the sample number 2000). To solve these issues, we used Changepoint detection (LAVIELLE 2005) over the slightly smoothed light signal in Fig. 4a, where the mean of the signal changes significantly at change points, especially at the transition point (dashed green lines). This divides the signal into multiple segments. Finally, we use the Otsu's binary result as a weighting system for assignment of binary values to each segment (Fig. 4b). For example, the largest segment in Fig. 4b is over the samples 0 to 3462, where the mean of binary values of Otsu's result is lower than 0.5, so we assigned binary label of zero to all samples over this period, while for samples from 3462 to 5000, we have mean value of higher than 0.5, so we assigned binary label of one to all samples in this period. In our implementations, we assigned binary values of one to indoors and zero to outdoors. After applying the weights, the resulting final binary classification of light from night time can be seen in Fig. 4c.

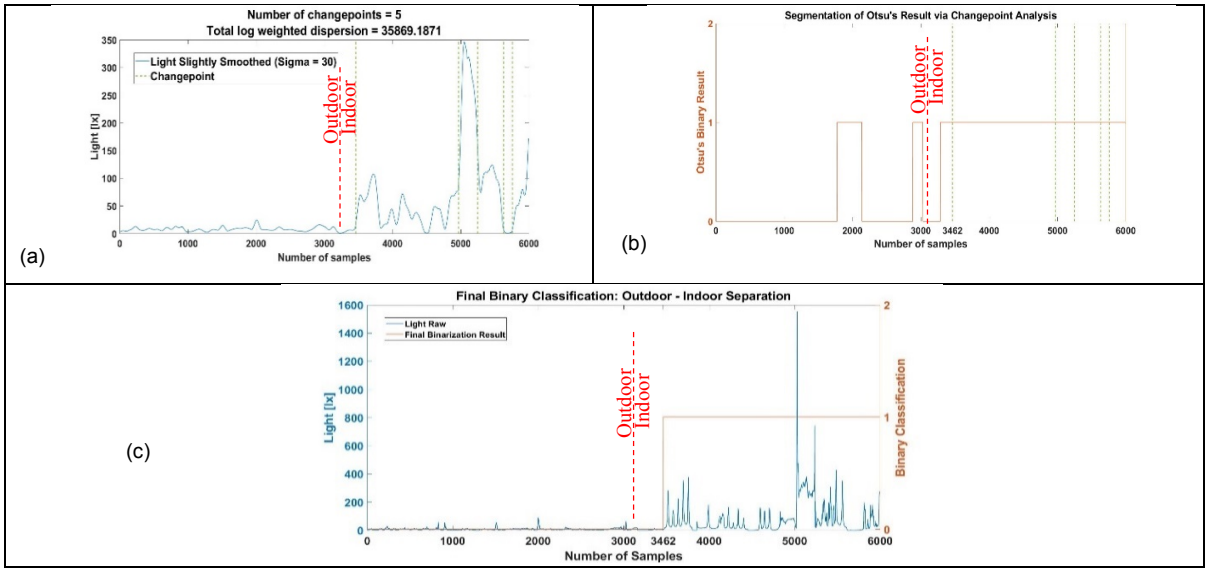


Fig. 4: (a) Changepoint analysis, where the signal is divided into 6 segments. (b) Segmentation of Otsu's result via mapping the change points over the Otsu's result. (c) Final binary classification of the light signal into indoor (binary values of ones) and outdoor areas (binary values of zeros)

However, in Fig. 4c, we are not sure if the sample number 3462 is exactly the transition point or not, because, as we mentioned before, after passing the true transition point (dashed red line), there were no sufficient light sources in the ground floor. Hence the point 3462 might be corresponding to the stairs or the next floor we visited after the ground floor. In order to find the correct transition point position, first we synchronized the light and Geolocation uncertainty signals based on their time dimension. Then we used the sudden increase of Geolocation uncertainty at the transition point as an additional indicator for a transition point (Fig. 5, dashed red line).

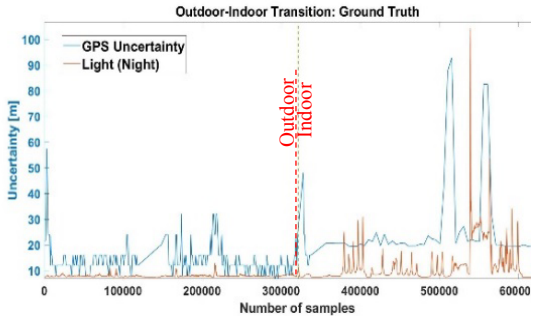
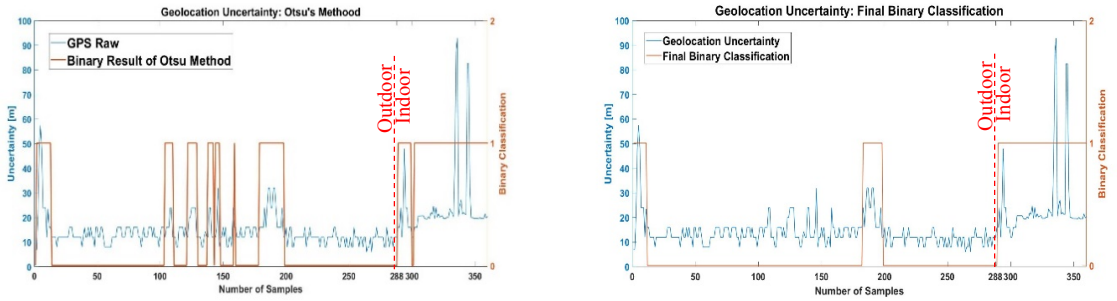


Fig. 5: The sudden rise in Geolocation uncertainty at transition point (at the dashed line) is used as an additional indicator of transition point. For synchronization, we resampled both the signals to the time unit of millisecond

For the GPS sensor, we used a similar approach. GPS is not accurate indoors due to the existence of walls and barriers, which results in a sudden increase of GPS uncertainty during outdoor-indoor transitions (Fig. 2b). First, we defined a bounding threshold of 70m to filter very high uncertainties, equivalent to a maximum gray-level of 255. Unlike the light signal where the threshold for binarization is calculated automatically, we used empirical uncertainty level equivalent to 13m, because the GPS uncertainty outdoors is usually below this level. In order to enhance the result of the Otsu's method, similar to the light signal, we perform the Changepoint detection analysis. As can be seen in Fig. 6b, the Changepoint analysis reduced number of outli-

ers significantly and helped to estimate the true transition point (sample 288). In the final binary classification of the Geolocation uncertainty, before the sample 288, there are only two small areas wrongly identified as indoors. As we mentioned before, we can remove the remaining outliers via fusion of multiple sensors.



(a) (b)

Fig. 6: (a) Binarization result of the Geolocation uncertainty signal via the Otsu's method. (b) Final binary classification after segmentation of Otsu's binary result via the Changepoint analysis. The outlier around the sample number 200 is due to the presence of a bridge before the TUM Main Campus building entrance. This type of outliers can be removed via sensor fusion

In Fig. 7a, due to the existence of metallic structures and high electromagnetic variations indoors, we expect higher magnetic disturbances indoors than outdoors. In this figure, from the sample 5026, the user entered the building. First, we define a moving-standard-deviation window with a small size along the raw magnetic signal, which produces the first standard-deviation-signal (Fig. 7b). Following this, we define another moving-standard-deviation window, but with a larger size, to compute second-standard-deviation signal from the first one (Fig. 7c). Afterwards, we use Gaussian filter to smooth the second signal (Fig. 7d). This helps to increase the contrast between indoor and outdoor parts of the second standard-deviation signal. The mean value of the smoothed signal is used as the threshold for the binary classification of the magnetic signal into indoors and outdoors (Fig.7e). The first moving standard deviation normalizes the disturbances (Fig. 7b), while the second moving standard deviation is used to increase the amplitude of highly disturbed parts of the first moving standard deviation signal (Fig. 7c), which correspond to indoor areas.

At the end, we fused the binary results of all sensors to increase accuracy of outdoor-indoor binary classification result (Fig. 8). The fusion is based on what the majority of the sensors score (Tab. 1). For instance, if two out of three sensors score a sample as outdoor, then we label it as outdoor. If the proximity sensor indicates that the light signal is not available (phone in pocket) for a short period of time, then we ignore the light readings during that period and we rely only on GPS uncertainty signal, because fusion of GPS and Magnetic sensors does not improve accuracies (if we use only GPS sensor, it gives better results than fusion of these two).

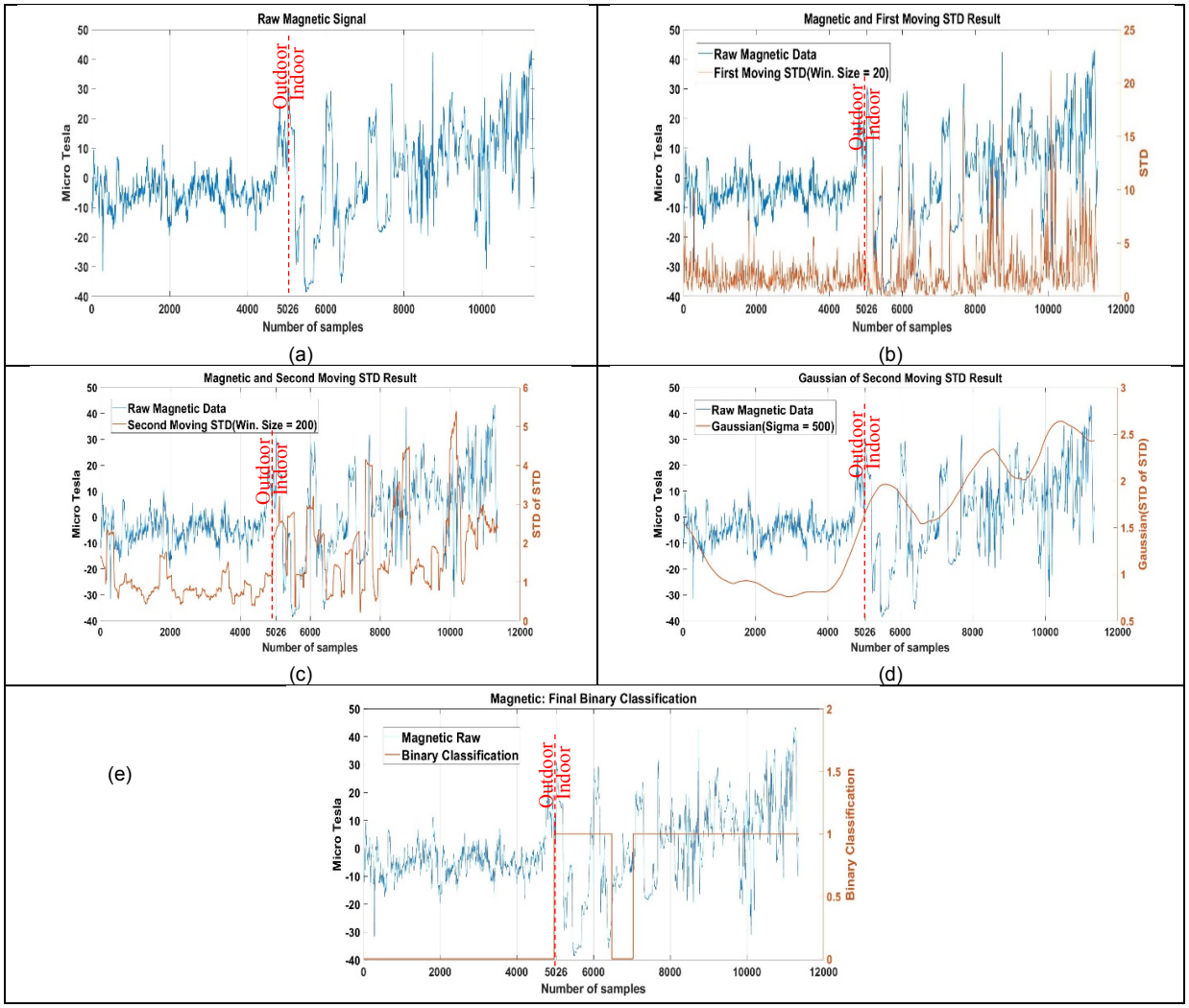


Fig. 7: (a) Magnetic disturbances from outdoor and indoor. From the sample 5026, the user entered a building. (b) The moving-standard-deviation algorithm applied over the raw magnetic signal with window size of 20. (c) Another moving-standard-deviation process, but with a larger window size of 200, which is applied over the previous moving-standard-deviation result obtained in b. (d) Gaussian smoothing of the resulting signal in c. (e) Final binary classification result

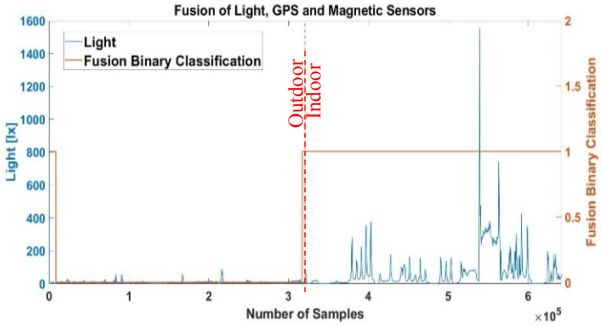


Fig. 8: Fusion result of multiple sensors. As can be seen in the figure, majority of indoor area is correctly labeled as indoor, while small section of the signal (start of the signal) is wrongly identified as indoor. The light signal is from night time and the dashed line is the true outdoor-indoor transition point

Tab.1: Truth table of our fusion concept for scoring what the majority of sensors indicate for outdoor-indoor transition identification. If the light signal is not available (phone in pocket), we rely only on GPS uncertainty.

Light	GPS	Magnetic	Fusion = Light ∧ (GPS ∨ Magnetic) ∨ (GPS ∧ Magnetic)	Indoor / Outdoor
F	F	F	F	Outdoor
F	F	T	F	Outdoor
F	T	F	F	Outdoor
F	T	T	T	Indoor
T	F	F	F	Outdoor
T	F	T	T	Indoor
T	T	F	T	Indoor
T	T	T	T	Indoor

In order to know if a light dataset is from day or night, we compute the maximum-bin-limit of its histogram. If this value is greater than or equal to 600, we consider it as a day dataset, and for values lower than 600, we assume it as a night dataset. If the user collects data only from indoor or only from outdoor, since we used image processing techniques, our approach will not fail. For example, Fig. 9 shows that if we collect light data during the day only from outdoor or only from indoor, our approach very accurately labels the entire signal as outdoor or indoor, respectively. In our approach, if the light signal is only from indoor, it mistakenly considers the dataset as a night dataset because its maximum-bin-limit becomes less than 600, so it applies fixed gray level of 10 lx for binarization. Since overall intensity of indoor light is usually much higher than 10 lx, the Otsu's method labels the entire signal as indoor (Fig. 9b). We also noticed that the light datasets are not reliable during the sunrise and sunset as the contrast between outdoor and indoor is not sufficient for binary classification via the Otsu's method. As a result, during these periods, we only rely on GPS uncertainty signal. To find out the sunrise or sunset periods, we can use the Sun angle with respect to the user's global coordination. There are online services that provide such information in real-time.

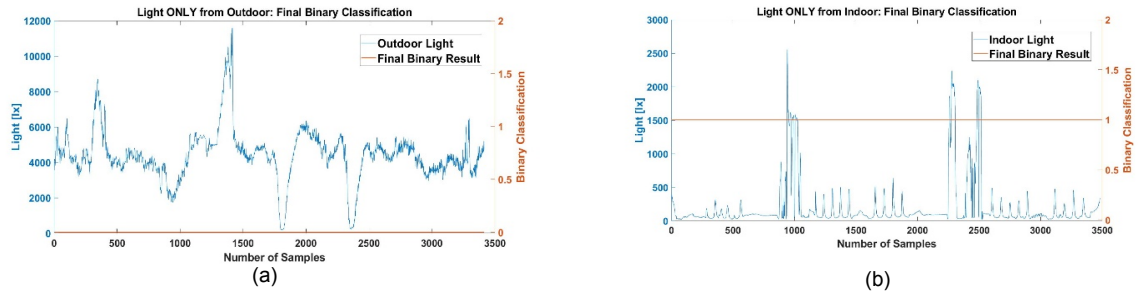


Fig. 9: Outdoor-indoor classification of light data only from outdoor or indoor, respectively

For the Geolocation uncertainty and magnetic signals, since they are not dependent on day or night, we do not have complexity of light data processing. For example, if we collect Geolocation uncertainty only from indoor or only outdoor, again our methodology works properly (Fig. 10).

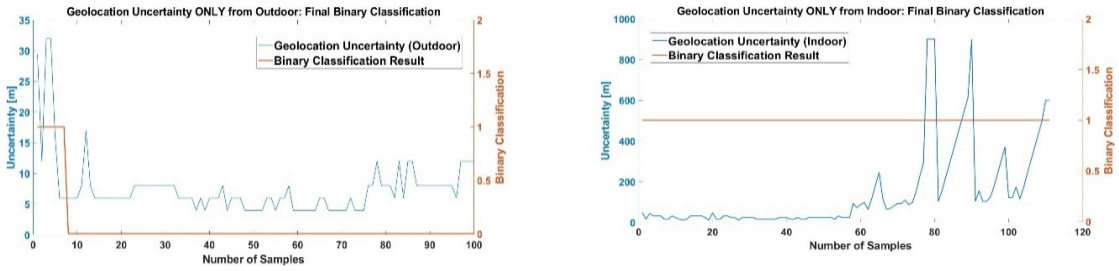


Fig. 10: Outdoor-indoor classification of Geolocation uncertainty data only from indoor or outdoor

2.3 Stair Removal

In the stair removal phase, we filter out data from stairs, elevators and outliers via the moving standard deviation algorithm. An exemplary result of this algorithm is shown in Fig. 11:

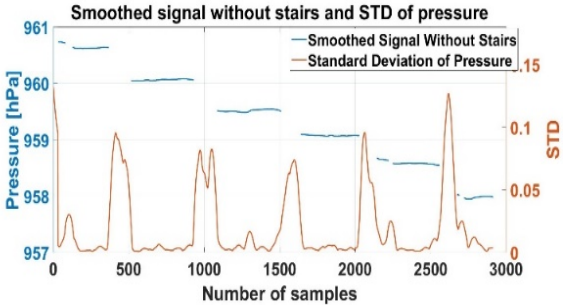
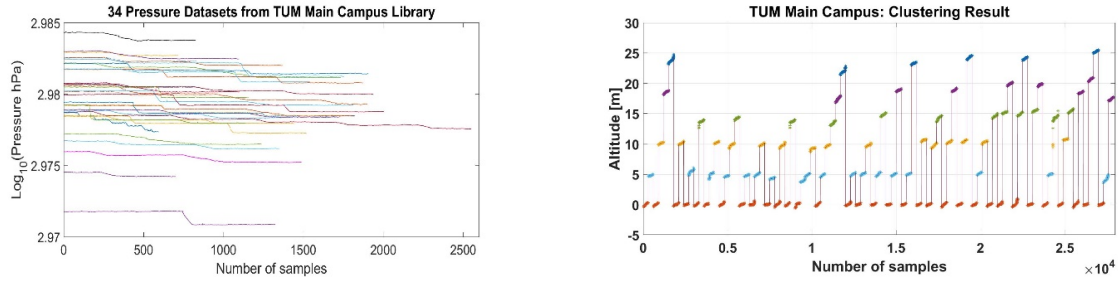


Fig. 11: Stair removal phase via the moving-standard-deviation algorithm. This algorithm filters out pressure readings from stairs and elevators. By setting a standard deviation threshold ($\sigma = 0.019$) and the sliding standard deviation window (size 28.5), we are able to remove pressure values with higher standard deviation than the threshold. At each movement, the algorithm computes the standard deviation of pressure values within the window. If the computed standard deviation is higher than the threshold, it filters out the corresponding pressure sample from the middle of the sliding window.

2.4 Number of Floors and Floor Altitude

After identification of floors pressure values via outdoor-indoor binary classification and stair removal phase, we estimated altitude values via the reformulation of the barometric formula in (Section 2.2). We aggregated multiple user altitude datasets, since not all users are expected to visit all floors. Since we extract the reference pressure by the same device we use for estimating floor altitude, unlike (XIA et al. 2015), there is no need for calibration between different barometric sensors. Since from the beginning the number of floors and their corresponding altitudes are unknown, for classification, we used K-Means because of its popularity and its relatively low processing demand and selected the Elbow method for estimating K (number of floors or clusters). The cluster centers represent the altitudes of the floors. In Fig. 12a, we presented a long-term pressure data collection from the TUM Main Campus Library building. We collected the datasets in 6 months, day and night, and in different weather conditions. Fig. 12b shows the altitude data aggregation of this building and the clustering result. In this figure, each color code corresponds to a specific floor, starting from ground floor to 5th floor (floors 0, 1, 2, 3, 4 and 5).



(a) (b)

Fig. 12: Pressure data collection and clustering result of aggregated altitude datasets.

2.5 CityGML LoD2 Plus Generator

We implemented a fully automatic CityGML LoD2 Plus building model generator via open tools and Java libraries such as Node-RED, citygml4j and OpenStreetMap queries, and visualized the models in a Virtual Reality environment using Unreal Engine. In another words, we insert raw sensor datasets as the input and obtain the corresponding building model as the output.

3 Evaluation

3.1 Data Collection Schedule

For evaluation of our proposed approach, we defined a long-term (six months) data collection schedule for each of the selected three buildings. We specified the floors to be visited in long-term period, day and night and in different weather conditions. For collecting each dataset, we walked about 800m outdoor and indoor, in total about 56km for all the collected datasets from all the buildings. The schedules can be found in (MOSLEHIRAD O. 2017), which describe all the details of collected datasets such as date, time, visited floors, temperature from AccuWeather (C°), temperature from the Weather Channel(C°), humidity from AccuWeather, humidity from the Weather Channel, ambient pressure from AccuWeather, ambient pressure from the Weather Channel, walking speed and cloud coverage. We wanted to know i.e. how the temperature, humidity or weather conditions can affect altitude readings in long-term visits of the buidings. At the end of the experiments, we noticed that the influence of temperature, humidity and weather variations have minor effects on the final floor altitude estimation results.

3.2 Influence of Different Phone Poses

All the evaluations are available in our public GitHub repository (MOSLEHIRAD 2017). A pre-study was conducted where the influence of different phone poses was estimated. These poses are: Phone (1) only in hand, (2) only in jacket pocket, (3) only in the swinging hand, (4) only in trouser pocket and (5) in a random assortment of above poses. We found that the influence on pressure due to different smart-phone poses is minor, with maximum vertical displacement of 40cm.

3.3 Influence of Stair Removal

We recorded data with three different walking velocities, approximately 1x, 1.5x and 2x, while climbing five flights of stairs in a building. The algorithm scores a precision of 93.8%, recall

95.5% and F-Score 94.6%, on correctly identifying the stairs, with the same moving standard deviation window length for all the data sets.

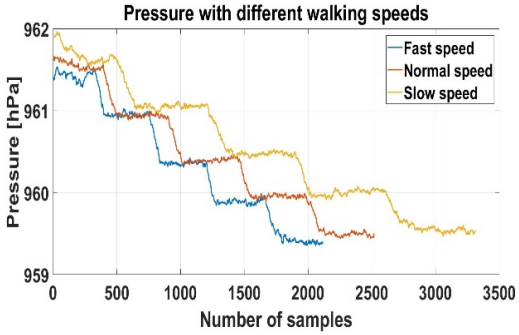


Fig. 13: Indoor pressure signals in three different walking speeds.

3.4 Evaluation of Reference Pressure Area (Outdoor – Indoor Transition)

We evaluated our approach for seven different buildings, each with four times visits for data collection during day and night. These buildings are located in Munich and the datasets as well as confusion matrices can be found in (MOSLEHIRAD 2017). Tab. 2 presents the accumulated binary classification result of all datasets from the seven buildings. According to the results, we can see that if the light is always available, we will have a better indoor area identification than GPS, magnetic and even fusion. However, light is not always available, so in this case, fusion gives a better reliability. Besides, fusion gives the best outdoor identification with true negative of 97,6% and precision of 98%, respectively. In terms of accuracy and F-Score, fusion comes after light, but as mentioned before, light is not always available. The less reliable sensor is the magnetic sensor since there are lots of magnetic noises outdoors, which affect the binary classification results.

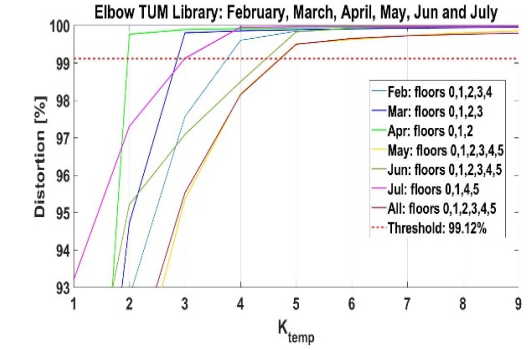
Tab. 2: Total evaluation of outdoor-indoor transition via light, GPS and magnetic sensors and fusion result

	Recall	True Negative	Precision	Accuracy	F-Score
Light	97,3%	95,7%	96%	96,6%	96,7%
GPS	89,1%	93,4%	85,2%	92,1%	87,1%
Magnetic	84,1%	85,7%	85,1%	84,9%	84,6%
Fusion	94,5%	97,6%	98%	96%	96,2%

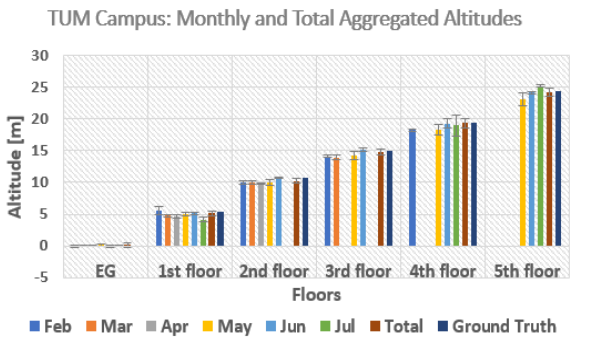
3.5 Evaluation of Number of Floors and Floor Altitude

We evaluated the estimated number of floors and floor altitude via 34, 25 and 20 datasets, collected form TUM Library, AO Laim and DeutschAkademie buildings, respectively, based on a schedule. We evaluated our approach based on monthly and total aggregation of altitude datasets. This helped us to compare the accuracy of partially or fully aggregated datasets. Fig. 14 (a, c and e) show the Elbow result for the three buildings. For instance, in (a), during April, the aggregated datasets for this month cover only the floors 0,1 and 2. The first K_{temp} after the threshold of 99.12% for this month is 2, and based on our implementation, we always increment K_{temp} by one to obtain the true K, which in this month is three (three visited floors). This threshold is given empirically and works correctly for all the datasets from all the buildings. All the selected buildings have the same number of floors (6 floors: Ground as **floor 0** to 6th floor as **floor 5**). As can be seen in Fig. 14 (b, d and f), the fully aggregated datasets result in a better floor altitude estimation as they are closer to the ground truth that is obtained via a laser range

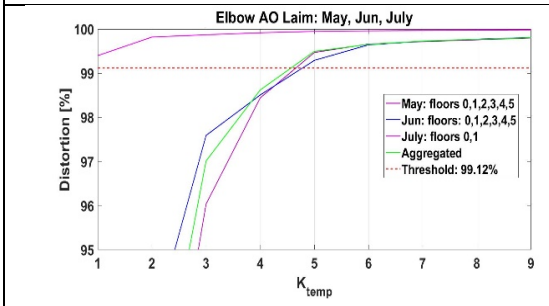
meter. The error bars in (b), (d) and (f) are based on the standard deviation of each cluster. Tab. 3, 4 and 5 illustrate the ground truth, estimated floor altitude and absolute errors for the three buildings.



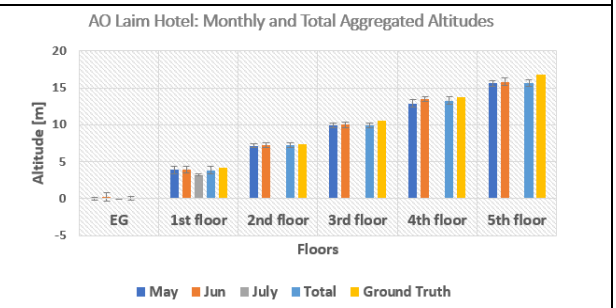
(a) TUM Elbow: Monthly and total data aggregation.



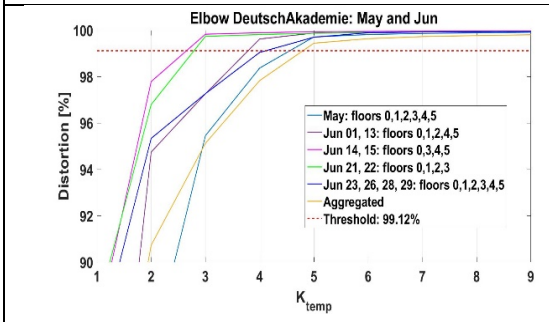
(b) TUM: Monthly and total of 34 dataset aggregation.



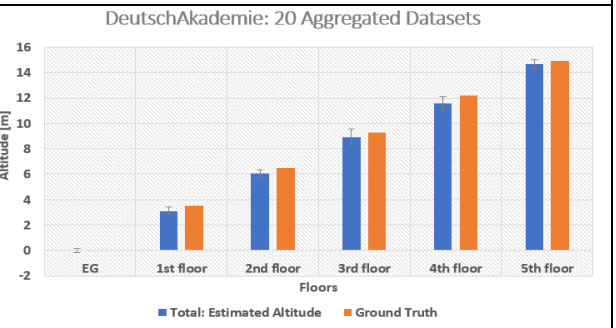
(c) AO Elbow: Monthly and total data aggregation.



(d) AO: Monthly and total of 25 dataset aggregation.



(e) De.Aka. Elbow: Monthly and total data aggregation.



(f) DeutschAkademie: Total of 20 dataset aggregation.

Fig. 14: Evaluation of the number of floors and floor altitude for three different buildings

Tab. 3: Ground truth, estimated floor altitude and absolute errors for the TUM Main Campus building

Floors	Ground	First	Second	Third	Fourth	Fifth
True Altitude (m)	0	5,3	10,68	15,05	19,47	24,41
Estimated Altitude (m)	0	4,81	10,03	14,48	18,86	23,74
Absolute Error (m)	0	0,48	0,65	0,57	0,61	0,66

Tab. 4: Ground truth, estimated floor altitude and absolute errors for the AO Laim Hotel

Floors	Ground	First	Second	Third	Fourth	Fifth
True Altitude (m)	0	4,17	7,31	10,5	13,7	16,8
Estimated Altitude (m)	0,02	3,86	7,24	9,92	13,26	15,68
Absolute Error (m)	0	0,31	0,07	0,58	0,44	1,12

Tab. 5: Ground truth, estimated floor altitude and absolute errors for the DeutschAkademie building

Floors	Ground	First	Second	Third	Fourth	Fifth
True Altitude (m)	0	3,54	6,51	9,31	12,2	14,9
Estimated Altitude (m)	0	3,1	6	8,9	11,59	14,67
Absolute Error (m)	0	0,4	0,45	0,38	0,61	0,23

3.6 Generated CityGML LoD2+ Models and Demonstrations in VR

Fig. 15 shows our generated enhanced CityGML models (LoD2+) for a couple of buildings and representation of a model in VR environment. We used different tools such as citygml4j open java library, OpenStreetMap queries and Node-Red to automate the entire model generation process from sensor signals. Some of the demonstrations can be found in (VIDEO 01 & VIDEO 02).

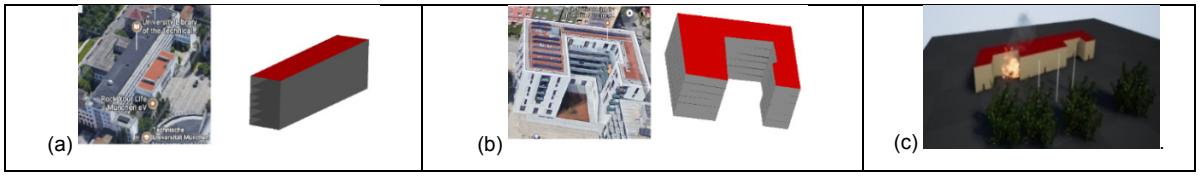


Fig. 15: CityGML LoD2 Plus models of TUM Main library, DeutschAkademie and VR visualization.

4 Conclusion and Future Work

This paper proposed our approach for the vertical mapping of buildings via crowdsourcing. We estimated each floor altitude via the barometric formula. For this, we introduced a novel approach for the extraction of reference pressure during the outdoor-to-indoor-transition of users, which is identified via sensor fusion. As a continuation of our previous work (PIPELIDIS et al. 2017), we proposed a different approach for outdoor-indoor transition identification: image-processing-based methods to have a more robust and general methodology. We evaluated the results for a higher number of buildings (seven buildings), while it was for only three buildings in the previous work. Thanks to the image processing algorithms and histogram analysis of the datasets, we solved the problem of what if the user collects data only from indoor or only from outdoor. In this scenario, we can determine if the user has been collecting data only from indoor or outdoor. Additionally, we solved an unsupervised classification problem, where the number of floors and the altitude of each floor were unknown. Finally, we proposed a way to map those characteristics by enhancing a CityGML building model and automated the whole process via the Node-RED. In future work, the intermediate reference pressures from each floor can improve the accuracy. Also reference pressure from multi-entrance buildings with different entrance levels is still an open question.

5 Literature

ALZANTOT, M. & YOUSSEF, M., 2012: UPTIME: Ubiquitous pedestrian tracking using mobile phones. IEEE Wireless Communications and Networking Conference (WCNC), 3204-3209.

BOETERS, R., OHORI, K.A., BILJECKI, F. & ZLATANOVA, S., 2015: Automatically enhancing citygml lod2 models with a corresponding indoor geometry. International Journal of Geographical Information Science, **29**(12), 2248-2268.

BOLLMEYER, C., ESEMANN, T., GEHRING, H. & HELLBRCK, H., 2013: Precise indoor altitude estimation based on differential barometric sensing for wireless medical applications. IEEE International Conference on Body Sensor Networks, 1-6.

GRZONKA, F., DIJOUX, A.; KARWATH, & BURGARD, B., 2010: Mapping indoor environments based on human activity, in Robotics and Automation (ICRA). IEEE International Conference on Robotics and Automation, 476-481.

KIM, J.S., YOO, S.J. & LI, K.J., 2014: Integrating IndoorGML and CityGML for Indoor Space. Springer Berlin Heidelberg, 184-196.

LAVIELLE M., 2005: Using penalized contrasts for the change-point problem. Signal Processing, **85**, 1501-1510.

LI, B., HARVEY, B. & GALLAGHER, T., 2013: Using barometers to determine the height for indoor positioning. International Conference on Indoor Positioning and Indoor Navigation, 1-7.

LOCH-DEHBI, S., DEHBI, Y. & PLMER, L., 2017: Estimation of 3d indoor models with constraint propagation and stochastic reasoning in the absence of indoor measurements. ISPRS International Journal of Geo-Information, **6**(3).

MOSLEHIRAD O., 2017: GitHub Repository. https://github.com/omidrad2017/Crowdsourced-vertical-indoor-mapping-Final.

OTSU, N., 1979: A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, **9**(1), 62-66.

PIPELIDIS G., 2017: RecordData. https://play.google.com/store/apps/details?id=com. recordData.basic&hl=en.

PIPELIDIS, G., MOSLEHIRAD, O., IWASZCZUK, D., PREHOFER, C. & HUGENTOBLER, U., 2017: A novel approach for dynamic vertical indoor mapping through crowd-sourced smartphone sensor data. International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sapporo, 1-8.

PRATT, W.K., 2007: Digital Image Processing 4th Ed. John Wiley & Sons, in Hoboken, NJ.

VIDEO 01: https://youtu.be/TPWhTybm5mA

VIDEO 02: https://www.youtube.com/watch?v=RwtW5eoYZO8&feature=youtu.be

XIA, H., WANG, X., QIAO, Y., JIAN, J. & CHANG, Y., 2015: Using multiple barometers to detect the floor location of smart phones with builtin barometric sensors for indoor positioning. Sensors, **15**(4), 7857-7877.

Hand Tracking using Kalman Filter for Safe Human-Robot Interaction

HASRET GÜMGÜMCÜ[1] & DANIEL LAUMER[1],
JAN DIRK WEGNER[1], PAUL BEARDSLEY[2] & KONRAD SCHINDLER[1]

Abstract: The goal of this bachelor thesis was to develop an algorithm to detect human activity with the use of an Intel RealSense camera and determine if the activity intrudes into a specified safety zone. The Intel RealSense is a depth camera, which means that we know the distance from the camera to the objects in the scene for every pixel. The Intel RealSense camera can also deliver tracked coordinates of hand joints which can be used as a signal of human activity. In order to be able to keep the safety zone small and at the same time reduce false positives, we use a Kalman Filter with which we are able to solve two problems at once: Predict future hand movements and at the same time filter the output to realize a smooth visualization. We were able to implement an application which is able to foresee quick hand movements and signal those in direction of the camera as potentially hazardous. The filter smooths the raw coordinates as desired and gives a nice visualization of the hand. To ensure an overall security we added another module which uses the 3D coordinates of the depth information to also check other objects than the hands against the security zone which however does not use prediction. Our developed method can be a starting point for further research purposes and a potential ground base to implement a profound safety zone for a real robot.

1 Motivation and Problem

Nowadays there is a huge change in the use and importance of vision based applications especially with automation in many big markets. Many autonomous processes need to have the ability to 'see', to conceive and understand the surroundings in order to react properly. However, there are still unsolved problems and questions in this area. Especially when we look at the research in robotics and suchlike topics, which depend on reliable vision based methods for the automation process to get along with the environment, a question arises: How can we ensure the safety of such human-robot interactions? As long as the person reacts as imagined by the programmer everything is fine, but a human is not a robot and can act unexpectedly. In this thesis we drill down on this safety problem and develop a security algorithm which protects interacting people of being hurt from a movement of the robot and at the same time the robot of getting damaged.

The initial setup for our thesis is the following: There is a robot on top of a table which is supposed to be able to interact with humans. This interaction could be for example playing board

[1] ETH Zürich, Institut für Geodäsie und Photogrammetrie, Stefano-Franscini-Platz 5, CH-8093 Zürich, Schweiz, E-Mail: hasret.g46@gmail.com, daniel.laumer@gmail.com, [jan.wegner, konrad.schindler]@geod.baug.ethz.ch
[2] Disney Research Zurich, Stampfenbachstrasse 48, CH-8006 Zürich, Schweiz, E-Mail: pab@disneyresearch.com

games. The robot has an Intel RealSense camera mounted on its top which is a so called depth camera that can deliver 3D coordinates for each pixel of the image.
The goal of this thesis is now the following:
Design an algorithm, which uses the functionalities of the Intel RealSense and filtering methods to establish a safety zone around the robot and issues a warning when something enters that zone.
Since it is a tabletop interaction, the main focus is on the hands. The Intel RealSense has a hand tracking option included, so it is possible to get the raw coordinates of 22 joints at any time a hand is in the field of view. An option would be to use these coordinates and check, if they are behind the safety zone. This approach would require a big safety zone in order to be safe, because if a hand is moving with high velocity directed to the robot, there should be enough time to shut down. Due to the fact that many hand movements will neither be directed to the robot nor reach it, it produces many false negatives. In this thesis we want to go a step further. The algorithm should distinguish if the hand is likely to enter the zone or not. To do that, a prediction of the movement is required. If the hand is still or just moving alongside the security zone, the algorithm should not issue any warning. However, if the hand movement is directed to the camera and is likely to be too close to the robot, it should shut down. As a result, the safety zone can be made much smaller and still offer the same level of security and at the same time, it allows much closer and smoother interaction with the robot itself without being interrupted with false negatives.

Fig. 1: The two possible solutions to solve the safety problem

2 Methodology

2.1 Safety Realization

To realize our application we define a safety zone. Since the field of view of the camera is limited to one direction a simple plane defined by 3 points $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$ is sufficient. To determine if any hand joint $\boldsymbol{p}$ is entering the safety zone we can use the following mathematical approach:

$$Safety\ Status = \begin{cases} 0\ for\ (\boldsymbol{p} - \boldsymbol{p}_1) \cdot \boldsymbol{n} > 0 \\ 1\ for\ (\boldsymbol{p} - \boldsymbol{p}_1) \cdot \boldsymbol{n} < \boldsymbol{0} \end{cases}$$

with 0 meaning outside and 1 meaning inside of the safety zone and $\boldsymbol{n}$ being the normal vector of the plane. To make sure that $\boldsymbol{n}$ and $\boldsymbol{p}_1$ look in the same direction we use a vector to a point far away from the origin (perpendicular to the camera sensor) and calculate its dot product with $\boldsymbol{n}$ and $\boldsymbol{p}_1$. Both results must have the same algebraic sign.
For this entire process we use the joint coordinates $\boldsymbol{p}$ of the predicted hand to detect dangerous interactions beforehand.
Although focus of our project is on safe human-robot interaction, we need to provide more than a safety status just for the predicted hand. For an overall safe environment, we add a depth map feature where any objects in the field of view range of the camera can be measured. Each pixel of the depth map goes through the same calculation as the predicted joint coordinates to find out if there is anything inside the safety zone. The only difference is that the security check is made with the real time raw data of the depth map and there is no prediction. Its function is just to have an overall secure environment.

2.2 Visualization

For the real time visualization of the application we can show the hand with the raw, predicted and filtered coordinates at the same time. As default we use the filtered hand because the visualization is smooth since that the jittery party of the raw hand is filtered. The hands are visualized by using a spherical representation for each joint with the joint coordinates as the center. A bone which connects two joints is represented by a cylinder where its size and the orientation is calculated with the coordinates of the two joints which it should connect.

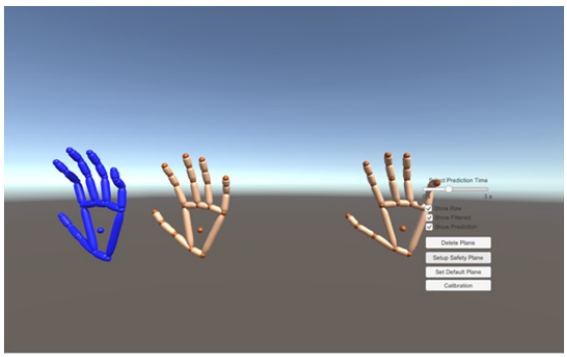


Fig. 2: Visualization of the hands, from left to right: Predicted, Filtered, Raw

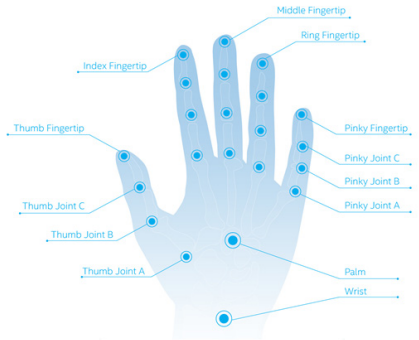


Fig. 3: All detectable joints with the Intel RealSense camera INTEL (2016)

For the depth map, there is a rectangular box in the upper left corner which shows every pixel red where its measured 3D coordinates are inside the safety zone and white otherwise (see Fig. 4).
The entire visualization of the application is made to get real-time information when implementing the software for any projects. The entire visualization is run in Unity.

2.3 Kalman Filter

To achieve the prediction of the hand we use the Kalman Filter, which was first introduced by KALMAN (1960). Kalman Filtering consists of two main parts, prediction and correction, where the state vector $\boldsymbol{x}$ describes position and velocity and P contains its variances and covariances.
The first step is the **prediction**. There we make a kind of educated guess of how the state of the system is going to be in the near future. For that we need the actual state and the prediction model. The prediction model is an equation which describes the nature of how the state of the system changes over time. It is a recursive function which is in the form:

$$\boldsymbol{x}_n = f(\boldsymbol{x}_{n-1})$$

Usually, the prediction model is a normal equation of motion by Newton. As proposed by KOHLER (1997), when tracking hands, the movement between two time steps is often assumed as linear and the acceleration is omitted and modeled as noise, so the model has the following form:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{n,apriori} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{n-1} + \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}_{n-1} \cdot \Delta t$$

Linear models like this can be directly transformed into matrix notation:

$$\boldsymbol{x}_{n,apriori} = A \cdot \boldsymbol{x}_{n-1}$$

For the predicted P matrix we perform a propagation of variance. Since the matrix $\boldsymbol{A}$ is already calculated, this is an easy task:

$$P_{n,apriori} = A \cdot P_{n-1} \cdot A^T$$

The second step is the **correction**. The prediction gets corrected by the actual measurements which are stored in the vector $\boldsymbol{z}_n$ and their errors and noise in the matrix R. The H matrix is called conversion matrix and handles the conversion of different units or scales. To perform the correction, we use a weighted mean of the measured and the predicted state:

$$\boldsymbol{x}_{n,apost} = \boldsymbol{x}_{n,apriori} + K \cdot \left(\boldsymbol{z}_n - H \cdot \boldsymbol{x}_{n,apriori}\right)$$

$$P_{n,apost} = P_{n,apriori} - K \cdot H \cdot P_{n,apriori}$$

The weight matrix K is the Kalman Gain Matrix. It is actually a kind of inverse of the ratio between the two covariance matrices of the measured and the predicted state and is calculated as follows:

$$K = \frac{P_n \cdot H^T}{H \cdot P_n \cdot H^T + R}$$

Both steps, prediction and correction, are repeated over the course of the filtering operation. In practice, all matrices could change along the process but most of them are assumed constant as proposed in (WELCH & BISHOP 2006). Only the value of $\mathbf{z}_n$ change as well as the outputs $\boldsymbol{x}_n$ and P_n.

3 Results

We implemented the developed method as a running software in C# that demonstrates the system's capabilities. Our main interface is shown in Fig. 4:

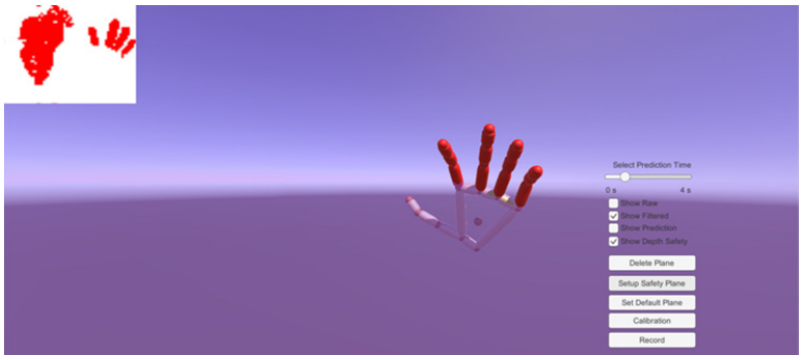


Fig. 4: Visualization of the algorithm

We represent the safety zone as a blue transparent plane and the visualized hand changes the color when its inside in order to verify it the algorithm is working. With one click you can also show the predicted hand and its intrusion into the zone. In the upper left corner, the safety status of the depth map is displayed as a small image stream.
To assess how good the filtering of the coordinates works, we chose a qualitative approach of recording some seconds and plotting the distance from the camera to one joint against the time. This way we can easily see how much noise the raw coordinates contain and what the Kalman Filter makes of this data. We can observe that the Kalman filtered outputs (red line in Fig. 5) are much smoother than the original raw measurements.

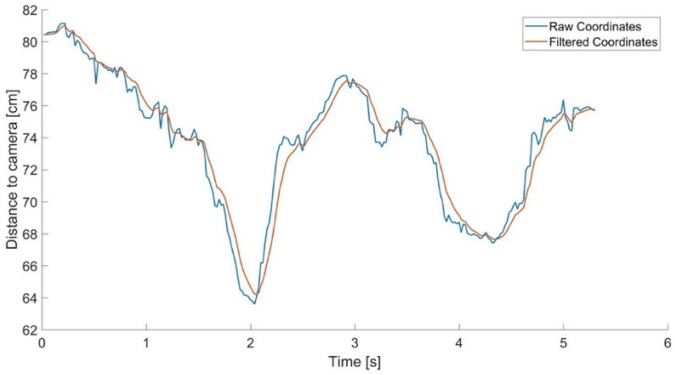


Fig. 5: Distance from the camera to the index fingertip over time, raw and filtered coordinates

We estimated measurement noise (values in R) with the empirical standard deviation of a still hand. We took the mean over all joints, which resulted in a standard deviation of 0.1 cm for the position and 0.7 cm/s for the velocity. In Fig. 6 we can see how variation of these values affects the filtering process. If the measurement errors are small, the filter stays close to the observations and there is virtually no filtering happening. With increasing uncertainty of the measurement, more and more filtering takes place.

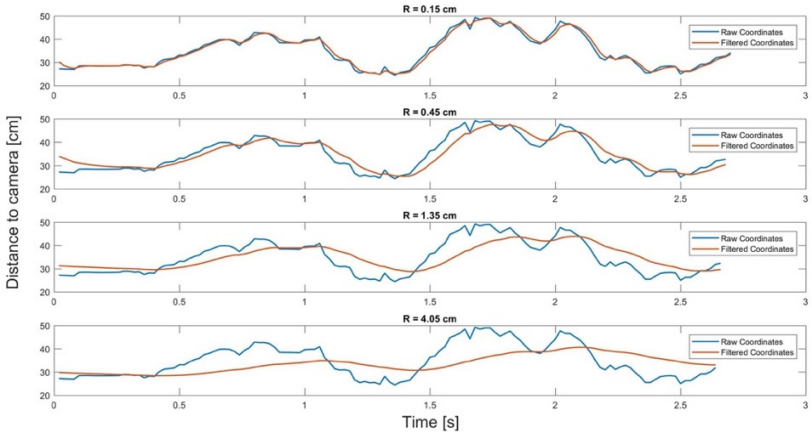


Fig. 6: Influence of the R value on the filtering process (Given values are the mean over the matrix)

4 Conclusions and Outlook

In our bachelor thesis we developed, implemented, and tested a hand-tracking tool for depth sensors. Our main methodological contribution is the adaption of a Kalman-filtering approach to human-robot interaction. Experiments with our method demonstrate that the method manages to guarantee a safety zone around the robot in order to avoid injury of humans and unnecessary damage to the robot.

In the future, we want to add additional parts like human heads, kids toys etc. to the safety zone predictor. Additionally, it is worth investigating if adding more depth sensors to the set up helps making predictions more robust. Another interesting topic would be experimenting with partial interference by daylight, which may cause the (infrared pattern-based) depth camera to fail.

5 References

INTEL, 2016: Intel RealSens SDK 2016 R2 Documentation. Intel RealSense SDK 2016 R2 Documentation.

KALMAN, R.E., 1960: A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME-Journal of Basic Engineering, **82**, Series D, 35-45.

KOHLER, M., 1997: Using the Kalman Filter to track Human Interactive Motion – Modelling and Initialization of the Kalman Filter for Translational Motion. University Dortmund, Informatics VII.

WELCH, G. & BISHOP, G., 2006: An Introduction to the Kalman Filter. University of North Carolina at Chapel Hill.

Untersuchungen zur dreidimensionalen Analyse von photogrammetrisch erfassten Betonproben

ANDREAS MAXIMILIAN PITER[1]

Zusammenfassung: Die Bestimmung der Luftporenkennwerte am Festbeton erfolgt in der Materialprüfung manuell mit einem Stereomikroskop in kleinen Bereichen des Betontestkörpers. Im Rahmen dieser Arbeit wurde die Porendetektion, ein Verfahren zur automatisierten und flächendeckenden Bestimmung der Lage von Luftporen entwickelt. Die Oberfläche des Betontestkörpers wurde mit den Methoden der Nahbereichsphotogrammetrie dreidimensional rekonstruiert. In der Porendetektion wird das Modell gefiltert, die Punktmenge anhand eines Schwellwertes klassifiziert und die klassifizierten Punkte zu Poren gruppiert. Die Validierung des Verfahrens anhand einer manuell erstellten Referenz zeigt den Einfluss der Parameter der Porendetektion auf die Ergebnisse.

1 Einleitung

Beton ist ein wichtiger Baustoff, dessen Eigenschaften bei der Herstellung durch Zugabe von verschiedenen Zusatzmitteln und Zusatzstoffen gezielt herbeigeführt werden. Dabei kann ungewollte Gasbildung die Betonfestigkeit verringern oder durch Zugabe von Luftporenbildnern der Frostwiderstand erhöht werden. In beiden Fällen muss der Luftporengehalt bereits bei der Herstellung kontrolliert werden (KAMPEN & RICHTER 2014). In Schadensfällen oder bei der Wirksamkeitsprüfung von Luftporenbildnern wird der Gehalt an Luftporen stichprobenartig am Festbeton getestet. Das in der Materialprüfung etablierte Verfahren zur Bestimmung von Luftporenkennwerten ist das Sehnenmessverfahren (BONZEL & SIEBEL 1981). Das Verfahren hat den Nachteil, dass die Bestimmung manuell erfolgt und somit nur kleine Bereiche bestimmt werden können. Deshalb ist ein Verfahren zur automatisierten und flächenhaften Detektion der Poren auf der Betonoberfläche zu entwickeln.

Poren sind Vertiefungen in der Oberfläche. Somit lässt sich anhand der Höheninformation auf Poren schließen. Zur dreidimensionalen Rekonstruktion der zu untersuchenden Betonoberfläche sind die Methoden der Nahbereichsphotogrammetrie geeignet. Für die Analyse der rekonstruierten Betonoberfläche hinsichtlich der Lage der Poren wurde keine vorhandene Funktion in gängiger Auswertesoftware gefunden. Im Rahmen dieser Arbeit wurde daher die „Porendetektion“ zur dreidimensionalen Analyse von photogrammetrisch erfassten Betonproben entwickelt. Sie wird im Folgenden vorgestellt.

[1] Student an der Leibniz Universität Hannover, Masterstudiengang Geodäsie und Geoinformatik, E-Mail: andreas.piter@online.de

2 Methodik

2.1 Photogrammetrische 3D-Rekonstruktion

Die dreidimensionale Rekonstruktion der 10 cm × 15 cm großen Oberfläche des Betontestkörpers erfolgte photogrammetrisch mit einem passiven und einem aktiven Triangulationsverfahren. Die beiden Datensätze werden im Folgenden als „Bündeltriangulation“ (passiv) und „Streifenprojektion“ (aktiv) bezeichnet. Sie dienen zum Test und zur Validierung der Porendetektion. Für weitere Informationen zu Messanordnung, Aufnahmeplanung und Genauigkeitsabschätzung wird auf PITER (2017) verwiesen.

2.2 Porendetektion

Der Fokus dieser Arbeit liegt auf der Porendetektion, die die Lage der Luftporen einer dreidimensional rekonstruierten Betonoberfläche liefert. Der Porendetektion liegt die Annahme zu Grunde, dass sich die Poren in der Höhe signifikant von ihrer Umgebung unterscheiden. Das Triangulated Irregular Network (TIN) der rekonstruierten Betonoberfläche wird zunächst gefiltert, um anschließend jeden Punkt auf einen signifikanten Höhenunterschied zu seiner Umgebung zu testen. Der Ablauf ist in Abb. 1 dargestellt.

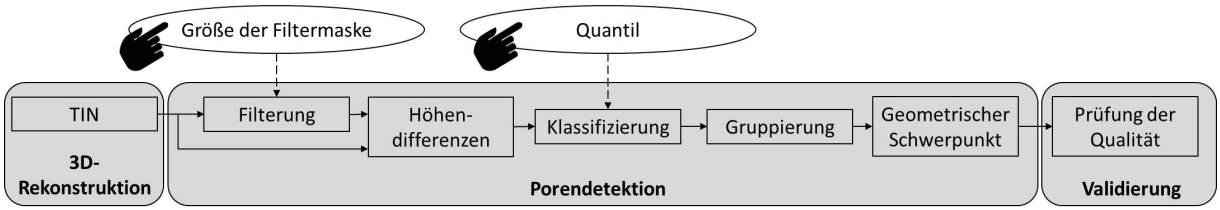


Abb. 1: Schematischer Ablauf der Porendetektion. Die Hand zeigt die vom Anwender zu wählenden Parameter.

2.2.1 Filterung

Es wird eine Kopie des TINs erstellt und geglättet (geglättetes TIN). Das Original bleibt unverändert (originales TIN). Die Umgebung, von der sich die Poren signifikant in der Höhe unterscheiden, wird durch das geglättete TIN dargestellt. Durch die Filterung wird jeder Punkt des geglätteten TINs an die Höhen in der Umgebung angepasst. Der Filter verändert die Höhe der Punkte, aber nicht die Lage. Untersuchungen haben gezeigt, dass globale Verfahren die Oberfläche schlecht approximieren. Hingegen ist ein lokaler Medianfilter für diese Anwendung am besten geeignet (PITER 2017).

Die Filteroperation wurde aus der digitalen Bildverarbeitung auf die Filterung von TINs übertragen. Hierbei ist die Größe der Filtermaske der Grad der Nachbarschaft. Alle Punkte, die mit dem betrachteten Punkt über eines der umliegenden Dreiecke verbunden sind, gehören zur ersten Nachbarschaft. Zur zweiten Nachbarschaft gehören alle Punkte, die über ein Dreieck mit der ersten Nachbarschaft verbunden sind. Weitere Nachbarschaften ergeben sich analog (siehe Abb. 2).

Der Anwender der Porendetektion greift durch die Wahl der Größe der Filtermaske in die Porendetektion ein und nimmt Einfluss auf das Ergebnis. Die Abb. 2 zeigt die Auswirkung des Parameters auf die gefilterte Oberfläche.

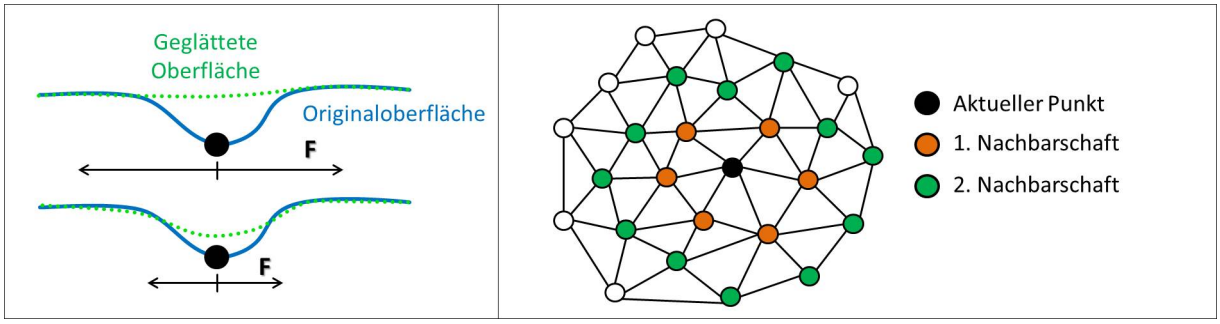


Abb. 2: Schema zur Auswirkung der Größe der Filtermaske auf die Glättung der Oberfläche im Höhenprofil (links). Schema zur Nachbarschaftsdefinition für TINs (rechts).

2.2.2 Klassifizierung

Jeder Punkt des originalen TINs wird einer von zwei Klassen zugeordnet. Alle Punkte, die zu einer Pore gehören, werden der Klasse Porenpunkt zugeordnet. Die restlichen Punkte werden der Klasse Nicht-Porenpunkt zugeordnet. Ob die Abweichung der Höhe eines Punktes von der Höhe in seiner Umgebung signifikant ist, wird auf der Grundlage der Höhendifferenzen (Gl. 1) lagegleicher Punkte des originalen und geglätteten TINs ermittelt.

$$\Delta Z_i = Z_{original,i} - Z_{geglättet,i} \quad \text{mit dem Punktindex i} \tag{1}$$

Die Häufigkeitsverteilung der Höhendifferenzen ist annähernd normalverteilt (vgl. Abb. 3). Für die meisten Punkte wird eine Höhendifferenz von $\Delta Z = 0$ erwartet. Alle Punkte, deren Höhe signifikant von der Umgebung abweicht, liegen außerhalb des Konfidenzintervalls

$$P\{\mu - k \cdot \sigma \leq \Delta Z \leq \mu + k \cdot \sigma\} = 1 - \alpha \tag{2}$$

um den Erwartungswert μ (NIEMEIER 2008, S. 82ff., 92ff.). Nach (Gl. 1) werden bei Punkten, die zu Poren gehören, lediglich negative Differenzen erwartet. Demnach erfolgt die Klassifizierung in Porenpunkt bei $\Delta Z_i < \mu - k \cdot \sigma$ und in Nicht-Porenpunkt bei $\Delta Z_i \geq \mu - k \cdot \sigma$.

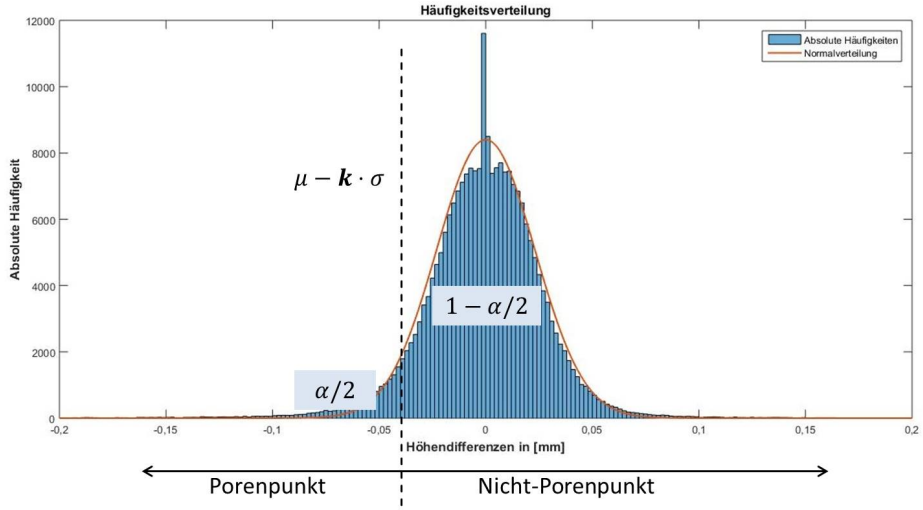


Abb. 3: Histogramm der absoluten Häufigkeitsverteilung der Höhendifferenzen, mit angenäherter Normalverteilung und dem Schwellwert zur Klassifizierung.

Das Klassifizierungsergebnis hängt von der Wahl der Irrtumswahrscheinlichkeit ab. Das damit verbundene Quantil k wird in der Porendetektion zur Bestimmung des Schwellwertes verwendet und vom Anwender gewählt.

2.2.3 Gruppierung

Eine Pore setzt sich aus mehreren Punkten der Klasse Porenpunkt zusammen. Welche Porenpunkte zusammen eine Pore bilden, wird mit einer Heuristik ermittelt (siehe Abb. 4). Die Heuristik beginnt mit einem zufällig ausgewählten Porenpunkt als Startpunkt. Ist der Startpunkt über eine Dreiecksseite mit weiteren Porenpunkten verbunden, so werden diese Punkte rekursiv geprüft, ob sie ebenfalls mit weiteren Porenpunkten über eine Dreiecksseite verbunden sind. Miteinander verbundene Porenpunkte werden zu einer Pore gruppiert. Die Gruppierung ermöglicht eine separate Betrachtung aller Poren, sodass die konvexe Hülle in der XY-Ebene berechnet werden kann. Sie stellt den Umriss der Pore dar. Der geometrische Schwerpunkt wird aus den Punkten der konvexen Hülle berechnet. Die geometrischen Schwerpunkte der Poren sind das Ergebnis der Porendetektion.

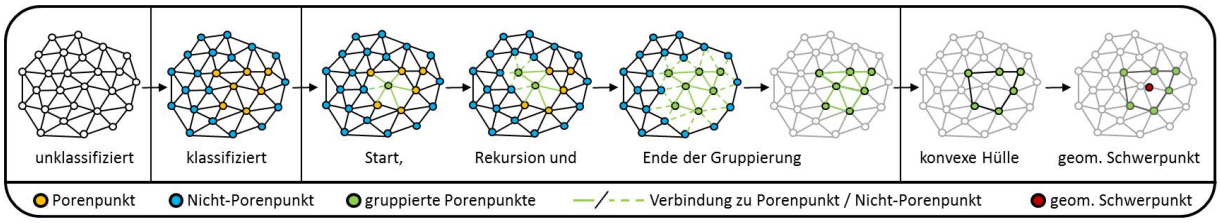


Abb. 4: Schematisches Beispiel zur Klassifizierung der Punkte im TIN, sowie der anschließenden Gruppierung und Repräsentation durch den geometrischen Schwerpunkt.

2.3 Validierung

2.3.1 Auswertestrategie

Die Porendetektion wurde mit je einer Referenz für die zwei aufgenommenen Datensätze „Bündeltriangulation“ und „Streifenprojektion“ validiert. Die Referenz entstand durch das manuelle Markieren der Poren in einem höhencodierten Rasterbild der Betonoberfläche, bei dem die Grauwerte proportional zur Höhe dargestellt werden. Die Schwerpunkte der detektierten Poren, die innerhalb eines Puffers um die jeweiligen in der Referenz enthaltenen Poren liegen, werden als True Positiv (TP) bezeichnet. Poren, die in der Referenz enthalten sind, jedoch nicht detektiert werden, heißen False Negativ (FN). Dagegen sind False Positiv (FP) die Poren, die detektiert werden, aber in der Referenz nicht vorhanden sind. Eine Bestimmung der True Negativ (TN) ist quantitativ nicht möglich.
Die Beurteilung der Porendetektion erfolgt nach HEIPKE et al. (1997) mit den Größen

$$\text{Vollständigkeit} = \frac{\text{TP}}{\text{TP+FN}} \cdot 100, \tag{3}$$

$$\text{Richtigkeit} = \frac{\text{TP}}{\text{TP+FP}} \cdot 100, \tag{4}$$

$$\text{Qualität} = \frac{\text{TP}}{\text{TP+FP+FN}} \cdot 100. \tag{5}$$

Das Ergebnis der Porendetektion ist von den zwei wählbaren Parametern „Größe der Filtermaske F“ und „Quantil k“ abhängig. Die Porendetektion wurde unter Variation der Parameter mit beiden Datensätzen mehrfach durchgeführt, um ihren Einfluss zu untersuchen.

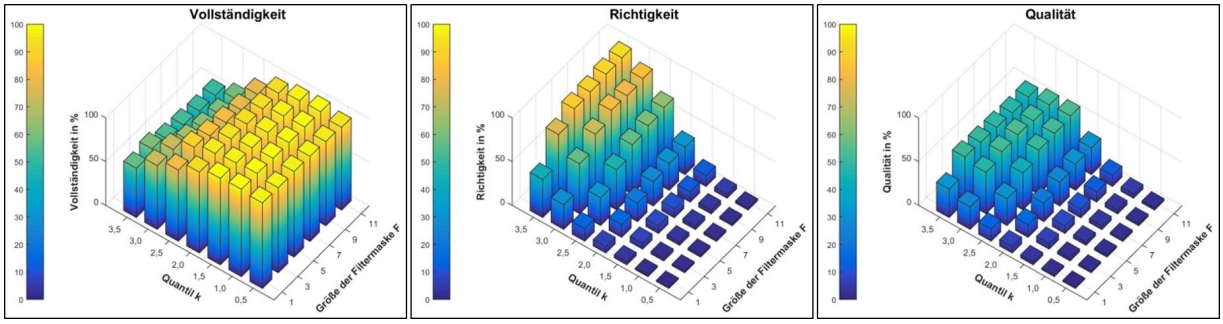


Abb. 5: Klassifizierungsergebnisse der Porendetektion in Abhängigkeit von der Größe der Filtermaske und des Quantils für den Datensatz „Bündeltriangulation“

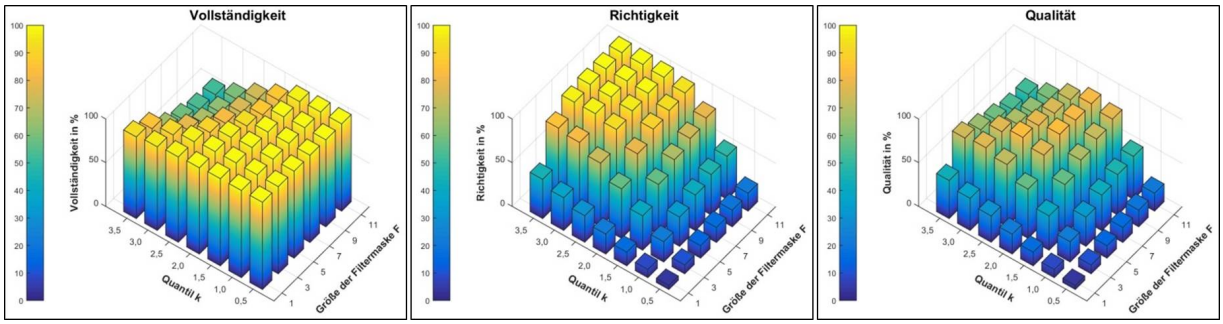


Abb. 6: Klassifizierungsergebnisse der Porendetektion in Abhängigkeit von der Größe der Filtermaske und des Quantils für den Datensatz „Streifenprojektion“

2.3.2 Ergebnisse

In den Abb. 5 und 6 sind die Vollständigkeit, Richtigkeit und Qualität für die Datensätze grafisch aufgeführt. Beide Datensätze weisen die gleichen Abhängigkeiten von den zwei Parametern auf. Auf alle drei Größen hat die Wahl des Quantils k den größten Einfluss. Je mehr Punkte zu Porenpunkten klassifiziert werden, desto höher ist die Vollständigkeit (nahe 100%). Dagegen ist dort der Anteil der richtig detektierten Poren sehr gering (nahe 0%). Mit zunehmenden Quantil k wird die Vollständigkeit geringer und die Richtigkeit höher.

Die Wahl der Größe der Filtermaske F hat einen geringeren Einfluss auf die drei Größen als das Quantil k. Jedoch nimmt die Richtigkeit gerade bei kleinen Filtermasken mit steigender Größe stark zu. Hieraus lässt sich schließen, dass die Filtermaske nicht zu klein gewählt werden sollte.

Die Qualität spiegelt die Balance aus der Vollständigkeit und Richtigkeit wider. In beiden Datensätzen ist zu erkennen, dass die Qualität bei einer bestimmten Kombination von Quantil k und Größe der Filtermaske F am höchsten ist. Der Datensatz „Streifenprojektion“ liefert mit 83,9% eine höhere maximale Qualität als der Datensatz der „Bündeltriangulation“ mit 54,5%. Ein Grund für die unterschiedliche Qualität der Ergebnisse aus den beiden Datensätzen ist die manuell erstellte Referenz. Im höhencodierten Orthomosaik aus dem

Datensatz der „Streifenprojektion“ sind die Poren aufgrund der hohen Detailauflösung und dem geringen Rauschen gut zu erkennen. Hingegen ist der Datensatz aus der „Bündeltriangulation“ stark verrauscht. Das erschwerte sowohl die automatische, als auch die manuelle Detektion der Poren.

3 Fazit & Ausblick

In dieser Arbeit wurde die Porendetektion zur dreidimensionalen Analyse von photogrammetrisch erfassten Betonproben entwickelt. Als Eingangsdaten diente ein TIN, das mit Methoden der Nahbereichsphotogrammetrie rekonstruiert wurde. Es wurde gezeigt, dass die Porendetektion die Schwerpunkte der Luftporen flächenhaft mit einer Qualität von bis zu 83,9 % liefert. Eine geeignete Kombination aus Quantil k und Größe der Filtermaske F kann iterativ, unter Berücksichtigung der aufgezeigten Abhängigkeiten und einer visuellen Evaluation der Ergebnisse, auf den jeweiligen Datensatz abgestimmt werden.
Die Porendetektion liefert die Grundlage für weiterführende Analysen, wie z.B. die Bestimmung der Abmessungen der Poren. Ausgehend vom berechneten Schwerpunkt könnte die Abgrenzung jeder Pore bestimmt und anhand des Höhenverlaufs getestet werden, ob ein für eine Pore typischer Höhenverlauf vorhanden ist. Die Richtigkeit der Porendetektion könnte durch diesen Test nachträglich gesteigert werden.
Die Automatisierung des bisher verwendeten Sehnenmessverfahrens (Bonzel & Siebel 1981) könnte Bestandteil zukünftiger Arbeiten sein.

4 Literaturverzeichnis

BONZEL, J. & SIEBEL, E., 1981: Bestimmung von Luftporenkennwerten am Festbeton. Betontechnischer Bericht Beton, **12**.

HEIPKE, C., MAYER, H., WIEDEMANN, C. & JAMET, O., 1997: Evaluation of automatic road extraction. International Archives of Photogrammetry and Remote Sensing, **32**(3) SECT 4W2, 151-160.

KAMPEN, R. & RICHTER, T., 2014: Betonzusätze – Zusatzmittel und Zusatzstoffe. Zement-Merkblatt Betontechnik B3. www.Beton.org, letzter Zugriff 09.08.2017.

NIEMEIER, W., 2008: Ausgleichungsrechnung: Statistische Auswertemethoden. 2., überarbeitete und erweiterte Auflage, de Gruyter.

PITER, A. M., 2017: Untersuchungen zur dreidimensionalen Analyse von photogrammetrisch erfassten Betonproben. Bachelorarbeit, Leibniz Universität Hannover, Deutschland.

Konzeption und Realisierung eines neuen portablen 360°-Stereokamerasystems

DAVID HOLDENER[1]

Zusammenfassung: Die Nachfrage nach der 3D-Vermessung von Innenräumen nimmt mit dem Trend der Digitalisierung in der Baubranche stetig zu. Eine effiziente Lösung zur Aufnahme des Innenbereichs ist das Mobile Mapping. Bildbasierte Systeme mit 360°-Panoramaabdeckung erlauben eine schnelle Datenerfassung und können zu georeferenzierten 3D-Bilddaten prozessiert und in cloudbasierte 3D-Geoinformationsservices integriert werden.

Das in dieser Arbeit entwickelte Kamerasystem erreicht eine 360°-Stereoabdeckung durch die ringförmige Anordnung von fünf horizontalen Stereobildpaaren. Die kostengünstige Realisierung basiert auf einer Kamerahalterung, welche mit dem 3D-Drucker gefertigt wurde, und aus Action-Kameras mit Fischaugen-Objektiv. Das Stereokamerasystem wurde erfolgreich kalibriert. Mit dem aufgezeigten Auswerteprozess können automatisch Tiefenkarten und Punktwolken aus den Stereobildern gerechnet werden. Ein Vergleich von 3D-Messungen mit Referenzdaten zeigt maximale Abweichungen von 3 cm auf typische Distanzen für Innenaufnahmen von 2-8 m. Trotz der kostengünstigen Implementierung kann so das Potential eines bildbasierten Aufnahmesystems aufgezeigt werden.

1 Einleitung

Bildbasiertes Mobile Mapping ist eine etablierte Methode zur dreidimensionalen Erfassung von Infrastrukturbauwerken wie Straßen und Bahntrassen. Um eine kontinuierliche und vollständige Abdeckung solcher komplexen Umgebungen zu gewährleisten, können die Aufnahmesysteme mit stereoskopisch ausgerichteten Panoramakameras ausgerüstet werden. Mit dem fortschreitenden Trend der Digitalisierung in der Architektur und Bauwirtschaft steigt aber auch die Nachfrage nach einer effizienten und genauen dreidimensionalen Erfassung von Innenräumen. Sei dies zur Grundlagenerfassung für eine spätere Planung oder die Baufortschrittskontrolle während der Ausführung.

Die bestehenden Systeme zur messtechnischen Erfassung von Innenräumen wie beispielsweise von BERGSLI & SCHROTH (2017) präsentiert, basieren häufig auf einer hybriden Aufnahmesensorik. Dabei wird die 3D-Information von Laserscannern erfasst und anschließend mit Bildern ergänzt. Eine Ableitung von georeferenzierten 3D-Bildern und deren Einbindung in cloudbasierte 3D-Geoinformationsservices ermöglicht auch fachfremden Anwendern Auswertungen vorzunehmen. Mittels Dense Image Matching können Tiefenwerte direkt aus Stereobilder gerechnet werden. Dies garantiert, im Gegensatz zum Laserscanning, die räumliche und zeitliche Kohärenz zwischen den Farb- und Tiefendaten eines 3D-Bildes und erübrigt eine Co-Registrierung zwischen Scannern und Kameras (NEBIKER et al. 2015).

[1] Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: david.holdener@hotmail.com

Entsprechend ist das Ziel dieser Arbeit die Entwicklung und Evaluierung eines 360°-Stereokamerasystems für flächenhafte 3D-Aufnahmen von Innenräumen und eine spätere Einbindung in ein Indoor Mobile Mapping System (IMMS). Dabei ist insbesondere den engen Platzverhältnissen und kurzen Aufnahmedistanzen bei Innenaufnahmen Rechnung zu tragen.

2 Konzeption des Aufnahmesystems

Für Außenaufnahmen existieren bereits fahrzeugbasierte Stereosysteme basierend auf mehrköpfigen Panoramakameras. Die Stereobasis wird dabei entweder vertikal (EARTHMINE 2014) oder horizontal (BLASER et al. 2017) angeordnet. Aus dem Bereich der Robotik kommt der Ansatz von omnidirektionalen Stereopanoramasystemen, bestehend aus zwei Kameras mit katadioptrischen Spiegeln (z.B. LUI & JARVIS 2010). Dabei können jedoch wie bei kommerziellen Systemen für Virtual Reality Anwendungen keine messtechnischen Genauigkeiten im cm-Bereich erreicht werden.
Aus der Präanalyse zur Erreichung einer 360°-Abdeckung mit cm-Messgenauigkeit entstanden die Anforderungen von mindestens fünf Stereobasen für eine optimale Überlappung und eine Basislänge von zirka 50 cm. Am besten kann eine horizontale ringförmige Anordnung der Stereobasen diese Anforderungen erfüllen (siehe Abbildung 1). Wird das System als IMMS über oder um den Kopf getragen, kann eine komplette 360°-Stereoabdeckung mit hoher Genauigkeit erreicht werden.

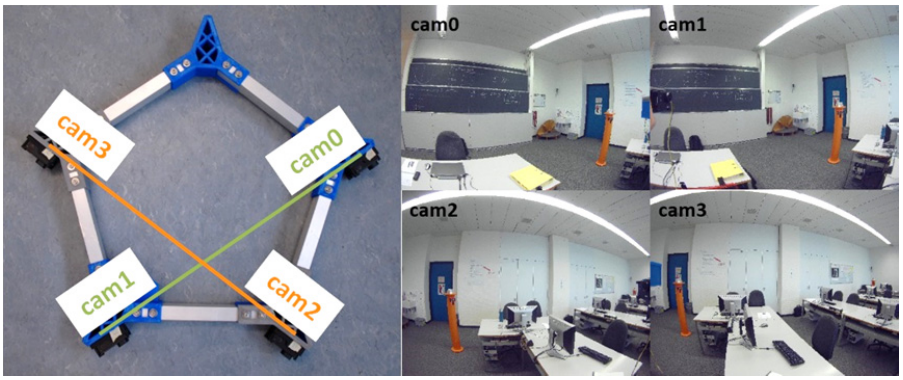


Abb. 1: Umgesetztes Kamera-Rig (links) mit den synchronen Aufnahmen von zwei der fünf Stereobildpaaren (rechts)

3 Implementierung und Kalibrierung

Die Implementierung des Systems erfolgte mit den kostengünstigen Action-Kameras Git2 von GitUp. Die Kameras haben eine Auflösung von 4608 x 3456 Pixel mit Kammerkonstanten von 3 mm bei einer Sensorpixelgröße von 1.34 µm. Dank der Fischaugen-Objektive wird ein Öffnungswinkel von 120° x 90° ermöglicht. Die Kamera-Halterungen und das Gestell wurden mittels additiver 3D-Drucktechnik hergestellt. Das Aufnahmesystem in der Ausführung, wie in Abbildung 1 ersichtlich, hat einen Gesamtdurchmesser von 70 cm und eine Basislänge von 60 cm.
Die synchrone Auslösung der Kameras erfolgt mit dem Einplatinencomputer Raspberry Pi 3 mit zugehörigem Touchscreen. Der Auslösemechanismus mit Pulsweitenmodulation (PWM) ermög-

licht eine mittlere Differenz der Synchronität zwischen den Kameras von 20 µs und maximale Abweichungen von 40 µs.

3.1 Kalibrierung

Die Kalibrierung erfolgte in einem Passpunktfeld mit Bündelblockausgleichung nach ELLUM & EL-SHEIMY (2017) und dem äquidistanten Kameramodell für Fischaugen-Objektive (ABRAHAM & FÖRSTNER, 2015). Die Implementierung konnte dabei vom Mobile Mapping System von BLASER et al. (2017) übernommen werden. Die resultierenden Parameter der Kalibrierung sind

- die inneren Orientierungsparameter (IOP) für jede Kamera;
- die relativen Orientierungsparameter (ROP) zwischen rechter und linker Kamera jedes Stereosystems;
- Boresight-Kalibrierung mit Verschiebungsvektorkomponenten (Lever Arm) und Verdrehungswinkel (Misalignment) zwischen den linken Kameras jedes Stereosystems und der Master-Kamera 0
- Äußere Orientierungsparameter (EOP) der Master-Kamera pro Epoche

Alle Parameter wurden in einer einzelnen Kalibrierung in einem Passpunktfeld (siehe Abb. 2) bestimmt. Dazu wurden in acht Epochen (32 Bilder) 3952 Bildmessung exklusive 61 Ausreißern durchgeführt. Die runden Passpunktmarken wurden mit einem Ellipsenfit nach dem Prinzip der kleinsten Quadrate angemessen.

Abb. 2: Kalibrierfeld

Die resultierenden Parameter haben allesamt eine Standardabweichung von unter 0.7 mm respektive 30 mdeg. Die vollständige Liste der erreichten Genauigkeiten ist in Tabelle 1 aufgezeigt. Auf Grund der teilweise hohen Korrelation zwischen den Parametern ist eine unabhängige Analyse der einzelnen Werte jedoch nicht sinnvoll.

Weiter wurde eine durchschnittliche Größe der Residuen von 0.7 Pixel erreicht. Die Residuen zeigen auf manchen Bildern ein systematisches Verhalten auf. Dies lässt auf den Rolling-Shutter-Effekt der Kameras oder eine ungenügende Synchronität schließen. Zudem nehmen die Residuen mit zunehmendem Bildradius zu, da dort die Bildverzerrungen grösser sind und im Fall der günstigen Action-Kameras nicht mehr so gut mit dem mathematischen Modell übereinstimmen.

Tab. 1: Empirische Genauigkeiten der Kalibrierung

Parameter	Kamera	St. Abw. der Lage [mm]			St. Abw. der Orientierung [mdeg]		
		X	Y	Z	ω	φ	κ
ROP Basis 1	0 - 1	0.3	0.3	0.6	16	24	3
ROP Basis 2	2 – 3	0.4	0.3	0.7	12	19	2
Boresight	0 - 2	0.5	0.3	0.5	31	20	28
EOP (mittel 8 Epochen)		0.4	0.4	0.4	16	12	11

3.2 Auswerteprozess

Die Stereobildauswertung konnte ebenfalls von BLASER et al. (2017) adaptiert werden. Das Ziel der Prozessierung sind georefernzierte 3D-Bilder im äquidistanten Modell und zugehörigen Tiefenwerten, im Optimalfall für jedes Pixel. In einem ersten Schritt werden dazu die Bilder nach den IOP und ROP korrigiert und mittels Epipolar-Rektifizierung in den Stereonormalfall gebracht. Diese Modellkonvertierung zur Normalisierung erfolgt nach dem epipolar äquidistanten Modell von ABRAHAM & FÖRSTNER (2005). Diese Konvertierung ist Voraussetzung für das anschliessende Dense Image Matching mit der Software SURE (ROTHERMEL et al., 2012). Die Disparitätskarte und deren Werte werden dann wieder in das äquidistante Projektionsmodell rückkonvertiert und stimmen so mit den Farbbildern überein. Aus diesen 3D-Bildern ist auch eine Ableitung der Punktwolke möglich.

4 Systemevaluierung

Zur Evaluierung der relativen Genauigkeit wurden manuelle Messungen im Passpunktfeld durchgeführt. Zudem wurden die abgeleiteten Punktwolken mit einem Referenzscan von einem terrestrischen Laserscanner (TLS) verglichen.

4.1 Manuelle Punktmessung

Mit Bildern des Passpunktfelds (Abb. 2) wurde die relative Genauigkeit innerhalb des Stereobildpaares untersucht. Die Bildmessungen fanden dabei in den nach IOP und ROP korrigierten Bildern mit einem Ellipsenfit nach den kleinsten Quadraten statt. Zwei Passpunkte formen zusammen eine 3D Referenzdistanz die als fehlerlos betrachtet werden kann. Bei einem Vergleich der photogrammetrisch gemessenen Distanz mit deren Referenz beeinflussen somit nur die Unsicherheiten von IOP, ROP und der Bildmessung sowie die Güte des mathematischen Modells die erreichbare Genauigkeit.

Auf typische Distanzen für Innenaufnahmen von 2 - 8 m wurden mit der 60 cm langen Basis in drei Epochen 78 3D-Distanzen gemessen. Die resultierende mittlere Differenz zwischen der photogrammetrischen bestimmten Distanz und Referenzdistanz beträgt 5.7 mm mit einer Standardabweichung der Differenzen von 5.6 mm (1σ). Davon ausgeschlossen sind zwei Ausreißer grösser als 70 mm. Die verbleidende maximale Abweichung beträgt 33 mm.

Der Einfluss der Distanz zwischen Kamera und Messobjekt sowie des Bildradius auf die resultierende Messgenauigkeit wurde vertieft analysiert. Auf Grund der schlechter werdenden Schnittgeometrie mit zunehmender Messdistanz wird auch eine Zunahme der Messunsicherheit erwartet. Das durchgeführte Experiment zeigt jedoch lediglich eine leichte nicht signifikante Abnahme der Messgenauigkeit auf (Abb. 3, links). Eine Abnahme der Messgenauigkeit wird im Falle von

Fischaugen-Objektiven auch mit zunehmendem Bildradius erwartet. Dies gilt insbesondere für Action-Kameras mit kostengünstiger Optik auf Grund der zunehmenden Diskrepanz zwischen mathematischem Modell und effektiven Eigenschaften des Objektivs. Die beobachteten Differenzen innerhalb dieses Experiments zeigen eine Zunahme der Messunsicherheit mit zunehmendem Bildradius auf (Abb. 3, rechts).

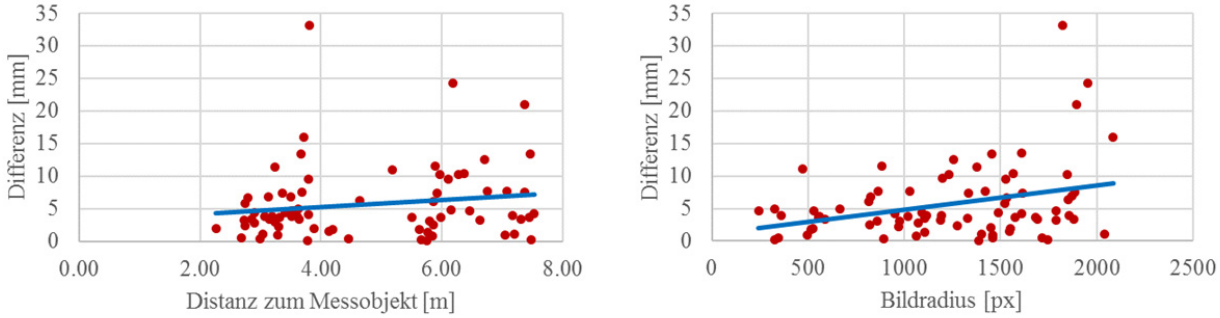


Abb. 3: Differenzen zwischen photogrammetrisch bestimmten 3D-Distanzen und Referenzdistanzen bezüglich der Distanz zum Messobjekt (links) und dem Bildradius (rechts)

4.2 Vergleich der Punktwolken

Zur Evaluierung des gesamten Auswerteprozesses wurde die Punktwolke des Stereosystems mit einer Referenzpunktwolke von einem TLS verglichen. Der Referenzscan wurde mit einer Leica P20 vom gleichen Standort wie die Kameraaufnahmen erfasst. Die Koregistrierung beider Punktwolken erfolgte mit dem ICP-Algorithmus. Die Testszene repräsentiert einen typischen Innenraum mit texturierten und untexturierten Flächen, Möblierung und Aufnahmedistanzen von 1.5 - 7 m. Der Vergleich zeigt erhebliche Differenzen im Bereich von ebenen und untexturierten Flächen wie Wänden und Decken (Abb. 4). Auf Grund der fehlenden Merkmale für eine saubere Zuordnung von Verknüpfungspunkten ist dies ist ein typisches Problem des Dense Image Matchings.

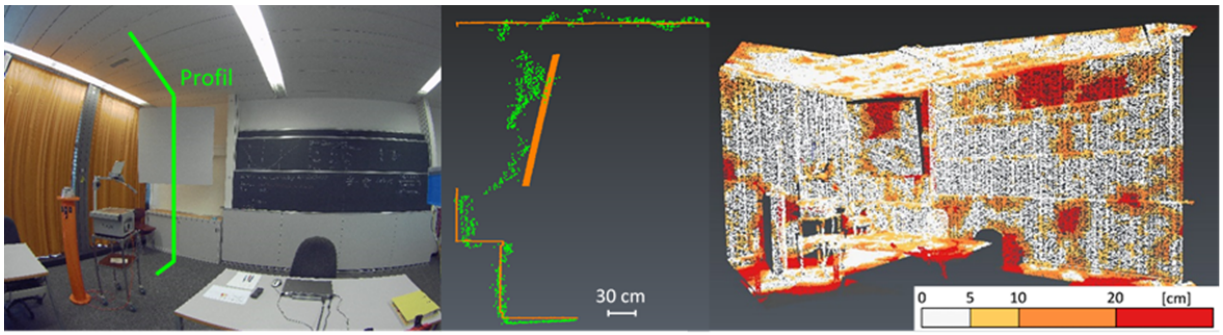


Abb. 4: Testszene der Punktwolkenevaluierung mit einem Profil und einem flächenhaften Vergleich

Weiter konnten auch in der Punktwolke systematische Abweichungen mit zunehmendem Bildradius festgestellt werden. Dank dem weiten Öffnungswinkel der Kameras sind die Überlappungen zwischen den Stereobereichen jedoch groß genug um die signifikant abweichenden Bereichen an den Bildrändern jeweils zu ersetzen.

5 Fazit & Ausblick

Die präsentierte Realisierung eines bildbasierten 3D-Aufnahmesystems bestehend aus einem Ring mit fünf horizontal angeordneten Stereobasen kann Innenräume mit einer 360°-Abdeckung dreidimensional erfassen. Die kostengünstige Realisierung erfolgte mit einem Rig aus dem 3D-Drucker, Action-Kameras (GitUp Git2) mit Fischaugen-Objektiven und einem Einplatinencomputer (Raspberry Pi 3) mit dazugehörigem Touchscreen. Die Kalibrierung des Systems mit zwei der fünf Stereobasen erzielte Genauigkeiten der ROP von unter 0.7 mm respektive 30 mdeg. Der Auswerteprozess für Fischaugen-Stereobildpaare nach ABRAHAM & FÖRSTNER (2005) wurde erfolgreich auf die Aufnahmen des Prototyps angewandt.
Die Nachteile der kostengünstigen Umsetzung mit Action-Kameras sind der Rolling-Shutter der Kameras, die ungenaue Synchronisierung und die teilweise großen Differenzen der Optik zum mathematischen Modell. Nichtsdestotrotz zeigte die Evaluierung mit Genauigkeiten im cm-Bereich das Potential des Systems für Innenaufnahmen auf. Die Tiefenkarten und Punktwolken zeigen typische Abweichungen für Bereiche mit großem Bildradius. Dank einer genügenden Überlappung können solche Bereiche jedoch ignoriert werden.
Durch die Integration von Positionierungssensorik könnte das präsentierte Aufnahmesystem für weiterführende Analysen als ganzheitliches IMMS evaluiert werden. Eine Lösung für die vielen untexturierten Flächen ist ebenfalls Teil der ausstehenden Untersuchungen.

6 Literaturverzeichnis

ABRAHAM, S. & FÖRSTNER, W., 2005: Fish-eye-stereo calibration and epipolar rectification. ISPRS Journal of Photogrammetry and Remote Sensing, **59**(5), 278-288.

BERGSLI, M. & SCHROTH, G., 2017: NavVis - Enabling Digital Value Creation Indoors. FIG Working Week. Helsinki, Finnland.

BLASER, S., NEBIKER, S. & CAVEGN, S., 2017: System Design, Calibration and Performance Analysis of a Novel 360° Stereo Panoramic Mobile Mapping System. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS Workshop. Hannover, Deutschland, 207-213.

EARTHMINE, 2014: earthmine Mars Collection System. http://www.earthmine.com/html/ products_mobile.html, [Zugriff vom 09.11.2017].

ELLUM, C. & EL-SHEIMY, N., 2002: The calibration of image-based mobile mapping systems. Proceedings of the 2nd Symposium on Geodesy for Geotechnical and Structural Engineering. The International Association of Geodesy (IAG). Berlin, Deutschland.

LUI, W.L.D. & JARVIS, R., 2010: Eye-Full Tower: A GPU-based variable multibaseline omnidirectional stereovision system with automatic baseline selection for outdoor mobile robot navigation. Robotics and Autonomous Systems, **58**(6), 747-761.

NEBIKER, S., CAVEGN, S. & LOESCH, B., 2015: Cloud-Based Geospatial 3D Image Spaces – A Powerful Urban Model for the Smart City. ISPRS International Journal of Geo-Information, **4**(4), 2267-2291.

ROTHERMEL, M., WENZEL, K., FRITSCH, D. & HAALA, N., 2012: SURE: Photogrammetric Surface Reconstruction from Imagery. Proceedings LC3D Workshop. Berlin, Deutschland.

Einsatz von Deep Learning zur Aktualisierung der Arealstatistik der Schweiz - erste Untersuchungen

MARKUS SCHÄR[1]

Zusammenfassung: Das Bundesamt für Statistik (BFS) erstellt periodisch die Arealstatistik über das gesamte Gebiet der Eidgenossenschaft. Da der personelle Aufwand für dessen Erfassung sehr hoch ist, ist der Bund auf der Suche nach einer neuen teil- oder vollautomatisierten Erfassungsmethode. In dieser Arbeit wurde mit Hilfe eines Deep Learning Frameworks, einer aus dem Machine Learning stammenden Methode, die Machbarkeit einer automatisierten Erfassung evaluiert. Dies erfolgte mit Hilfe der Ausarbeitung (Konzept, Implementierung) eines Workflows, sowie erster Evaluierungen anhand von vorliegenden Daten der Arealstatistik des BFS. Untersucht wurden insbesondere der Einfluss der Architektur und der Parametrisierung der für Deep Learning eingesetzten neuronalen Netze sowie die Eignung der Luftbilder als Grundlage. Erste Ergebnisse zeigten auf, dass sich durch eine Basiskonfigurierung der verwendeten neuronalen Netze bereits gute Ergebnisse erzielen lassen, die jedoch hinsichtlich einer zuverlässigen automatisierten Erfassung auf ein Optimierungspotenzial der Methodik hinweisen.

1 Einleitung

Ein Bild sagt mehr als tausend Worte. In den 1920er Jahren wurden in der Schweiz die ersten Luftbilder erstellt. Dadurch erhielt man einen neuen Blickwinkel auf die Landschaft Schweiz und somit auch neue Informationen, die sich zur Generierung von Landeskarten nutzen ließen. Der Mensch erkannte mit Leichtigkeit auf den Luftbildern Häuser, Flüsse, Wälder und andere Areale. Im Zeitalter der Digitalisierung hat dieser Ansatz der Arealstatistik neue Bedeutung gewonnen, indem versucht wird, einer Maschine die Klassifizierungsfähigkeit spezifischer Areale zu übertragen.

Die vom Bundesamt für Statistik (BFS) periodisch im Abstand von 12 Jahren erstellte Arealstatistik dient zur Erfassung von Veränderungen der Landnutzung in der Schweiz. Diese wird anhand von flächendeckenden Luftaufnahmen durchgeführt, indem zwei Mitarbeiter des BFS die Landnutzung mit einem Stichprobenraster von 100m Maschenweite in 72 Klassen kategorisieren (ca. 4.1 Mio. Stichprobenpunkte mit einer Abdeckung von je 1ha) und diese dem Kategorienkatalog GEOSTAT zuteilen. Eine Kontrolle der erzielten Kategorisierung erfolgt anhand eines reduzierten Datensatzes, der von einer unabhängigen Person begutachtet und im Falle von Unstimmigkeiten vor Ort abgeglichen wird. Um den Aufwand und die Aktualisierungszeitspanne zu reduzieren, sollen Verfahren untersucht werden, mit denen sich die Arealstatistik des BFS automatisieren lässt. Es handelt sich dabei um eine *Big Data* Problemstellung, für dessen Lösung sich „State of the Art“ Verfahren des maschinellen Lernens anbieten und nachfolgend als Deep Learning Framework umgesetzt werden.

[1] Fachhochschule Nordwestschweiz, Institut für Vermessung und Geoinformation, Gründenstrasse 40, CH- 4132 Muttenz, E-Mail: markus.schaer@g2014.ch

Als Bilddatenbank eignet sich das Schweizerische Orthophotomosaik, welches durch das Bundesamt für Landestopografie zur Verfügung gestellt wird und die Schweiz in zwei Bodenauflösungs-Gebiete aufteilt (Abb. 1).

Abb. 1: Aufteilung der GSD – Gebiete (BUNDESAMT FÜR LANDESTOPOGRAFIE 2017)

Je nach Region beträgt die Bodenauflösung 25cm (Abb. 1 violett) oder 50cm (Abb. 1 grün). Die Standardabweichung (1σ) der Lagegenauigkeit von SWISSIMAGE für beide GSD liegt bei einem Pixel. Beim aktuellen Datensatz wurden vier Spektralkanäle erfasst. Rot, Grün. Blau und Nahes Infrarot.

2 Methodik

Deep Learning ist ein Subthemengebiet von Artificial Intelligence und Machine Learning. Artificial Intellegence, dessen Wurzeln bis 1950 zurückreichen hat durch die wachsende Rechenleistung von Maschinen und der damit verbundenen Möglichkeit des Deep Learnings einen enormen Aufschwung erhalten.

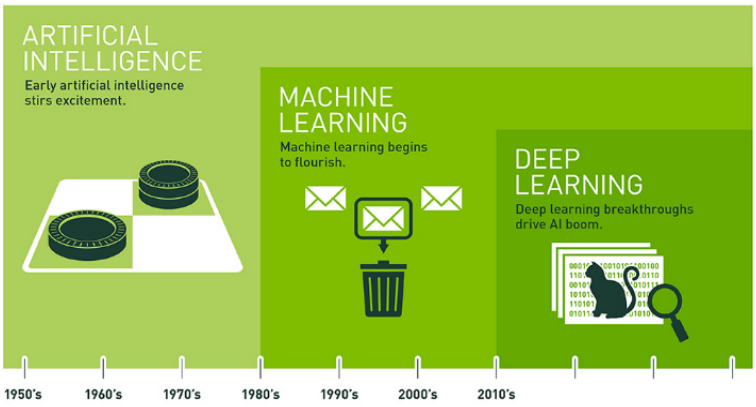


Abb. 2: Hierarchy of AI, Machine Learning and Deep Learning (COPELAND 2016)

Der oft benutzte, umfassende Begriff des maschinellen Lernens, deckt bereits einfachere der Maschine vorgegebenen Aufgaben ab. Ein klassisches Beispiel für eine einfache Anwendung des maschinellen Lernens ist die SPAM – Regel, welche Nachrichten beispielsweise anhand des Betreffs automatisch vorsortiert.

Deep Learning, auf Deutsch tiefergehendes Lernen, ist dahingehend entwickelt worden, menschliche Wahrnehmungsmechanismen maschinell anzuwenden, indem Netzwerkstrukturen des zentralen Nervensystems nachgebildet werden. Die Methodik hat sich zur automatisierten Bearbeitung von Bildinformationen als besonders geeignet erwiesen und ermöglicht mit Hilfe eines «synthetischen Nervensystems» die überwachte oder unüberwachte Kategorisierung von Daten.
Grundsätzlich besteht ein für Deep Learning verwendetes neuronales Netz aus drei Hauptelementen. Den funktional getrennten Layern, den darin enthaltenen Neuronen und den übergreifenden Verbindungen.

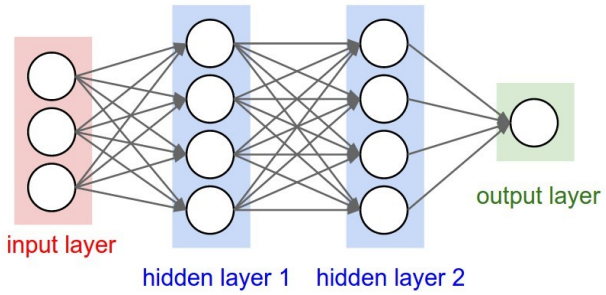


Abb. 3: 3-Layer Netzstruktur (STANFORD UNIVERSITY 2017)

Nach REY & WENDER (2010) werden die Layer in drei Bereiche aufgeteilt. Dem Netz werden wiederholt über den Input Layer (Abb. 3 rot) Informationen in Form von Zahlen übergeben. Diese Informationen werden in den Hidden Layern (Abb. 3 blau) verarbeitet und weitergegeben. Der Output Layer (Abb. 3 grün) führt anhand der erhaltenen verarbeiteten Informationen die Kategorisierung durch.
In den Layern befinden sich die sogenannten Neuronen. Diese führen die eigentliche Modifikation der Eingabewerte durch. Dabei spezialisieren sich die einzelnen Neuronen im Rahmen der vorgegebenen mathematischen Parameter auf bestimmte Merkmale des Gesamtbildes. Diese Merkmale verarbeiten die Neuronen in neue Merkmale welche wiederum in Zahlenform über die Verbindungen an neue Neuronen weitergegeben werden. Die Extraktion solcher Merkmale basiert beispielsweise auf vorgegebenen Filtern, Kantenoperatoren oder Faltungen.
Die Neuronen werden durch die Verbindungen Layer-übergreifend verbunden. Jede Verbindung besitzt eine Gewichtung. Damit ein neuronales Netz und somit der Deep Learning Algorithmus eingesetzt werden kann, muss das Netz «angelernt» werden. Dazu werden anhand eines bereits kategorisierten Datensatzes dem neuronalen Netz der Input sowie der Output bereitgestellt. Dies ist ein Trainingsprozess welcher idealerweise eine große Anzahl an unabhängigen und repräsentativen Daten erfordert und dabei die Gewichte der Verbindungen solange anpasst, bis das Netz zwischen In – und Output ein möglichst gutes Bindeglied bildet. Die Gewichte der Verbindungen repräsentieren das «Wissen» eines Deep Learning Algorithmus. Die Berechnung der Gewichte geschieht in der Regel mithilfe des iterativen Gradientenabstiegsverfahrens (Delta-Regel), der zu einer Minimierung von Fehlklassifizierungen führt.

3 Implementation

Die Implementierung wurde in drei Schritte aufgeteilt, der Datenaufbereitung, der Lern- und Auswertungsphase des Deep Learning Algorithmus und der Auswertung (Abb. 4). Sämtliche Schritte wurden mithilfe der Programmiersprache Python (Python Software Foundation, V. 3.5.3., Beaverton OR) implementiert. Mit Hilfe der High – Level API „Keras“ und dem von Google entwickeltem Deep Learning Framework „TensorFlow“ wurde das neuronale Netz implementiert und für die Aufgabenstellung trainiert.

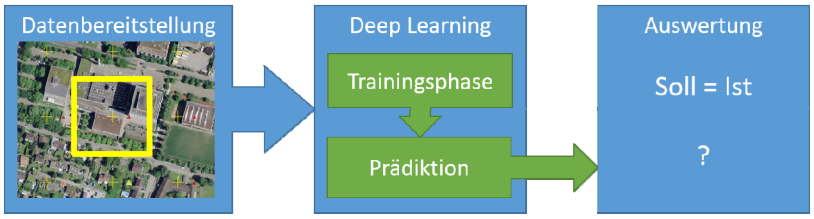


Abb. 4: Generalisierter Workflow

In einem ersten Schritt wurden sämtliche Stichprobenpunkte der Arealstatistik über den Raum mit den Luftbildern verknüpft. Zu jedem Stichprobenpunkte wird zentrisch ein quadratischer Bildausschnitt aus den Luftbildern ausgestanzt. Der Bildausschnitt wird als Input dem Deep Learning Algorithmus übergeben und die bekannte Sollkategorie aus einer bestehenden Arealstatistik als Output definiert.
In einem zweiten Schritt wird das neuronale Netz für die Aufgabenstellung trainiert. Die Daten werden für TensorFlow aufbereitet und mit der Schnittstelle Keras übergeben. Nach der Initialisierung des neuronalen Netzes, werden die Daten nacheinander an das Netz übergeben.
Zuletzt werden die berechneten Gewichte der Knoten fixiert und ein unabhängiger Datensatz dem Netz zur Prädiktion übergeben. Dabei wird der Output dem Netz selbst überlassen. Das Ergebnis wird anschließend mithilfe der aus der Fernerkundung bekannten Fehlermatrix mit der der Sollkategorie verglichen.

4 Ergebnis

Der Kategoriekatalog der Arealstatistik besitzt unterschiedliche Stufen mit bis zu 72 unterschiedlichen Kategorien. Durch die detailierte Unterteilung der schweizerischen Landschaft entstehen stark unter- und überrepräsentierte Kategorien.

Tab. 1: Über- / unterrepräsentierte Kategorien, Stufe 72

Kategorie	**Anteil unausg.**	**Anteil ausg.**	**Kategorie**	**Anteil unausg.**	**Anteil ausg.**
Normalwald	25 %	6 %	Reihen und Terrassenhäuser	0.03 %	0.09 %
Ackerland	17 %	5 %	Lawinen- und Steinschlagverbauungen	0.02 %	0.11 %
Günstige Alpweiden	10 %	6 %	Flugplätze	0.02 %	0.08 %
Naturwiesen	7 %	5 %	Abwasserreinigungsanlage	0.02 %	0.04 %
Stehendes Gewässer	5 %	6 %	Landschaftseingriffe	0.01 %	0.04 %
Gesamtgenauigkeit	36.5 %	37.0 %	**mittlere Klassengenauigkeit**	24.5 %	30.7 %

In Tabelle 1 sind die fünf Kategorien mit höchster und kleinster Repräsentation aufgeführt. Drei der 72 Kategorien machen zusammen mehr als die Hälfte aller Datenpunkte aus. Bei der Auswertung stellte sich heraus, dass unterrepräsentierte Kategorien eine deutlich schlechtere Klassifikationsgenauigkeit aufweisen.
Um den Einfluss der Verteilung der Kategorien zu untersuchen wurde ein Datensatz erstellt, bei dem die Verteilung ausgeglichener ist. Dadurch blieben die Klassifikationsergebnisse jedoch beinahe gleich, sowohl für die einzelnen Kategorien als auch in der gesamten Auswertung (Gesamtgenauigkeit). Verbessert hat sich aber die mittlere Klassengenauigkeit welche auf eine bessere Genauigkeitsverteilung unter den Kategorien hinweist. Insbesondere unterrepräsentierte Kategorien welche sich durch individuellere Merkmale von anderen Kategorien abheben, werden deutlich besser kategorisiert.
Teils sehr ähnlichen Merkmale in unterschiedlichen Kategorien führen zu folgender Problematik.

Abb. 5a & 5b: Öffentliches Gebäude (links), Industrie – Gewerbe (rechts)

Eine Differenzierung zwischen den Abbildungen 5a und 5b auf alleiniger Basis des Luftbildes durchzuführen, ist nicht eindeutig möglich. Es fehlen eindeutige Merkmale. Dennoch werden die Bilder in der Arealstatistik des BFS unterschiedlichen Kategorien zugeordnet. Solche Situationen scheinen den Deep Learning Algorithmus zu überfordern.

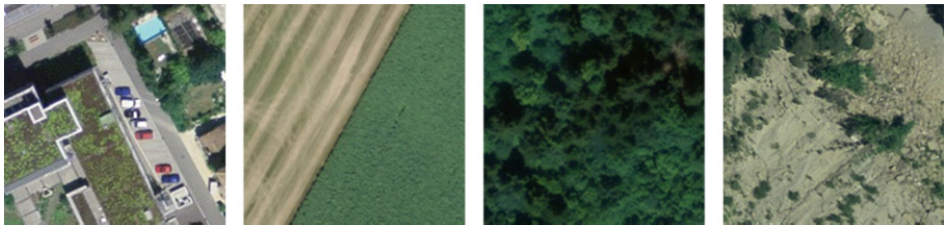


Abb. 6a–6d: v.l.n.r.: Siedlungsflächen, Landwirtschaftsflächen, Bestockte Flächen, Unproduktive Flächen auf Stufe 4 Kategorien

Sind Kategorien wie in Abbildung 6 mit ähnlichen Merkmalen zusammengefasst, resultiert eine klarere Kategorisierung.
Dennoch ließen sich mit Hilfe einer Parameteroptimierung am Deep Learning Algorithmus folgende Ergebnisse erzielen (Tab. 2).

Tab. 2: Resultierte Gesamtgenauigkeiten

Anzahl Kategorien:	4	17	27	72
Gesamtgenauigkeit:	83%	63%	59%	56%

5 Diskussion

Die große Hürde eine automatisierte Erfassungsmethode für die Arealstatik zu finden, welche mit den vorabgehenden menschlich interpretierten Datensätzen harmonisierbar ist, kann mit Deep Learning gemeistert werden. Die Datenverarbeitung ist mit einem Deep Learning Framework und mit Hilfe von High – Level API's ohne stark vertiefte Kenntnisse implementierbar. Die ersten erhaltenen Ergebnisse überzeugen und bestätigten den erarbeiteten Ansatz.
Bei der Wahl des neuronalen Netzes und den damit verbundenen benötigten Parametern sind gewisse Kenntnisse über die Methodik unerlässlich. Es hat sich gezeigt, dass Parameter sowie die Lernregel einen starken Einfluss auf das Resultat haben.
Ein ausgeglichener Datensatz ohne starke über- oder unterrepräsentierte Kategorien scheint eine wichtige Voraussetzung zu sein, um ein Netz effizient trainieren zu können. Dabei ist ein Zusammenspiel zwischen der Anzahl Datensätze und den Merkmalen der jeweiligen Kategorie feststellbar. Kategorien welche sich anhand ihrer Merkmale vom restlichen Datensatz abgrenzen, benötigen weniger Trainingsdaten und können trotz einer Unterrepräsentierung klassiert werden. Um ein besseres Ergebnis zu erhalten, müsste der Datensatz noch weiter optimiert werden.
Das BFS benutzt für die Klassierung der Arealstatistikpunkte teilweise zusätzliche Informationen aus anderen Quellen. Diese zusätzlichen Informationen fehlen dem Deep Learning Framework um zuverlässig eine Klassifikation durchführen zu können. Die Anzahl an Informationen welche dem Framework übergeben werden können ist grundsätzlich unbegrenzt. Damit ähnliche Kategorien zukünftig getrennt werden könnten, müsste geprüft werden, welche Informationen neue Merkmale hinzufügen würden und wie diese an das Deep Learning Framework übergeben werden können.

6 Literaturverzeichnis

BUNDESAMT FÜR LANDESTOPOGRAFIE, 2017: SWISSIMAGE - Das digitale Farborthophotomosaik der Schweiz. https://www.swisstopo.admin.ch/content/swisstopo-internet/de/ home/products/images/ortho/swissimage/_jcr_content/contentPar/tabs/items/dokumente/tabPar/downloadlist/downloadItems/588_1464190449870.download/infosi201003deu.pdf, [7.4.2017].

COPELAND, M., 2016: The Difference Between AI, Machine Learning, and Deep Learning? |NVIDIA Blog. The Official NVIDIA Blog. https://blogs.nvidia.com/blog/2016/07/29/ whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/, [21.7.2017].

REY, G.D. & WENDER, K.F., 2010: Neuronale Netze: Eine Einführung in die Grundlagen, Anwendungen und Datenauswertung. Bern: Hogrefe, vorm. Verlag Hans Huber, 15.

STANFORD UNIVERSITY, 2017: CS231n Convolutional Neural Networks for Visual Recognition. Stanford University. http://cs231n.github.io/neural-networks-1/, [26.7.2017].

Bei sämtlichen aufgeführten Bildausschnitten aus dem Luftbilddatensatz SWISSIMAGE liegt das Urheberrecht beim Bundesamt für Landestopografie (swisstopo, Vertrag JA100071).

Klassifikation von Fahrzeugen aus RGB- und LiDAR-Daten mit Convolutional Neural Networks

ROBIN NIESSNER[1], HENDRIK SCHILLING[2], STEFAN HINZ[1] & BORIS JUTZI[1]

Zusammenfassung: In den letzten Jahren wurde ein signifikanter Fortschritt in der Detektion, Identifikation und Klassifikation von Objekten und Bildern durch Convolutional Neural Networks (CNNs) gemacht. In dem vorliegenden Beitrag wird die Eignung von CNNs zur Analyse von RGB und LiDAR Fernerkundungsdaten anhand des Beispiels der Fahrzeugklassifikation untersucht. Hierfür werden drei unterschiedliche Methoden zum Trainieren von CNNs analysiert, die es ermöglichen, CNNs auf Datensätzen mit wenigen hunderten Trainingsbeispielen zu trainieren. Zwei der vorgestellten Methoden basieren auf dem Konzept des Transferlernens. Merkmale, die mit Hilfe eines vortrainierten CNNs erstellt werden, werden bei der ersten Methode mit einer Support Vector Machine (SVM) klassifiziert. Bei der zweiten Methode wird ein CNN' auf einen anderen Datensatz angepasst und ein Teil des CNN wird neu trainiert. Bei der dritten Methode werden eigene flache neuronale Netze designt und trainiert. Die Evaluation der Methoden erfolgt auf einem von der IEEE Geoscience and Remote Sensing Society (GRSS) bereitgestellt Datensatz aus kombinierten RGB und LiDAR Aufnahmen eines urbanen Bereiches. Es wird gezeigt, dass mit diesen Methoden sowohl auf RGB als auch auf LiDAR Daten Klassifikationsergebnisse von hoher Genauigkeit erzielt warden können. Anhand der Ergebnisse wird deutlich, dass durch Methoden des Transferlernens auf RGB Daten trainierte Merkmale auf LiDAR Daten übertragbar sind. Mit Hilfe dieser Merkmale werden auf LiDAR Daten genauere Klassifikationsergebnisse erzielt als auf RGB Daten. Die besten Klassifikationsergebnisse werden mit den selbst designten und trainierten neuronalen Netzen erreicht, welche aus deutlich weniger Ebenen aufgebaut sind als gängige tiefe neuronale Netze. Auch hier werden mit LiDAR Daten höhere Klassifikationsgenauigkeiten erreicht als mit RGB Daten.

1 Einleitung

Verkehrsbezogene Daten sind das zentrale Thema für die Überwachung und das Monitoring urbaner Regionen. Daher ist die automatisierte Analyse dieser Daten, um eine eine parametrisierte Charakterisierung abzuleiten, unerlässlich. Typische Parameter sind beispielsweise die Fahrzeugpositionen, die Anzahl der Fahrzeuge, die Verkehrsdichte und der Verkehrsfluss. Zur Erkennung und Klassifizierung von Fahrzeugen kann der Gradient zwischen Fahrzeug und Hintergrund hilfreich sein, da er häufig eine starke Kennlinie aufweist. Daher werden auf Gradienten basierende Algorithmen verwendet, um Fahrzeuge in Bildern zu bestimmen, z.B. durch Histogram of Oriented Gradients (DALAL & TRIGGS 2005), Local Binary (OJALA et al., 1994) oder SIFT Descriptors (LOWE 1999). Um die Klassifikationsleistung zu erhöhen und die Falschalarmrate zu reduzieren, können zusätzliche Daten betrachtet werden (HINZ & STILLA 2006; TÜRMER et al., 2013).

[1] Karlsruher Institut für Technologie, Institut für Photogrammetrie und Fernerkundung, Englerstr. 6, D-76131 Karlsruhe, E-Mail: [robin.niessner, stefan.hinz, boris.jutzi]@kit.edu
[2] Fraunhofer Institut für Optronik, Systemtechnik und Bildauswertung, Gutleuthausstr. 1, D-76275 Ettlingen, E-Mail: hendrik.schilling@iosb.fraunhofer.de

Das Kombinieren von Daten gewonnen aus verschiedenen Arten von Sensoren für die Analyse ist eine Strategie zur Steigerung der Klassifikationsleistung, insbesondere wenn die verwendeten Daten komplementär sind. Daher können radiometrische Daten in Form von RGB-Bildern und die mit einem LiDAR-Sensor gemessenen geometrischen Daten für die Fahrzeugklassifizierung vorteilhaft sein. Zur Einschränkung des Suchraums für die Fahrzeugerkennung sind geometrische Daten prinzipiell von Vorteil. Die LiDAR-Daten können zum Extrahieren von Fahrzeugen genutzt werden. Weiterhin können Merkmale für die Objektklassifizierung anhand der LiDAR-Daten gewonnen werden (JUTZI & GROSS 2009; WEINMANN et al., 2015). Um Fahrzeughypothesen abzuleiten, können in diesem Zusammenhang verschiedene Merkmale berechnet werden, indem radiometrische (optische) und geometrische (Elevations-) Daten zur Klassifizierung verwendet werden. Die Fusion der Daten erfolgt meist durch einen (adaptierten) State-of-the-Art-Klassifikator (SCHILLING et al., 2018).
Aktuelle Arbeiten verwenden Convolutional Neural Networks (CNNs). Diese mehrschichtigen neuronalen Netze sind so ausgelegt, dass sie optimale Merkmale aus Trainingsdaten für ein gegebenes Klassifikationsproblem lernen und vielversprechende Ergebnisse für Erkennungs- und Klassifikationsaufgaben zeigen.
In diesem Beitrag wird das Potential von Convolutional Neural Networks speziell für die Fahrzeugklassifikation untersucht. Daher werden drei Ansätze angegangen, um Klassifikatoren basierend auf Convolutional Neural Networks zu trainieren. Der Hauptbeitrag ist:

- CNNs basierend auf RGB- und LiDAR-Daten führen zu Klassifikationsergebnissen mit hoher Genauigkeit
- Merkmale, die von RGB-Daten abgeleitet werden und durch Transferlernen in LiDAR-Daten übertragen werden, führen zu besseren Klassifikationsergebnissen als die alleinige Verwendung von RGB-Daten
- Neuronale Netzwerke mit weniger Ebenen als weit verbreitete neuronale Netzwerke mit vielen Ebenen können bei einfachen Klassifikationsaufgaben mit wenigen Trainingsdaten zu besseren Klassifikationsergebnissen führen

2 Problemstellung - Klassifikation von Fahrzeugen

In diesem Beitrag werden drei auf CNNs basierende Ansätze vorgestellt, um Fahrzeuge in RGB und LiDAR zu klassifizieren. In Abschnitt 2.2.1 und Abschnitt 2.2.2 wird ein vortrainiertes CNN verwendet. Diese Ideen basieren auf dem Konzept des Transferlernens. In Abschnitt 2.3 wird ein CNN von Grund auf neu entworfen und trainiert. Durch diese Ansätze können die normalerweise benötigten enormen Mengen an Trainingsdaten reduziert werden, um CNNs zu trainieren.

2.1 RGB und LiDAR Datensatz

Zum Training und Evaluation der Experimente wird ein Datensatz verwendet, welcher vom Image Analysis and Data Fusion Technical Committee (IADF TC) of the IEEE Geoscience and Remote Sensing Society (GRSS) bereitgestellt wird. Die Daten bestehen aus einem RGB- und LiDAR-Datensatz. Der Datensatz wurde mit einem flugzeuggetragenem System über dem belgischen Hafen Zeebrügge aufgenommen. Die RGB-Daten sind Orthophotos mit einer

Auflösung von 5 cm. Die LiDAR-Daten werden als digitales Oberflächenmodell (DSM) mit einem Punktabstand von 10 cm bereitgestellt. Die Auflösung der RGB-Daten wird auf 10 cm reduziert, um mit der GSD der LiDAR-Daten zu übereinzustimmen. Da nur ein Datensatz verfügbar ist, werden die Trainings- und Validierungsdaten lokal getrennt, um die Korrelation zwischen den beiden Sätzen zu minimieren. Zu Trainingszwecken werden die Daten erweitert, indem sie um 90°, 180° und 270° gedreht werden, um mehr Trainingdaten zu erstellen und den Klassifikator robuster gegenüber Rotationsvarianzen zu machen. Weiterhin wird, um große Bias in den Aktivierungsfunktionen der CNNs zu vermeiden, von allen Kanälen der Bilder der Mittelwert subtrahiert.

2.2 Transferlernen

Die folgenden zwei Ansätze basieren auf dem Konzept des Transferlernens. Die Idee hinter diesen Ansätzen ist, dass CNNs, die auf einen Datensatz trainiert sind, übertragen werden können, um einen anderen Datensatz zu klassifizieren.
Verwendet wird das VGGNet (SIMONYAN & ZISSERMAN, 2014), ein CNN basierend auf der AlexNet-Architektur von (KRIZHEVSKY et al., 2012) welches auf dem ImageNet-Datensatz vortrainiert wurde (DENG et al., 2009). Die Merkmale, die das CNN in den ersten Ebenen erlernt, werden als allgemeine Merkmale betrachtet. Die folgenden Ebenen in der Architektur des CNN enthalten abstraktere Merkmale. Abbildung 1 zeigt die vortrainierten Merkmale der ersten Ebene des VGGNet.
Die Merkmale erscheinen wie eine Art von Gabor- und Blob-Merkmalen. Diese allgemeinen Merkmale sind für viele 2D-Bilderkennungsaufgaben geeignet.

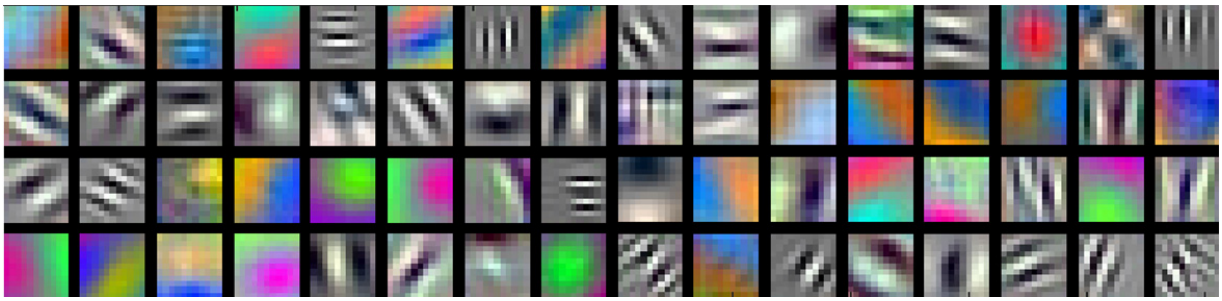


Abb. 1: Merkmale des vortrainierten VGGNet der ersten Ebene berechnet in *MATLAB*. Das Aussehen der Filtermasken ist ähnlich wie bei Gabor- und Blob-Filtern

2.2.1 CNN Merkmalsvektor

In diesem Ansatz wird das CNN lediglich als Merkmalsextraktor verwendet. Um dies zu erreichen, wird die letzte vollständig verbundene (fully connected) Ebene des CNNs entfernt, welche normalerweise die Klassifikationsebene ist. Wie in Abbildung 2a dargestellt, entspricht der resultierende Merkmalsvektor der Ausgabe der Faltung der vorletzten Ebene. Die Größe des Merkmalsvektors hängt von der CNN-Architektur ab, in diesem Falle ist die Größe *4096 x 1*.
Die extrahierten Features werden verwendet, um einen Standardklassifikator, eine lineare Support Vector Machine (SVM), zu trainieren. Um die RGB-, LiDAR-Höhen- und Intensitätsdaten zu fusionieren, werden die Merkmalsvektoren für das SVM-Training verkettet. Die Dimensionen der resultierenden Vektoren sind *4096 x 1, 8192 x 1* und *12288 x 1*.

2.2.2 Fine-Tuning

In diesem Ansatz wird das vortrainierte CNN als Merkmalsextraktor und Klassifikator verwendet. Das Trainieren eines vortrainierten CNN auf einem anderen Datensatz wird Fine-Tinung genannt. Auch hier müssen kleine Anpassungen am CNN vorgenommen werden. Theoretisch gibt es keine Einschränkungen, wie viele Änderungen der Architektur des CNN durchgeführt werden. Da nur eine kleine Menge von Trainingsdaten zur Verfügung steht, soll die Anzahl der Parameter, die neu trainiert werden müssen, so klein wie möglich gehalten werden. Das VGGNet wurde auf dem ImageNet-Datensatz trainiert. Somit besteht die Klassifikations-Ausgabeebene aus 1000 Elementen für die 1000 Klassen des ImageNet. Da nur die Hintergrundklasse und die Fahrzeugklasse getrennt werden sollen, muss die Klassifizierungsebene in eine Zwei-Elemente-Ebene geändert werden. Die grün hervorgehobenen Elemente in Abbildung 2b ersetzen die vorherige Klassifizierungsebene. Der Rest des CNN bleibt gleich. Die Gewichte zwischen der vorletzten *fully-connected* Ebene und der Klassifizierungsebene müssen bei diesem Ansatz neu trainiert werden. Hierfür werden sie zufällig initialisiert und mit Hilfe des stochastischen Gradientenverfahrens optimiert.

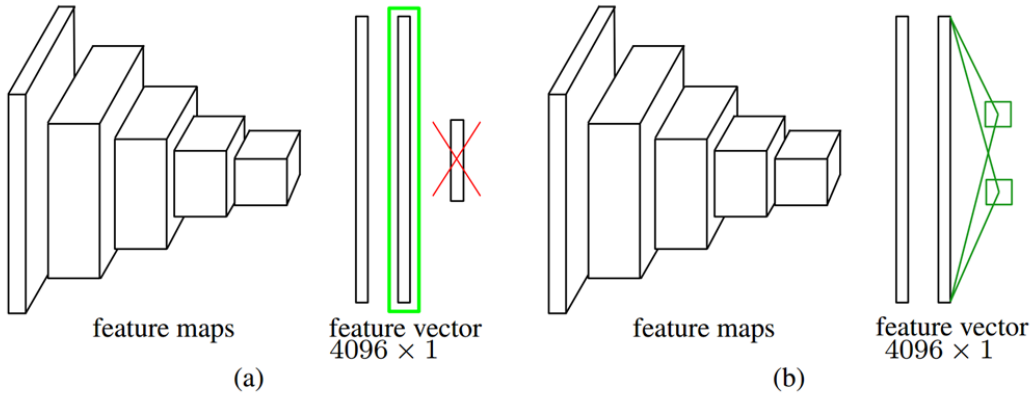


Abb. 2: (a) Änderungen im VGGNet, um die Ausgabe der letzten Ebene als Merkmalsvektor zu verwenden. Die Klassifikationsebene wird entfernt und die vorletzte Ebene mit der Größe *4096 x 1* wird als Merkmalsvektor verwendet. (b) Änderungen zur Feinabstimmung des VGGNet auf die Klassifizierungsaufgabe. Die Klassifizierungsebene und die entsprechenden Gewichte werden durch eine Zwei-Elemente-Klassifizierungsebene mit Gewichtsmatrizen der entsprechenden Größe ersetzt.

Ein Nachteil dieses Ansatzes ist, dass die Dimensionalität der Eingabedaten nicht geändert werden kann, da die vortrainierten Gewichtungen in der ersten Ebene des CNN nicht geändert werden sollen. Die Anzahl der Kanäle des Eingabebildes ist auf die Größe der Daten beschränkt, mit denen das CNN ursprünglich trainiert wurde. Dies impliziert, dass die RGB- und LiDAR-Daten nicht auf der Eingangsebene des CNN fusioniert werden können.

2.3 Selbsterstellte CNNs

Dieser Ansatz basiert auf dem Entwerfen und vollständigen Trainieren adaptierter CNN Architekturen. Aktuelle CNNs bestehen aus Millionen von Parametern und erfordern eine große Menge an Trainingsdaten. Da ein CNN für ein binäres Klassifizierungsproblem trainiert wird, erscheint ein CNN mit wenigen Parametern für diese Problemstellung ausreichend.

In Tabelle 1 sind die Architekturen der zwei leistungsstärksten CNNs dieser Studie aufgeführt. *S* ist die Schrittweite *stride* und *P* das *padding* der Filter der jeweiligen Faltungsebenen. Als Pooling-Methode wird das Max-Pooling verwendet und während des Trainings wird ein Dropout auf der ersten vollständig verbundenen (fully connected) Ebene angewandt, um ein Überanpassung (Overfitting) zu verhindern. Die entworfenen flachen CNNs bestehen nur aus einem Bruchteil von Parametern im Vergleich zu tiefen CNNs wie AlexNet. Die Größe des Filters in der ersten Schicht des Medium-CNN wurde mit der gleichen Größe wie die des Filters aus der AlexNet-Architektur gewählt. Die Größe der Filter des Large-CNN wurde so gewählt, dass sie auf der Skala der Objekte in dem Bild liegt, das klassifiziert werden soll.

Tab. 1: Die Architektur des von Grund auf neu konzipiert und trainierten CNNs.

	Medium-CNN
Conv1	*16x8x8, S=3,P=0 x3 pool*
Conv2	*32x5x5, S=1, P=0, x2 pool*
Conv3	*64x4x4, S=1, P=0*
Full4	*128, dropout*
Full5	2, softmax
Parameter	180.000

Im Gegensatz zum Ansatz von Abschnitt 2.2.2 können die RGB-, LiDAR-Höhen- und Intensitätsdaten in der Eingabeebene fusioniert werden. Hierfür werden die Kanäle der Bilder verkettet.

3 Ergebnisse und Evaluierung

Das Training der CNNs wurde an einem ausgeglichenen Datensatz mit je 399 Trainingsbeispielen für die Hintergrund- und Fahrzeugklasse vor der Datenerweiterung durchgeführt. Da es in einem realistischen Szenario mehr Hintergrund als Fahrzeuge gibt, werden 426 Fahrzeug- und 6017 Hintergrundmuster für die Validierung des Trainings verwendet, die mit einer Watershed-basierten Segmentierungsmethode erzeugt wurden (SCHILLING & BULATOV, 2016). Da die Datenverteilung für die Evaluation sehr unausgeglichen ist, ist die Gesamtgenauigkeit nicht geeignet, die Qualität der Klassifikation darzustellen. Der F-Score, das harmonische Mittel der Präzision und des Recalls, eignet sich besser für diese Aufgabe. Die Trainingsergebnisse werden mit einem Klassifizierungsansatz basierend auf den Histogramm von orientierten Gradienten (HOG)-Merkmalen mit einer *Random Forest* Klassifikation verglichen.

3.1 Klassifikationsergebnis

Das Training wird unter Verwendung des *MatConvNet* Frameworks in *MATLAB* (VEDALDI & LENC, 2015) durchgeführt. Für die Klassifizierungsebene wird ein Softmax-Klassifikator mit einer Cross-Entropy-Loss-Funktion verwendet, um die Gewichte zu aktualisieren. Als Optimierungsfunktion wird das Stochastische Gradientenverfahren (SGD) gewählt. Das CNN-Training wird mit Lernraten (LR) von *0,01*, *0,001* und *0,0001* durchgeführt. Die folgenden Beispiele des CNN-Trainings zeigen nur die besten Kombinationen von LR und Datentypen.

3.1.1 CNN Merkmalsvektor

Das Training der SVM mit den CNN-Merkmalen erfolgt durch eine 10-fache Kreuzvalidierung. Die SVM mit dem besten Ergebnis wird verwendet, um die Validierungsdaten zu klassifizieren. Wie in Abbildung 3 gezeigt, erreicht die Klassifikation mit den RGB-Daten den niedrigsten F-Score (*0,794*), obwohl die CNN-Merkmale ursprünglich in einem RGB-Datensatz trainiert wurden. Das beste Klassifikationsergebnis wird durch die Fusion des Elevations- und Intensitätsdatensatzes erreicht (F-Score = *0,929*). Das Hinzufügen der RGB-Daten zur Fusion des LiDAR-Datensatzes übertrifft noch immer die Ergebnisse mit RGB-Daten und liefert gute Ergebnisse, verbessert jedoch nicht die Ergebnisse der Fusion der LiDAR-Daten.

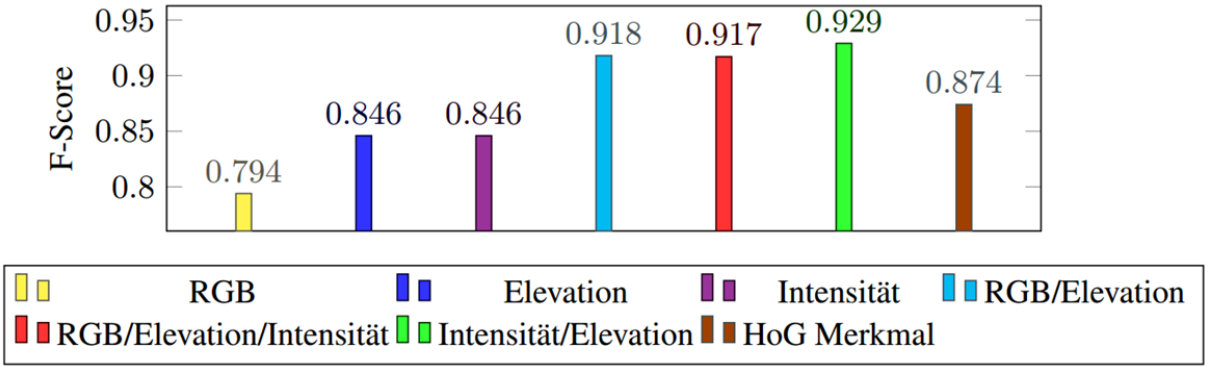


Abb. 3: *F-Scores* für das CNN Merkmalstraining mit einer SVM.

3.1.2 Fine-Tuning

Für das Fine-Tuning wird das gleiche vortrainierte CNN (VGGNet) verwendet, wie es bereits beim CNN-Merkmalsvektor Ansatz verwendet wurde. Für diesen Ansatz stehen nur Klassifikationsergebnisse für jeden Datentyp, nicht die fusionierten Daten zur Verfügung. Wie in Abbildung 4 gezeigt, übertrifft die Klassifikation auf der LiDAR-Elevation (F-Score: *0,903*) und Intensität (F-Score: *0,902*) wieder die Ergebnisse der RGB-Daten (F-Score: *0,867*). Im Vergleich zu den Ergebnissen aus Abbildung 3 zeigt die Klassifikation mit einzelnen Sensordaten verbesserte Ergebnisse für alle Datensätze.

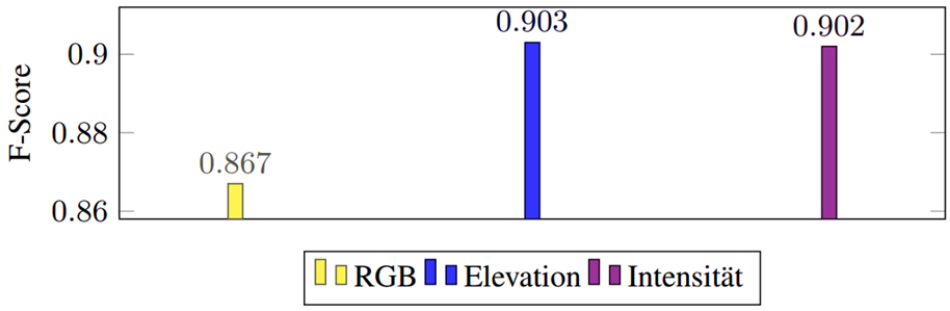


Abb. 4: *F-Scores* für das Fine-Tuning des VGGNet.

3.1.3 Selbsterstellte CNNs

Für die von Grund auf neu entwickelt und trainierten CNNs wird das Medium-CNN betrachtet. Beim Testen mehrerer Architekturen von CNNs zeigte dieses die beste Leistung sowohl bei der

Dauer des Trainings als auch bei den Klassifikationsergebnissen. Wie in Abbildung 5 und Abbildung 6 gezeigt, kann das Training mit der Medium-CNN Architektur im RGB-Datensatz (F-Score: *0,803*, gelbe Linie mit Dreiecken) die Ergebnisse, die durch das Training mit den LiDAR-Daten erreicht werden, nicht erreichen. Das beste Klassifikationsergebnis wird durch die CNN-Datenfusion von Elevation und Intensität erreicht (F-Score: *0,955*, grüne Linie mit Punkten).
Verglichen wird auch das Ergebnis des Medium-CNN-Trainings mit einer HoG-Feature-Klassifikation (F-Score: *0,874*, braune Linie mit Quadraten), die auf fusionierten RGB- und LiDAR-Datensätzen trainiert wurde.
Während dieser Klassifizierungsansatz bessere Ergebnisse liefert als die Klassifikation von RGB-Daten mit dem Medium-CNN, liegt er innerhalb des gleichen Bereichs wie das auf dem Intensitätsdatensatz trainierte Medium-CNN und die Fusion von RGB- und Elevationsdaten. Für die Elevationsdaten, die fusionierten Daten des LiDAR und die fusionierten RGB- und LiDAR-Datensätze übertrifft das Medium-CNN die HoG-Feature-Klassifikation.

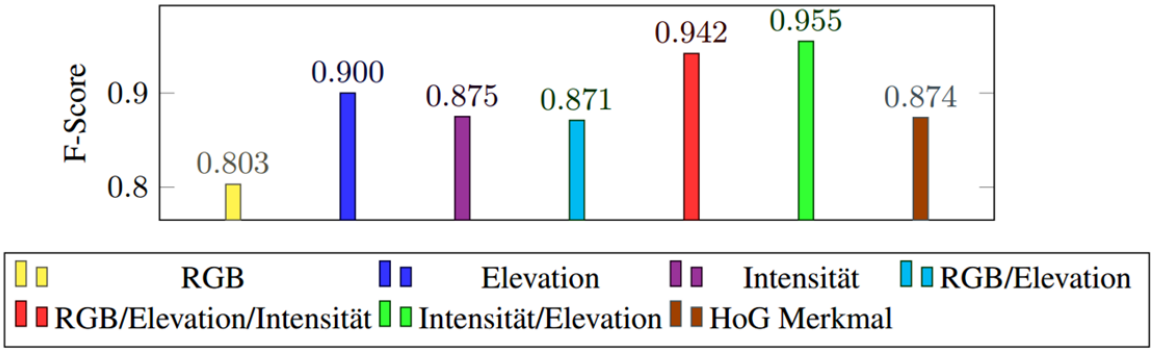


Abb. 5: *F-Scores* für die Trainingsergebnisse, die von verschiedenen Daten des Medium-CNN abgeleitet wurden, mit dem HoG-Merkmal-Klassifizierungsergebnis zum Vergleich.

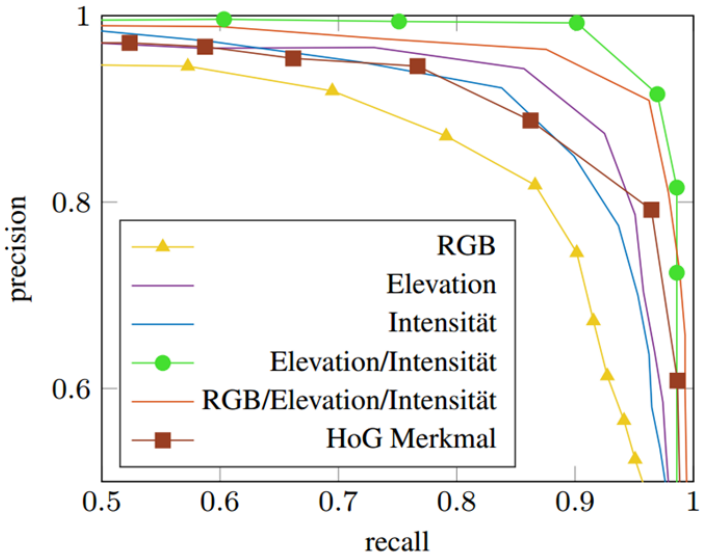


Abb. 6: PR-Kurven der Klassifikationsergebnisse, die aus verschiedenen Daten für das Medium-CNN abgeleitet wurden, mit dem HoG-Feature-Klassifizierungsergebnis als Vergleich.

3.2 Sliding-Window Klassifikation

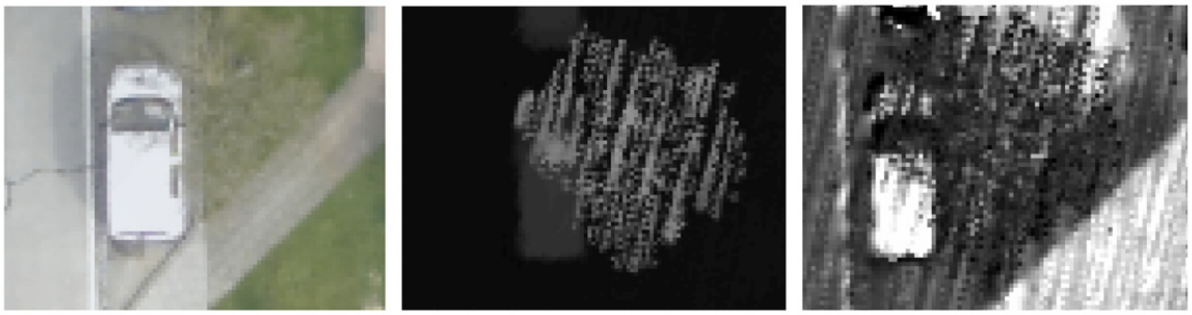
Schließlich soll ein Detektionsergebnis des CNN-basierten Klassifikationsansatzes auf einem gegebenen Datensatz gezeigt werden, um die Leistungsfähigkeit des Klassifikators zu untersuchen. Das Medium-CNN wird für eine Sliding-Window Klassifikation auf der Fusion von LiDAR-Daten gewählt, da es die beste Leistung in Abschnitt 3.1 zeigt. Die Detektion erfolgt über ein gleitendes Fenster mit einer Größe von *80 x 80* Pixeln. Die Abtastrate ist mit 10 Pixel vorgegeben, so dass sichergestellt ist, dass jedes Fahrzeug enthalten ist und die Anzahl der zu berechnenden Bilder minimiert wird. Für einen Testbereich der Größe *5000 x 5000* Pixel werden *243.049* Bilder berechnet. Abbildung 7 zeigt Beispiele aus dem klassifizierten Bereich. Zur Visualisierung wird das RGB-Bild als Hintergrund verwendet, obwohl die Daten auf der Fusion von LiDAR-Daten klassifiziert wurden. Die Klassifikationsergebnisse werden unter Verwendung einer Heatmap dargestellt, bei der die Klassifizierungswerte der klassifizierten Pixel dargestellt werden. Auf einem großen offenen Feld, bei denen die Fahrzeuge verteilt sind, hat der Klassifikator keine Probleme, alle Fahrzeuge mit geringer Fehlerrate (falsch positiv) zu erkennen.

Eine anspruchsvollere Aufgabe ist die Detektion von Fahrzeugen, die nahe beieinander stehen oder sich im Hinterhof befinden. In Abbildung 7c wird ein Beispiel gezeigt, bei dem Fahrzeuge nebeneinander geparkt sind, mit einigen Ausnahmen kann der Klassifikator die Fahrzeuge noch trennen. Abbildung 7d zeigt einen interessanten Fall, bei dem die Fahrzeuge völlig unterschiedliche Größen haben. Der Klassifikator kann kleine Fahrzeuge bis hin zu großen Wohnwagen erkennen, obwohl die Anzahl der Trainingsbeispiele für Wohnwagen sehr gering ist.

Abb. 7: Abbildungen (a) und (b) zeigen Beispiele aus dem gesamten klassifizierten Bereich. Abbildung (c) zeigt ausschnittsweise ein Beispiel für die Klassifikation auf Parkplätzen und Abbildung (d) zeigt, dass sich der Klassifikator auch für Hinterhöfe mit Fahr Fahrzeugen unterschiedlicher Größen eignet. Es ist zu beachten, dass die Klassifizierung der Fahrzeuge anhand der Fusion von LiDAR-Daten durchgeführt wird, jedoch werden die Ergebnisse unter Verwendung der RGB-Daten als Hintergrund angezeigt, um die Sichtbarkeit zu erhöhen. Das Heatmap-Overlay verwendet die vom Medium-CNN berechneten Klassifizierungswerte.

Eine herausfordernde Aufgabe ist gegeben, wenn die Fahrzeuge von Vegetation oder ähnlichen Objekten bedeckt sind. Obwohl das Fahrzeug in der RGB-Darstellung in Abbildung 8a vollständig sichtbar ist, hat unser Klassifikator es nicht als Fahrzeug klassifiziert. In den Elevationsdaten in Abbildung 8b ist das Fahrzeug kaum sichtbar, obwohl der letzten Puls des LiDAR-Signals verwendet wird, um die 2D-Bilder darzustellen. Auch in der

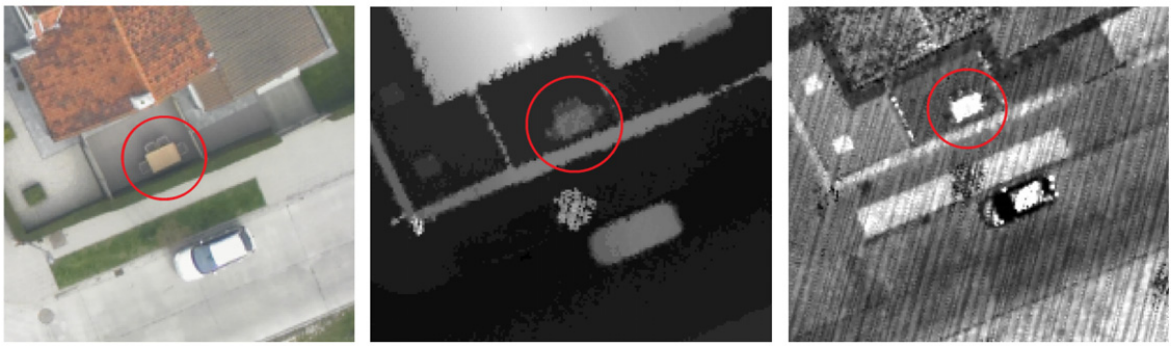
Intensitätsdarstellung in Abbildung 8c sind große Bereiche des Fahrzeugs nicht sichtbar. Da die Merkmale, die das CNN in den ersten Ebenen erlernt, meistens Kantendetektoren sind, kann es das Fahrzeug nicht korrekt klassifizieren.



(a) (b) (c)

Abb. 8: Fehlklassifizierung eines von einem Baum bedeckten Fahrzeugs. (a) zeigt das Fahrzeug in den RGB-Daten. (b) das Fahrzeuge ist kaum sichtbar. (c) Im Intensitätsbild ist die Front des Fahrzeugs größtenteils von einem Baum bedeckt.

Andere Hindernisse sind Objekte, die Fahrzeugen in ihrer Geometrie sehr ähnlich sind. Die Abbildung 9a zeigt Gartenmöbel, die in der RGB-Darstellung bedingt anders aussehen als ein Fahrzeug. Betrachtet man die Darstellung von Elevation (9b) und Intensität (9c), ähneln die Möbel einem kleinen Fahrzeug oder dem Dach eines Fahrzeugs. Der Detektor muss für verschiedene Größen empfindlich sein, da das Ziel darin besteht, alles von einem kleinen Fahrzeug bis zu einem Wohnwagen zu erfassen.



(a) (b) (c)

Abb. 9: Fehlklassifizierung eines Objekts, das bedingt einem Fahrzeug ähnelt. RGB-Bild (a) zeigt Gartenmöbel, die fälschlicherweise als Fahrzeug bestimmt wurden. Im Elevationsbild (b) und im Intensitätsbild (c) haben die Möbel die gleiche Art von Kanten wie ein Fahrzeugdach, was zu einer Fehlklassifizierung führt.

4 Diskussion und Fazit

In Tabelle 2 wird einen Überblick über alle erzielten Klassifizierungsergebnisse gegeben.
Die besten Ergebnisse wurden durch das Medium-CNN erzielt, das auf die Fusion der LiDAR-Daten ausgelegt und trainiert wurde, dicht gefolgt von der Fusion von RGB- und LiDAR-

Datensätzen. Für 5 von 6 Untersuchungen konnte zudem die Klassifikationsleistung durch die Fusion der Sensordaten im Vergleich zur Verwendung der einzelnen Daten erhöht werden. Die einzige Ausnahme ist die Fusion von RGB- und Elevationsdaten für die selbsterstellten CNNs. Der Fine-Tuning-Ansatz konnte die besten Klassifikationsergebnisse für die einzelnen RGB- und LiDAR-Daten erzielen. Betrachtet man die Gesamtleistung, liefert die LiDAR-basierte Klassifizierung bessere Ergebnisse als die RGB-basierte Klassifizierung.

Tab. 2: Übersicht der erreichten F-Scores für alle Trainings für jeden Datentyp und jede Kombination.

Daten	CNN Merkmalstraining	Fine-Tuning	Selbsterst. CNNs
RGB	0,794	0,867	0,803
Elevation	0,846	0,903	0,900
Intensität	0,846	0,902	0,875
RGB/Elevation	0,918	-	0,871
Elevation/Intensität	0,929	-	0,955
RGB/Elevation/Intensität	0,917	-	0,942

Dies ist ziemlich überraschend, da die CNNs, die für die Transferlernansätze (CNN Feature, Fine-Tuning) verwendet wurden, mit einem RGB-Datensatz vortrainiert wurden. Dies führt zu der Schlussfolgerung, dass in RGB-Daten gelernte Merkmale in LiDAR-Daten übertragbar sind. Die LiDAR-Daten scheinen für diese Klassifizierung besser geeignet zu sein. Da nur wenige Hundert Trainingsbeispiele zur Verfügung stehen, kann die Einheitlichkeit der Geometrie der Fahrzeuge zu stabileren geometrischen Merkmalen als die Radiometrie aus den RGB-Daten führen, da Fahrzeuge unterschiedlicher Farben in den LiDAR-Daten meist das gleiche Aussehen haben. Ein Grund für die abnehmende Leistung durch Hinzufügen der RGB-Daten zur Fusion der LiDAR-Daten könnte eine ungenaue Koregistrierung der 3D-LiDAR-Daten auf den 2D-RGB-Daten sein. Dies führt zu unscharfen Kanten. Da die Filter in der ersten Schicht der CNNs hauptsächlich Kantendetektoren sind, kann sich das Ergebnis hierdurch verschlechtern. Weiterhin zeigt die Klassifikationsleistung des Medium-CNN mit nur 180.000 Parametern, dass tiefe CNNs für binäre Klassifizierungsaufgaben in dieser Fernerkundungsanwendung nicht erforderlich sind.
Bei zukünftigen Untersuchungen wollen wir weitere CNN-Architekturen wie siamesische CNNs testen, um das Problem einer unpräzisen Koregistrierung anzugehen und mehr Aufmerksamkeit auf die Eigenschaften jedes Eingangskanals zu richten. Darüber hinaus beabsichtigen wir, CNN-Klassifizierungen für Hyperspektraldaten und Kombinationen von RGB-, Hyperspektral- und LiDAR-Datensätzen zu testen.

5 Literaturverzeichnis

DALAL, N. & TRIGGS, B., 2005: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, **1**, IEEE, 886–893.

DENG, J., DONG, W., SOCHER, R., LI, L.-J., LI, K & FEI-FEI, L., 2009: Imagenet: A large-scale hierarchical image database. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, IEEE, 248–255.

HINZ, S. & STILLA, U., 2006: Car detection in aerial thermal images by local and global evidence accumulation. Pattern Recognition Letters, **27**(4), 308–315.

JUTZI, B. & GROSS, H., 2009: Nearest neighbour classification on laser point clouds to gain object structures from buildings. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(Part 1), 4–7.

KRIZSHEVSKY, A., SUTSKEVER, I. & HINTEN, G. E., 2012: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, 1097–1105.

LOWE, D. G., 1999: Object recognition from local scale-invariant features. In: Proceedings of the Seventh IEEE International Conference on Computer Vision, **2**, 1150–1157.

OJALA, T., PIETIKAINEN, M. & HARWOOD, D., 1994: Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In: Proceedings of the 12th IAPR International Conference on Pattern Recognition, **1**, 582–585.

SCHILLING, H. & BULATOV, D., 2016: Segmentation methods for detection of stationary vehicles in combined elevation and optical data. In: International Conference on Pattern Recognition (ICPR), International Society for Optics and Photonics, 592–597.

SCHILLING, H., BULATOV, D. & MIDDELMANN, W., 2018: Object-based detection of vehicles using combined optical and elevation data. ISPRS Journal of Journal of Photogrammetry and Remote Sensing, **6**(6), 85-105.

SIMONYAN, K. & ZISSERMAN, A. 2014: Very deep convolutional networks for large-scale image recognition. CoRR.

TÜRMER, S., KURZ, F., REINARTZ, P & STILLA, U., 2013: Airborne vehicle detection in dense urban areas using hog features and disparity maps. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **6**(6), 2327–2337.

VEDALDI, A. & LENC, L., 2015: Matconvnet: Convolutional neural networks for Matlab. In: Proceedings of the 23rd ACM international conference on Multimedia, ACM, 689–692.

WEINMANN, M., SCHMIDT, A., MALLET, C., HINZ, S., ROTTENSTEINER, F. & JUTZI, B., 2015: Contextual classification of point cloud data by exploiting individual 3d neighbourhoods. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **2**(3), 271.

Simulation raumzeitlicher Effekte des Angebotes und der Nachfrage von Hackschnitzeln in Kärnten: Ein Agentenbasierter Ansatz

FLORIAN BREITWIESER[1], JOHANNES SCHOLZ[1] & PETER MANDL[2]

Zusammenfassung: Energetische Nachhaltigkeit gewinnt in Gesellschaft und Politik immer mehr an Bedeutung, deshalb untersucht die hier vorgestellte Studie das Potential von Agenten-basierter Modellierung im Zusammenhang mit der Versorgungskette des nachhaltigen Energieträgers Hackschnitzel. Im Untersuchungsgebiet, dem österreichischen Bundesland Kärnten, wird ein Agenten-basiertes Modell umgesetzt und beschreibt eine vereinfachte Marktsituation vom Wald bis zu den Heizwerken. Das Modell besteht aus drei Agententypen: den Waldeigentümern auf der Angebotsseite, den Heizkraftwerken auf der Nachfrageseite und den intermediär agierenden "Händlern". Die Eigentümer besitzen unterschiedliche wirtschaftliche Interessen, auf welche die Händler reagieren und transportkosten-optimiert handeln. Das Modell geht von einer autarken Holzversorgung im Untersuchungsgebiet aus und gibt die raumzeitliche Entwicklung der Transportdistanz, Erntemuster und Ernteverteilung in jährlichen Schritten wider. Die Ergebnisse bilden verschiedene Simulationsszenarien, die eine differenzierte Marktentwicklung abbilden sollen.

1 Einleitung

Erneuerbare Energiequellen werden in Österreich durch Wasserkraft, Windenergie und Biomasse, z.B. in Form von Hackschnitzeln, zur Nah/Fernwärmeeinspeisung bereitgestellt. Durch gesellschaftlichen sowie politisch forcierten Stellenwert der nachhaltigen Energiebereitstellung gewinnt die wirtschaftlich optimierte Nutzung der Ressource Holz zur Energiegewinnung in Form von Hackschnitzeln immer mehr an Bedeutung (MADLENER et. al. 2007).

Diese Arbeit hakt genau in diesem Themengebiet ein und beschäftigt sich mit der Modellierung, Simulation und Analyse der raumzeitlichen Zusammenhänge in der Versorgungskette von Hackschnitzeln. Aufbauend auf einer Studie von SCHOLZ et. al. (2014) wird der Hackschnitzel-Markt mit einem Agenten-basieren Modell abgebildet. Innerhalb dieses Modells konkurrieren die Biomassenheizwerksbetreiber um den in begrenztem Ausmaß zur Verfügung stehenden Rohstoff Hackschnitzel und müssen dabei die verschiedenen Interessen der Waldeigentümer sowie die Transportkosten berücksichtigen.

Trotz der Aktualität des Forschungsfeldes finden sich in der Literatur nur wenige Arbeiten, die sich auf lokalem bzw. globalem Maßstab auch mit den dynamischen räumlichen Aspekten im Zusammenhang mit erneuerbaren Energiequellen beschäftigen (CINTAS et. al. 2015; SACHELLI et. al. 2013; ARBEITSPLATTFORM FÜR HOLZ UND WALD IN KÄRNTEN 2009).

[1] Technische Universität Graz, Institut für Geodäsie – Arbeitsgruppe Geoinformation, Steyrergrasse 30/I, A-8010 Graz, E-Mail: florian.breitwieser@student.tugraz.at, johannes.scholz@tugraz.at

[2] Alpen-Adria-Universität Institut für Geographie und Regionalforschung, Universitätsstrasse 65-67, A-9020 Klagenfurt, +43 463 2700 3216, peter.mandl@aau.at

Methodisch wird zur Simulation einer solchen Marktsituation bzw. Versorgungskette auf ein Agenten-basiertes Modell (ABM) zurückgegriffen. Ein ABM beschreibt und simuliert Systeme, denen autonom agierende Teilnehmer zugrunde liegen (CROOKS & HEPPENSTALL 2012). Zentrales Merkmal ist die Interaktion dieser Agenten untereinander bzw. das Zusammenspiel mit einer modellierten Umwelt (BATTY 2009; MANDL 2003). Des Weiteren zeigen die Arbeiten von BATTY et. al. (2012), JOHNSTON (2012), KOSTADINOV et. al. (2014) und O'SULLIVAN et. al. (2012), dass ABM bereits vielfältig zur Simulation von Realwelt-Phänomenen eingesetzt wird. Der Artikel ist folgendermaßen aufgebaut: Abschnitt 2 beschreibt das umgesetzte Modell. In Abschnitt 3 werden die Ergebnisse vorgestellt und Abschnitt 4 bietet eine abschließende Diskussion und einen Ausblick.

2 Problemstellung

Das österreichische Bundesland Kärnten bildet das im Modell berücksichtigte Testgebiet. Mit einem Waldflächenanteil von rund 57% zählt Kärnten, nach der Steiermark, zu den waldreichsten Bundesländern in Österreich. Im Testgebiet befinden sich 18 größere und kleinere, durch Biomasse befeuerte, Nah- und Fernwärmeheizwerke, welche um den Rohstoff Hackschnitzel konkurrieren. Die berücksichtigten Heizwerke weisen einen jährlichen Bedarf von ca. 240.000 fm (Festmetern) Holz in Form von Hackschnitzeln auf. Erklärtes Ziel dieser Betriebe ist es, den Bedarf durch das lokal zur Verfügung stehende Hackschnitzelvolumen zu decken. Laut Angaben der Landwirtschaftskammer Kärnten im Durchschnitt der Holzentnahme aus dem Kärntner Wald in den letzten Jahren bei etwa 2,5 Mio. Erntefestmeter. Wichtigste Holzsortimente sind Sägerundholz (60%), Industrieholz (25%) und Brennholz (ca. 15%). (ARBEITSPLATTFORM FÜR HOLZ UND WALD IN KÄRNTEN 2009).
Somit stellt sich diese Studie der Problemstellung, die auftretenden raumzeitlichen Effekte dieses Konkurrenzkampfes um den jährlich zur Verfügung stehenden Rohstoff Hackschnitzel zu untersuchen. Außerdem soll untersucht werden ob sich das Bundesland Kärnten bei der aktuellen Lage des Holzmarktes auch zukünftig autark mit dem Rohstoff Hackschnitzel versorgen kann.

3 Methoden

Dieser Abschnitt beschreibt den methodischen Ansatz dieser Arbeit. Räumliche und tabellarische Daten werden im Zusammenhang mit ABM zur Simulation des lokalen Hackschnitzelmarktes herangezogen. Zudem gibt dieser Teil der Arbeit einen kurzen Überblick über die Eckdaten des Testgebiets Kärnten.

3.1 Datengrundlage

Für das Modell wurden folgende Daten bereitgestellt: Informationen über Wald, beispielsweise Vorratsmenge, Alter und Waldart wurden auf Basis von Rasterdaten mit einer räumlichen Auflösung von 30 m mal 30 m durch die Landwirtschaftskammer Kärnten und das Bundesforschungszentrum für Wald (BFW) bereitgestellt. Diese Daten wurden zur Modellierung des jährlichen Zuwachses und der jährlich zur Verfügung stehenden Einschlagsmenge herangezogen. Zur Berechnung der Transportkosten wurde ein Digitales Geländemodell (DGM), mit einer Auflösung

von 10 m mal 10 m und der Straßendatensatz der Graphen Integrationsplattform (GIP) kombiniert. Das DGM wurde über das Open-Government-Portal des Landes Kärnten bezogen. Das Ziel der GIP ist es einen gemeinsamen österreichweiten Verkehrsgraphen und eine digitale Karte für das Verkehrsnetz zu schaffen (HEIMBUCHNER 2017). Die Daten für die Biomasseheizwerke (räumliche Position, jährlicher Bedarf usw.) wurden durch die Autoren erfasst.

3.2 Modellaufbau

Methodisch wird diese dynamische Problemstellung in Form eines ABM umgesetzt. Um die Realität entsprechend abzubilden, wird ein raumzeitliches Modell definiert, welches folgende drei Agententypen berücksichtigt. Die Angebotsseite bilden die Waldeigentümer und die Aufgabe haben den Wald zu bewirtschaften. Außerdem entscheidet er ob und wann Holz entnommen wird. Auf der Nachfrageseite stehen die Heizwerke. Sie weisen einen jährlichen zu deckenden Bedarf an Biomasse in Form von Hackschnitzel auf. Intermediär stehen die Händler, sie bilden das Bindeglied zwischen Waldeigentümern und Heizwerken.

Das logische Grundprinzip des Modells ist in Abb. 1 dargestellt. Durch ein einfaches Waldwachstumsmodell wird auf Basis der räumlichen Daten der jährliche Zuwachs berechnet. Dieses Modell basiert auf den Ertragstabellen (MARSCHALL 1975) und waldbaulichen Maßnahmen nach WEINFURTER (2013). Basierend auf dem Vorrat und dem Alter des Waldes sowie bezugnehmend auf den Waldentwicklungsplan und aus dem Forstgesetz 1975 abgeleiteten Regeln, entnimmt der Eigentümer Holz aus dem Wald und bietet die entnommene Menge an. Bei der Entnahme werden Faktoren wie Ernteverluste und Hackgutanteil berücksichtigt. Die Händler haben die Aufgabe den Bedarf der Heizwerke transport- und kosten-optimiert zu decken, wobei genau ein Händler einem Heizwerk zugeordnet ist.
Außerdem werden im Modell die Eigentumsstrukturen des Waldes und die daraus resultierenden Einzelinteressen der Eigentümer berücksichtigt. Die ca. 20.000 Waldeigentümer, welche den Wald bewirtschaften und entscheiden ob Holz entnommen wird, können folgendermaßen unterschieden werden (ECKMÜLLNER 1964; JÖBSTL 1986):

- Kleinwald (< 200 ha)
- Großwald (200-1.000 ha)
- forstwirtschaftliche Betriebe (> 1.000 ha) und die staatlichen Österreichischen Bundesforste AG

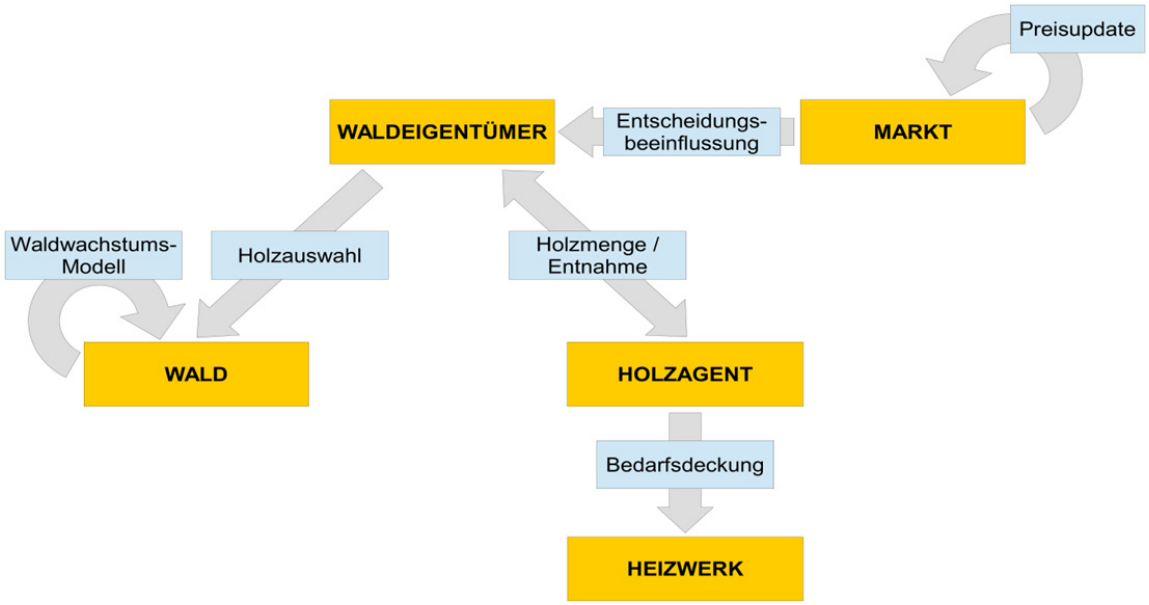


Abb. 1: Aufbau und Zusammenspiel der Modellkomponenten

Die privaten Kleinwaldbesitzer weisen ein volatiles Nutzungsverhalten auf, da eine gleichmäßige Nutzung oft nicht möglich ist. Grundsätzlich wird also Holz entnommen, wenn Material- oder finanzieller Bedarf besteht (ECKMÜLLNER 1964; ARBEITSPLATTFORM FÜR HOLZ UND WALD 2009). Im Gegensatz dazu stehen die kommerziell ausgerichteten forstwirtschaftlichen Betriebe (> 1.000 ha), welche aufgrund laufender Kosten langfristig konstante Erntemengen bereitstellen. Forstbetriebe sind hinsichtlich der Holzendnutzung weniger preiselastisch (JÖBSTL 1986; SCHWARZENBAUER et. al. 2009). Die Entscheidungen der Eigentümer beeinflussen dabei die jährlich zur Verfügung stehende Menge an Holz und in weiterer Folge auch das Hackschnitzelvolumen. Als weiterer Faktor kommt der Markt hinzu. Dieser beeinflusst das Verhalten der Agenten von außen durch die Entwicklung des Holzpreises.
Die Simulation des Modells läuft rundenbasiert ab und setzt auf Iterationszyklen von einem Jahr. Für die Simulation werden verschiedene Szenarien definiert, welche im Hinblick auf die raumzeitliche Entwicklung der Transportdistanz, Entnahmeverteilung und der Holzvorratsentwicklung analysiert werden.

3.3 Validierung

Die Validierung des Modells erfolgt über eine Sensitivitätsanalyse und einen Vergleich mit bestehenden historischen Entnahmezahlen in Bezug auf die Zeitspannen (1994-1996, 2000-2002 und 2007-2009) der Österreichischen Waldinventur (ÖWI). Durch die Verwendung einer aus historischen Holzpreisen extrapolierten Preisentwicklung können Entnahmemengen der Eigentumsklassen der Simulation mit jenen der ÖWI verglichen werden. Wie in Abb. 2 ersichtlich, ergibt sich im Mittel eine Abweichung von 1,8%. Über die drei Vergleichszeiträume kann im schlechtesten Fall eine Abweichung von ca. 4% ermittelt werden.

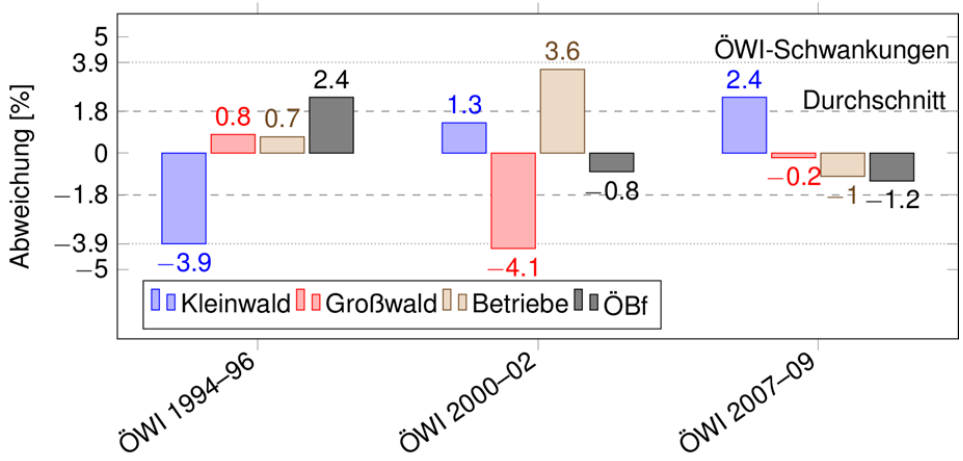


Abb. 2: Ergebnisse der Validierung. Prozentuale Abweichung der Entnahmemengen pro Eigentumsklasse im Modell bezogen auf die im Zuge der ÖWI erhobenen Entnahmemengen

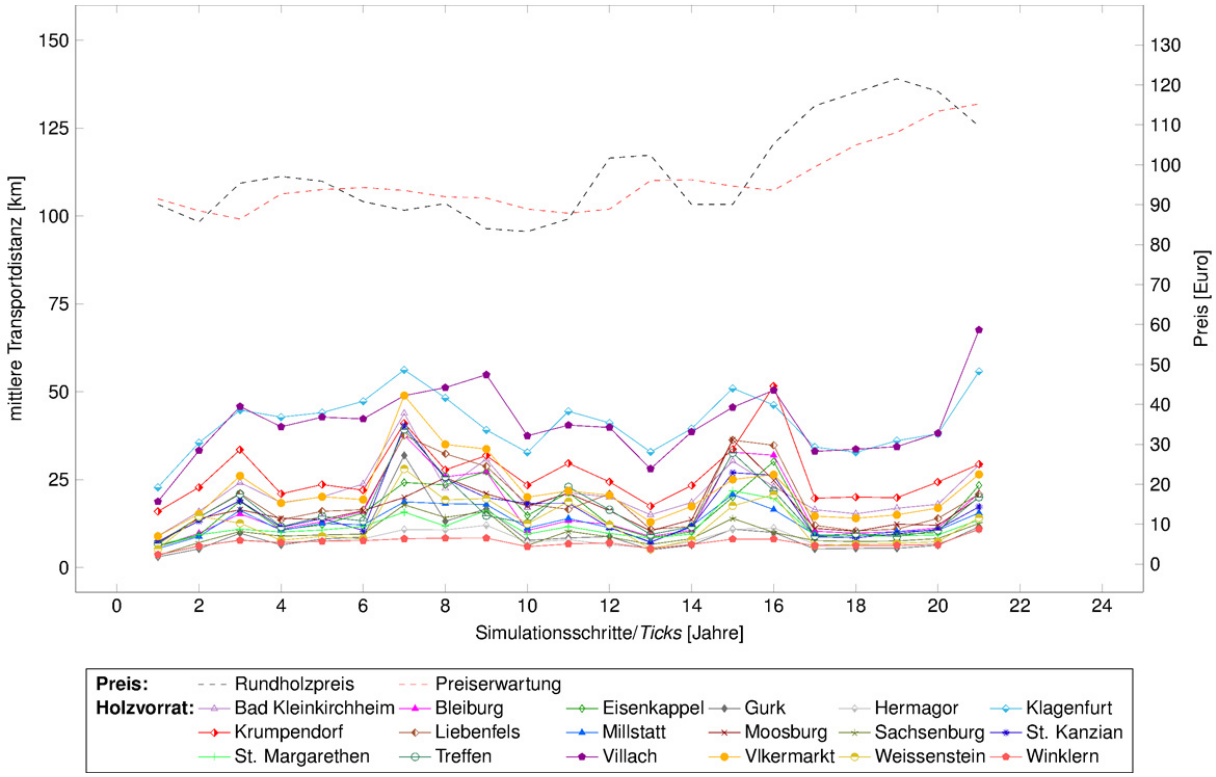


Abb. 3: Entwicklung der mittleren Transportdistanz zu den Heizwerken pro Holzbeschaffung über den Simulationsverlauf von 21 Jahren. Dieser ist im Bezug zur extrapolierten Holzpreisentwicklung und der mittleren Preiserwartung der Waldeigentümer dargestellt. Liegt die Erwartung unter dem Preis müssen weniger Transportkosten in Kauf genommen werden und umgekehrt.

4 Ergebnisse

Die Ergebnisse der Arbeit sind 6 Simulationsszenarien. Zwei Szenarien sind dabei ressourcenorientiert und beobachten die Hackschnitzel-Nutzung ohne Einfluss des Marktes. Die vier marktorientierten Szenarien beschreiben differenzierte Entwicklungen des Holzpreises.
In Abb. 3 ist die Entwicklung der mittleren Transportdistanz im Zusammenhang mit einer positiven Holzpreisentwicklung ersichtlich. Steigt der Preis, nimmt auch die angebotene Holzmenge seitens der Eigentümer zu, wobei in weiterer Folge die Transportdistanz für das betrachtete Jahr (Tick) sinkt. Fällt der Preis, verhält es sich umgekehrt und die mittlere Transportdistanz nimmt zu.
Abb. 4 zeigt die Entwicklung der Entnahmeverteilung. In den ersten Iterationsschritten (Tick 1-3) weisen die Heizwerke eine relativ konzentrische Entnahmestruktur rund um den eigenen Standort auf. Mit Fortdauer der Simulation sowie durch schlechte Preislage (Tick 10-12 und 19-21) zeigen die Entnahmemuster eine immer weiter fortschreitende Zerstreuung und Entfernung vom Standort des Heizwerkes auf. Die Händler müssen in diesem Fall versuchen den Bedarf der Heizwerke zu befriedigen, ohne auf andere Randbedingungen Rücksicht nehmen zu können.

5 Fazit und Ausblick

Die Nachfrage der Biomasseheizwerke kann durch den Wald im Bundesland Kärnten nicht gedeckt werden. Die Versorgungssicherheit ist somit im Laufe der Jahre nicht gewährleistet.
Durch Modellierung der Marktteilnehmer in Form von kognitiven Agenten können diese individuell auf Marktsituationen reagieren. Besonders durch den hohen Klein- und Kleinstwaldanteil ist die konstante Versorgungssicherheit nur schwer sicherzustellen, da diese Agenten besonders volatil auf Änderungen des Marktumfeldes (z.B. Holzpreis) reagieren. Zudem haben diese Marktteilnehmer die Tendenz, ihre Ressourcen zu „horten“, wenn der Marktpreis nicht den Vorstellungen entspricht.
Das hier dargestellte Modell kann zukünftig noch in mehreren Punkten erweitert werden. Zum einen wäre die Verfeinerung des Marktmodells erstrebenswert, sodass man den ausländischen Markt und die Papier- und Sägeholzindustrie miteinbezieht. Außerdem könnte man Waldgemeinschaften und Genossenschaften als separate Agenten integrieren. Auch die Integration von Schadereignissen wie Stürmen (Windbruch und -wurf) oder Wildschäden wäre denkbar.

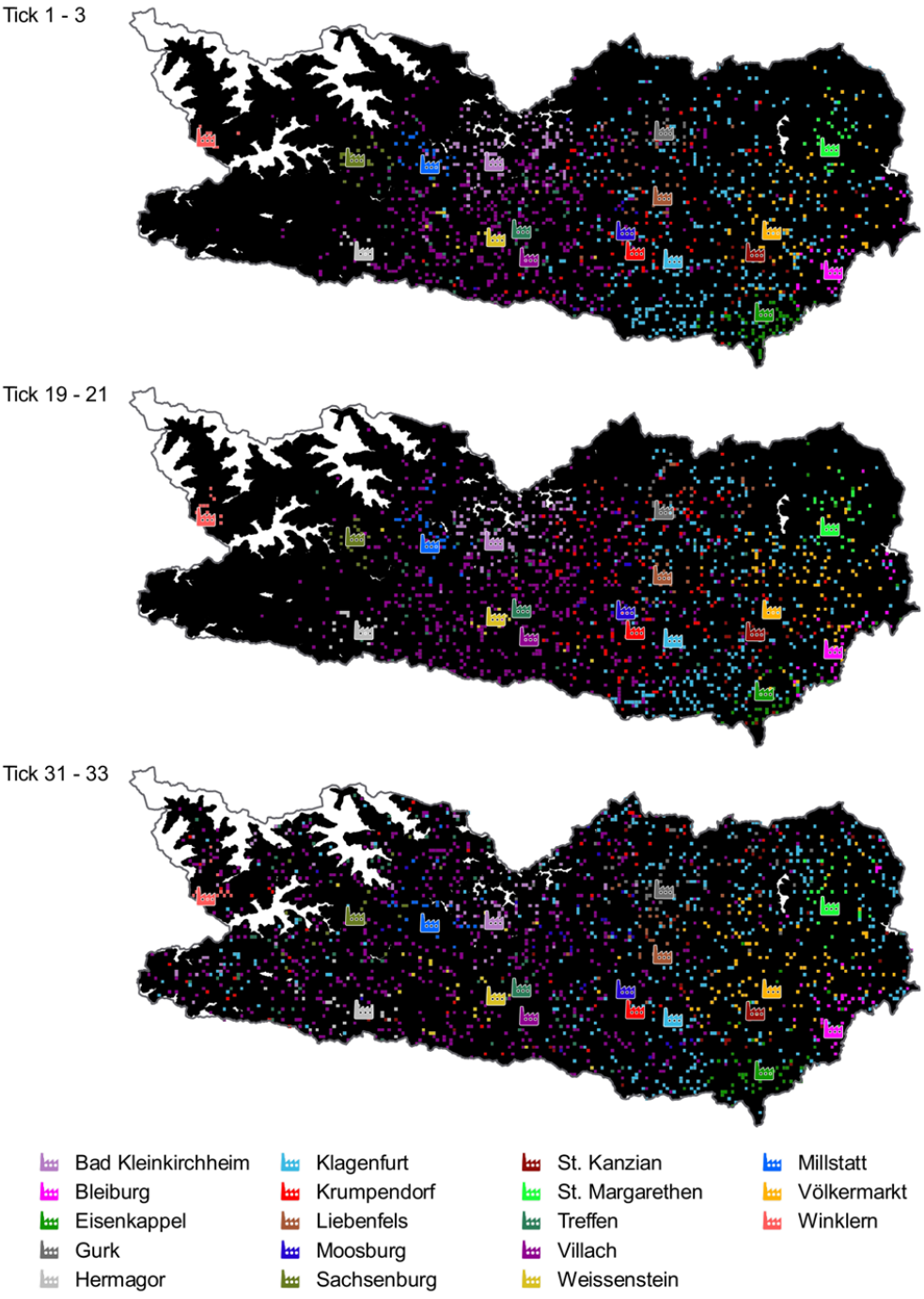


Abb. 4: Raumzeitliche Entwicklung der Holzentnahmebereiche der Heizwerke. Mit knapper werdendem Holzvorrat verschieben sich die Entnahmebereich von lokal um die Standorte immer weiter in periphere Bereiche.

6 Literaturverzeichnis

BRUMMITT, N. & LUGHADHA, E.N., 2003: Biodiversity: Where's Hot and Where's Not. Conservation Biology, **17** (5), 1442-1448.

ARBEITSPLATTFORM WALD UND HOLZ IN KÄRNTEN, 2009: Bilanz und Strategieplan über Aufkommen, Nutzen und Potentiale. Landwirtschaftskammer Kärnten, ausgearbeitet von Vertretern der Kärntner Forst- und Holzwirtschaft.

BATTY, M., 2009: Urban modeling. International encyclopedia of human geography, Elsevier, Oxford.

BATTY, M., CROOKS, A. T., SEE, L.M. & HEPPENSTALL, A.J., 2012: Perspectives on agent-based models and geographical systems. In: Heppenstall, A. J., Crooks, A. T., See, L. M. & Batty, M. (Hrsg.). Agent-based models of geographical systems, Springer, 1-15.

CINTAS, O., BERNDES, G., COWIE, A. L., EGNELL, G., HOLMSTRÖM, H. & ÅGREN, G. I., 2015: The climate effect of increased forest bioenergy use in Sweden: evaluation at different spatial and temporal scales, Wiley Interdisciplinary Reviews: Energy and Environment.

CROOKS, A.T. & HEPPENSTALL, A.J., 2012: Introduction to agent-based modelling". In: Agent-based models of geographical systems. In: Heppenstall, A. J., Crooks, A. T., See, L. M. & Batty, M. (Hrsg.), Agent-based models of geographical systems. Springer, 85-105.

ECKMÜLLNER, O., 1964: Probleme des Bauernwaldes; nachgedruckt in: Institut für Forstliche Betriebswirtschaft und Forstwirtschaftspolitik". Eigenverlag, 91-106.

JÖBSTL, H., 1986: Zum Einschlagsverhalten der Forstbetriebe. Holzpreis und Einkommensteuer als Bestimmungsgrößen der Marktanpassung, Eine Daten- und Modellanalyse. Teil 1". In: Zentralblatt für das Gesamte Forstwesen, **1**, 1-14.

HEIMBUCHNER, K., 2017: Graph integration platform - gip. Retrieved from http://www.gip.gv.at/home-en.html

JOHNSTON, K.M., 2013: Agent analyst: Agent-based modeling in ArcGIS. Redlands, CA: Esri Press.

KOSTADINOV, F., HOLM, S., STEUBING, B., THEES, O. & LEMM, R., 2014: Simulation of a Swiss wood fuel and roundwood market: An explorative study. In agent-based modeling, In: Forest Policy and Economics, **38**, 105-118.

MADLENER, R., KOWALSKI, K., & STAGL, S., 2007: New ways for the integrated appraisal of national energy scenarios: the case of renewable energy use in Austria. In: Energy Policy **35 (12)**, 6060-6074.

MANDL, P., 2003: Multi-agenten-Simulation und Raum - Spielwiese oder tragfähiger Modellierungsansatz in der Geographie. Klagenfurter Geographische Schriften, Heft **23**, 5-34.

MARSCHALL, J., 1975: Hilfstafeln für die Forsteinrichtung, Österreichischer Agrarverlag.

O'SULLIVAN, D., MILLINGTON, J., PERRY, G. & WAINWRIGHT, J., 2012: Agent-Based Models – Because They're Worth It?" In: Heppenstall, A. J., Crooks, A. T., See, L. M. & Batty, M. (Hrsg.). Agent-based models of geographical systems. Springer, 109-123.

SACCHELLI, S., FAGARAZZI, C. & BERNETTI, I., 2013: Economic evaluation of forest biomass production in central Italy: a scenario assessment based on spatial analysis tool. In: Biomass and Bioenergy, **53**, 1-10.

SCHWARZBAUER, P., HUBER, W. & STERN, T., 2009: Das Angebotsverhalten der österreichischen Forstwirtschaft. Ökonometrische Schätzungen von Angebotsfunktionen. In: Schriftenreihe des Instituts für Marketing und Innovation, Band **3**.

SCHOLZ, J., MANDL, P., KOGLER, C. & MÜLLER, M., 2014: A Spatial Agent-based Model for Assessment and Prediction of Woodchips Availability for Heating Plants in Austria. In: Extended Abstract Proceedings of the GIScience 2014. Department für Geodäsie und Geoinformation, Technische Universität Wien.

WEINFURTER, P., 2013: Waldbau in Österreich auf ökologischer Grundlage: Eine Orientierungshilfe für die Praxis. Landwirtschaftskammer Österreich, Ländliches Fortbildungsinstitut Österreich.

Entwicklung und Implementierung eines „Weather Extremity Index (WEI)“ zur Bewertung von Starkregenereignissen unter Berücksichtigung räumlicher und zeitlicher Verteilungen

THORSTEN KELM[1], ULRIKE KLEIN[1], FABIAN NETZEL[2],
CHRISTOPH MUDERSBACH[2], MARC KRÜGER[3] & ANGELA PFISTER[3]

Extended Abstract

Durch die in den letzten Jahren immer häufiger vorkommenden extremen Unwetterereignisse und die mediale Berichterstattung sowie die direkte und schnelle Verbreitung privater Fotos und Videos in den sozialen Medien, rücken die auftretenden Extremwetterlagen immer weiter in den Fokus der Öffentlichkeit. Präsente Meldungen über den fortschreitenden Klimawandel unterstreichen diesen Aspekt. Permanent wird über schwere Unwetter, Hochwasser und wochenlange Trockenphasen berichtet. Dadurch steigt stetig das Interesse an der Beschreibung der Schwere solcher Ereignisse.

Bei der Beschreibung von Extremereignissen werden Niederschläge nicht allein durch die Menge des Niederschlages, sondern zusätzlich durch die Abhängigkeit des Zeitraums, in dem dieser gefallen ist, klassifiziert. Dementsprechend kann ein extremes Niederschlagsereignis sowohl von kurzer Dauer mit hoher Intensität sein als auch über mehrere Stunden oder Tage andauernde Niederschläge mit geringen Niederschlagsmengen mit sich bringen. Als Maß der Intensität von Niederschlägen wird die Menge der Niederschlagshöhe in Liter pro Quadratmeter $[l/m^2]$ bzw. in Millimetern $[mm]$ angegeben (PATT & JÜPNER 2013).

Eine erste Einteilung kann entsprechend der Intensität bzw. der Häufigkeit des Auftretens, mithilfe des statistischen Wiederkehrintervalls T_n, erfolgen. Dieses lässt sich im statistischen Mittel als Zeitspanne der Länge von n Jahren, in der ein Ereignis einen Wert einmal erreicht oder überschreitet, definieren (MANIAK 2010; DIN 4049-3 1992). Für die Allgemeinheit ist dies nicht bzw. nur schwer nachvollziehbar, da aus statistischer Sicht nicht der Zeitraum, sondern die Schwere ausgedrückt wird. Beispielsweise wird von der Öffentlichkeit aus der Bezeichnung $T_n = 100\,a$ bzw. „Jahrhundertereignis“ fälschlicherweise oft abgeleitet, dass ein Ereignis mit diesem Ausmaß erst in 100 Jahren erneut eintritt. Daher wird seit einigen Jahren versucht, die Schwere von Ereignissen durch einen Index anhand einer definierten Skala sowie einer verbalen Beschreibung analog zu der von Erdbeben oder Windstärken zu beschreiben. Für Starkregenereignisse kommt dabei deren Lage im Raum und damit einhergehend dem zeitlichen Verlauf des Niederschlages eine große Bedeutung zu. Die bisher mit den eingesetzten Verfahren durchgeführten Analysen liefern keine Aussage über die räumliche Verteilung eines

[1] Hochschule Bochum, Fachbereich Geodäsie, Lennershofstraße 140, D-44801 Bochum,
E-Mail: thorsten.kelm@hs-bochum.de
[2] Hochschule Bochum, Fachbereiche Bauingenieurwesen, Lennershofstraße 140, D-44801 Bochum,
E-Mail: fabian.netzel@hs-bochum.de
[3] Emschergenossenschaft und Lippeverband, Kronprinzenstraße 24, D-45128 Essen,
E-Mail: marc.krueger@eglv.de

Niederschlagsereignisses. Es werden ausschließlich Auswertungen pro Bodenstation durchgeführt. Die Kombination von mehreren Stationen ist mit den bisherigen Verfahren nicht möglich.

Im Rahmen der vorliegenden Masterarbeit wurde in Kooperation mit Emschergenossenschaft und Lippeverband (EG/LV) ein GIS-basiertes Werkzeug zur automatisierten Auswertung des Weather Extremity Index (WEI) nach MÜLLER & KASPAR (2014) entwickelt. Anhand des WEI wird die Extremität bzw. das Ausmaß durch die Berücksichtigung der Seltenheit über die interpolierten statistischen Wiederkehrzeiten und deren räumliche Ausdehnung sowie ihre zeitliche Dauer beschrieben. Nach MÜLLER & KASPAR (2014) stehen die Intensität und die räumliche Ausdehnung in einem direkten Verhältnis. Demnach lässt sich ein extremes Wetterereignis dadurch beschreiben, dass es selten auftritt und eine große Ausdehnung besitzt (MÜLLER & KASPAR 2014).

Als Fallbeispiel wurde hierzu das Stadtgebiet von Dortmund in Nordrhein-Westfalen, das zum Großteil innerhalb der Verbandsgrenzen von EG/LV liegt, ausgewählt. Dieses Stadtgebiet eignet sich zum einen aufgrund der vorliegenden Datengrundlage und zum anderen durch die in der Vergangenheit zahlreich aufgetretenen schweren Unwetterereignisse gut zur Plausibilisierung und Verifizierung der WEI-Methodik.

Die Ausgangsdaten bildeten die langjährigen und lückenbereinigten Niederschlagszeitreihen mit einer zeitlichen Auflösung von einer Minute und Zeitreihenlängen von teilweise mehr als 80 Jahren, die von der EG/LV als betreibenden Wasserwirtschaftsverband zur Verfügung gestellt wurden. Diese Datenreihen ermöglichen über die Extremwertstatistik eine ortsbezogene Aussage zu den Wiederkehrzeiten. Das Arbeitsblatt DWA A-531 (DWA 2012) definiert 21 Dauerstufen D von fünf Minuten bis sechs Tagen und Wiederkehrzeiten T_n von ein bis 100 Jahren. Die aufgezeichneten langjährigen Zeitreihen der terrestrischen Niederschlagsstationen stellen dabei Stichproben aus der Grundgesamtheit aller aufgetretenen Niederschläge dar, wobei aus den Parametern der Stichprobe mit geeigneten Verteilungsfunktionen auf die Grundgesamtheit geschlossen werden kann und Regenhöhen mit einer statistischen Eintrittswahrscheinlichkeit bzw. Wiederkehrzeit über Extrapolationen bestimmt werden können.

Zur Berechnung des Index galt es zunächst die Datengrundlage zu schaffen. Für jede der verwendeten Bodenstationen wurde eine extremwertstatistische Auswertung durchgeführt. Zur Ableitung der statistischen Wiederkehrzeiten T_n werden die Regenhöhen h_N aus den Niederschlagsmessungen der langjährigen terrestrischen Niederschlagsstationen für das zu untersuchende Zeitfenster ermittelt. Diese Werte bilden die Basis für eine räumliche Interpolation, bei der die ortsbezogenen statistischen Wiederkehrzeiten auf ein Untersuchungsgebiet übertragen werden und so die Ausgangsdaten des WEI bestimmen.

Nach der Umsetzung des GIS-Werkzeugs wurde es auf drei signifikante Unwetter aus 1949, 1968 und 2008 im Untersuchungsgebiet angewendet. Im Anschluss wurden deren Ergebnisse gegenübergestellt und miteinander verglichen. Alle Ereignisse führten zu schweren Schadensausmaßen. Allerdings wurde deutlich, dass sich das Ausmaß und die Lage sowie die Dauer deutlich unterscheiden. Durch die Anwendung des WEI konnte eine statistisch gesicherte Beschreibung der Schwere durchgeführt werden.

Der WEI eignet sich für das analysierte Untersuchungsgebiet, Niederschlagsereignisse über die Seltenheit und räumliche Ausdehnung zu klassifizieren und zu charakterisieren sowie anhand

von verschiedenen Kennwerten miteinander zu vergleichen. Jedoch wurde deutlich, dass dies nicht ohne Einschränkungen möglich ist. Der WEI ist kein dimensionsloser Index, ohne Wertebegrenzung und somit ohne Obergrenze. Diese Vorgabe musste für die Masterarbeit aufgrund der Datengrundlage eingeschränkt werden. Durch die Begrenzung auf ein Untersuchungsgebiet und damit Festlegung der maximalen Ausdehnung wird eine obere Grenze eingeführt. Infolge der Normierung lassen sich WEI verschiedener Niederschlagsereignisse in fest definierten Untersuchungsgebieten miteinander vergleichen. Bisher konnte sich allerdings keiner der eingesetzten Ansätze komplett durchsetzen, sodass in Zukunft weiterhin Forschungsbedarf besteht, da das Interesse und die Nachfrage aufgrund der immer häufiger auftretenden Extremereignisse in Deutschland stetig steigt.

Literaturverzeichnis

DIN 4049-3, 1994: Hydrologie. Begriffe zur quantitativen Hydrologie, Beuth Verlag.

DWA, 2012: Arbeitsblatt DWA-A 531: Starkregen in Abhängigkeit von Wiederkehrzeit und Dauer, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e.V. (DWA), Hennef.

MANIAK, U., 2010: Hydrologie und Wasserwirtschaft: Eine Einführung für Ingenieure. Springer.

MÜLLER, M. & KASPER, M., 2014: Event-adjusted evaluation of weather and climate extremes. In: Natural Hazards and Earth System Sciences, **14**(2), 473-483.

PATT, H. & JÜPNER, R., 2013: Hochwasser-Handbuch: Auswirkungen und Schutz. Springer Vieweg.

Deep Convolutional Neural Networks for Semantic Segmentation of Multispectral Sentinel-2 Satellite Imagery: An Open Data Approach to Large-Scale Land Use and Land Cover Classification

LUKAS LIEBEL[1]

Abstract: Large-scale land use and land cover (LULC) maps with up to global coverage still incorporate large amounts of manual work and are thus only available after production times of several years. In the domain of computer-vision, deep convolutional neural networks (CNNs) dominate the state-of-the-art in image recognition, a task closely related to LULC classification. By adopting a state-of-the-art CNN architecture for the classification of multispectral Sentinel-2 image patches, the close relation of both tasks is exploited. Since no suitable dataset for training is available as of now, a quasi-automatic method for deriving labels from existing sources of volunteered geographic information, in particular the OpenStreetMap, is presented. The results indicate that CNNs for image recognition are well-suited for LULC classification with only minor adaptions and are furthermore robust to label noise, originating from the error-prone data acquisition method.

1 Introduction

One of the core fields in earth observation is gathering information about the surface of the earth. The acquired data allows for a better understanding of our living environment and thus serves as a basis for decision making in numerous domains. Spaceborne remote sensing enables applications on a global scale. A typical product created for such applications are *land use and land cover* (LULC) maps, which classify the covered area into regions of similar characteristics regarding their surface coverage and employed usage.

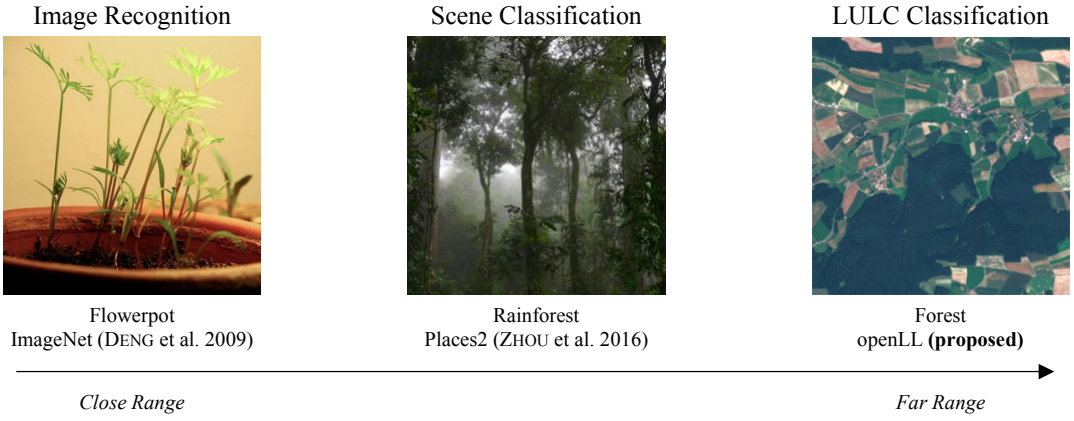


Fig. 1: Relation of image classification tasks from the computer vision domain and LULC classification

[1] Technical University of Munich, Chair of Remote Sensing Technology, Arcisstr. 21, D-80333 München, E-Mail: lukas.liebel@tum.de

Automatic methods for LULC classification exist. However, they often struggle with the difficulties of large-scale LULC classification, in particular the low intra-class and high inter-class similarity.
Pixel-wise spectral features, as used by most traditional methods, are often not expressive enough to cover the variance of the target classes. Recent approaches, thus, consider spatial context by using image regions or patches for classification instead of single pixels. Utilizing image patches for classification hints a close relation to the tasks image recognition and scene classification from the computer vision domain, as illustrated in Fig. 1.

2 Approach

Techniques based on deep learning were able to achieve unmatched results for image recognition in the last years. As LULC classification is closely related to this task, existing methods can be adopted. This can potentially help to overcome shortcomings of the currently employed traditional classification methods. *Convolutional neural networks* (CNN), which represent the state-of-the-art in image recognition, are well-known for being able to generalize well to unseen samples, even in very fine-grained classification. However, CNN architectures for image recognition cannot directly be applied to LULC classification, since remote sensing imagery differs from common photographs in several decisive properties, such as spatial, spectral, and radiometric resolution. Few remote sensing datasets exist, which have been used for adapting image recognition CNNs for LULC classification (PENATTI et al. 2015; CASTELUCCIO et al. 2015). They are, however, not suitable for training a deep CNN for large-scale LULC from scratch, since this task comprises solving optimization problems with millions of parameters.
In order to still be able to evaluate the applicability of this promising technique to the given task, a comprehensive dataset for training was created. Since manual annotations are prohibitively expensive, a method for deriving image patches and ground-truth labels from open data is presented. Multispectral image patches and cloud masks were acquired through the *Copernicus* earth observation program, in particular, the *Sentinel-2* mission. Corresponding ground truth annotations for the creation of the proposed dataset *Open LULC Labels* (openLL) were derived from the *OpenStreetMap* (OSM) database. As semantic information is provided by existing crowdsourced data, the proposed method is considered a quasi-automatic annotation approach. This concept allows for scaling the approach to an arbitrarily large area and number of samples with constant and almost negligible manual effort.
The proposed dataset enables for training a CNN for large-scale LULC classification. With both state-of-the-art performance and available source code for the *TensorFlow* framework, the *Inception-v3* CNN architecture (SZEGEDY et al. 2016) is well-suited and was thus chosen for the experiments. The architecture was adapted to the requirements of multispectral Sentinel-2 image patches and optimized on the training set.

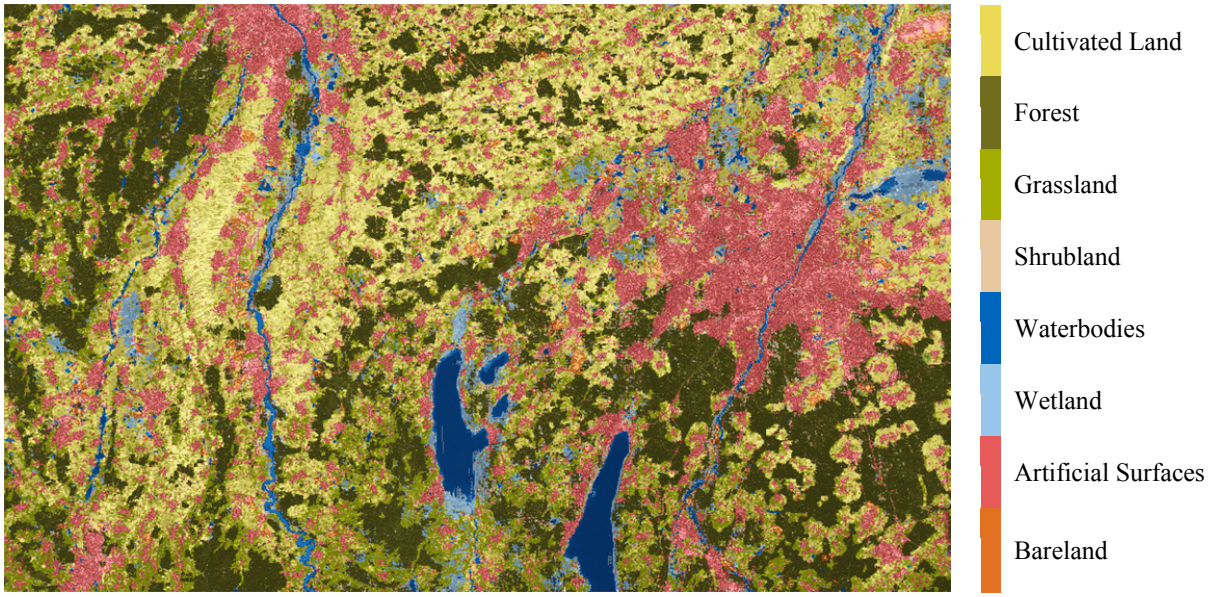


Fig. 2: Land use and land cover map produced by the proposed CNN-based classification approach

3 Results and Conclusion

The classification performance of the trained CNN was assessed using the manually annotated test set of openLL. As a reference, the results were compared to existing LULC maps, namely *GlobeLand30* (GLC30; CHEN et al. 2015) and *CORINE Land Cover* (CLC), which represent the state-of-the-art but involve a large amount of manual work. They achieved a recall of 36% and 44%, and a high overall accuracy of 76.82% and 79.88%, respectively, which was expected. The CNN classifier was applied by using a sliding window approach. With an average recall of 45% and an overall accuracy of 50.72%, the presented approach achieved solid results. Fig. 2 shows a qualitative classification result, enhanced by the suggested multi-look classification procedure. Some of the remaining misclassifications can be traced back to issues in the translation of OSM tags to LULC labels. Further experiments were conducted on the robustness of the classifier to label noise, introduced by the error-prone method employed for creating training data. In addition to that, the propagation of errors from faulty training data to classification results was explored.

For a selected class (clouds opaque), dedicated test areas were manually annotated using a binary classification scheme. A direct comparison to the Sentinel-2 Level-1C cloud masks, used as training data, resulted in a recall of only 5.29% and an overall accuracy of 74.15% for the original cloud masks, and a recall of 21.26% and an overall accuracy of 78.29% for the CNN classification result. The CNN thus proved to be able to cope with conflicting training data.

These results support the hypothesis that an image recognition CNN can successfully be re-trained for the LULC classification task using quasi-automatically annotated training data from open data sources. Since the current implementation is a mere feasibility analysis and is not optimized for achieving competitive classification performance, suggestions for enhancing the proposed approach are given.

4 References

CASTELLUCCIO, M., POGGI, G., SANSONE, C. & VERDOLIVA, L., 2015: Land Use Classification in Remote Sensing Images by Convolutional Neural Networks. Online: arXiv:1508.00092 [cs.CV] (last accessed: 18-01-2017).

CHEN, J., CHEN, J., LIAO, A., CAO, X., CHEN, L., CHEN, X., HE, C., HAN, G., PENG, S., LU, M., ZHANG, W., TONG, X. & MILLS, J., 2015: Global Land Cover Mapping at 30 m Resolution: A POK-based Operational Approach. ISPRS Journal of Photogrammetry and Remote Sensing, **103**, 7-27.

DENG, J., DONG, W., SOCHER, R., LI, L. J., LI, K. & FEI-FEI, L., 2009: ImageNet: A Large-scale Hierarchical Image Database. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 248-255.

PENATTI, O. A. B., NOGUEIRA, K. & DOS SANTOS, J. A., 2015: Do Deep Features Generalize from Everyday Objects to Remote Sensing and Aerial Scenes Domains? Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 44-51.

SZEGEDY, C., VANHOUCKE, V., IOFFE, S., SHLENS, J. & WOJNA, Z., 2016: Rethinking the Inception Architecture for Computer Vision. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2818-2826.

ZHOU, B., KHOSLA, A., LAPEDRIZA, A., TORRALBA, A. & OLIVA, A., 2016: Places: An image database for deep scene understanding. Online: arXiv preprint arXiv:1610.02055 [cs.CV] (last accessed: 18-01-2017).

A Semi-automatic Algorithm for Wetland Detection using Multi-temporal Optical Satellite Data

CHRISTINA LUDWIG[1], STEVE KASS[2] & MICHAEL RIFFLER[1]

Abstract: Wetlands are valuable ecosystems providing a variety of important ecosystem services such as water purification, flood control and food supply. Due to climate change and increasing human water demand, wetlands are increasingly threatened and degraded. An effective monitoring for wetlands is therefore necessary to preserve and restore these endangered ecosystems. In this study, a semi-automatic algorithm for water and wetness detection based on multi-temporal optical imagery and topographic data is presented. Suitable spectral indices sensitive to water and wetness were identified using stability selection. Water and wetness were mapped based on monthly image composites using split-based Otsu thresholding. The monthly water and wetness masks were aggregated to water and wetness frequencies, which were fused in the end to derive a wetland classification and the Water Wetness Probability Index. The algorithm was applied to two study sites in Kenya/Uganda and Niger using Sentinel-2 MSI imagery. For both study sites overall accuracy was above 88%. User's and producer's accuracy were generally higher for water than for wetness, but mostly reaching more than 80% in both classes. Due to the high degree of automation and low processing time, the proposed method is applicable on a large scale and is a first step towards the implementation of global wetland mapping and monitoring service.

1 Introduction

Wetlands are very important ecosystems providing habitat for a variety of flora and fauna as well as many valuable ecosystem services such as flood control and food supply. However, with advancing climate change and increasing human water demand for agriculture, livestock and, cities, wetlands are becoming increasingly threatened and degraded especially in water scarce regions (RAMSAR CONVENTION 2016). In order to preserve and restore wetlands, a cost-effective and efficient monitoring of these ecosystems is necessary. Satellite data offers great potential for wetland monitoring and has been the subject of many studies. Yet, a large scale monitoring system that captures the seasonal dynamics and long-time trends of wetlands has not been implemented so far.

The most common classification methods for wetland mapping are unsupervised (DOGAN et al. 2009; CHEN et al. 2014), supervised (BWANGOY et al. 2010; CORCORAN et al. 2011; FLUET-CHOUINARD et al. 2015) and hybrid classification combining supervised and unsupervised classification approaches (MWITA et al. 2012; LANE et al. 2014). For open water body mapping, DONCHYTS et al. (2016) used Otsu thresholding and a Random Forest classifier based on the

[1] GeoVille Information Systems and Data Processing GmbH, Sparkassenplatz 2, AT-6020 Innsbruck, E-Mail: [ludwig, riffler]@geoville.com

[2] Université du Luxembourg, Faculté des Lettres, des Sciences Humaines, des Arts et des Sciences de l'Education, 11, Porte des Sciences, L-4366 Esch-sur-Alzette, E-mail: steve.kass@uni.lu

mNDWI and HAND index to map water bodies. MARTINIS et al. (2009) applied split-based Otsu for water body mapping using SAR imagery.
In this study, a semi-automatic algorithm for wetland detection using multi-temporal optical imagery and topographic data is proposed. Water and wetness detection is hereby performed using a modified form of split-based Otsu Thresholding. To evaluate the performance of the method in different bio-geographical regions it was applied to two study sites in northern Africa: the Sio-Siteko Wetland (Kenya / Uganda) and the Namga-Kokorou Wetland Complex (Niger). In section 2, the components of the processing chain are outlined. The results of the study sites are briefly summarized in section 3 and section 4 gives a conclusion and outlook.

2 Data and Methods

2.1 Data

The algorithm is mainly based on a time series of high-resolution optical imagery. In this study, the algorithm is demonstrated using Sentinel-2 MSI imagery, but it is also applicable to the Landsat imagery for historical analyses. In addition, a digital elevation model generated from the NASA Shuttle Radar Topography Mission (SRTM) in 2000 is used to calculate the Topographic Wetness Index (TWI).

2.2 Methodology

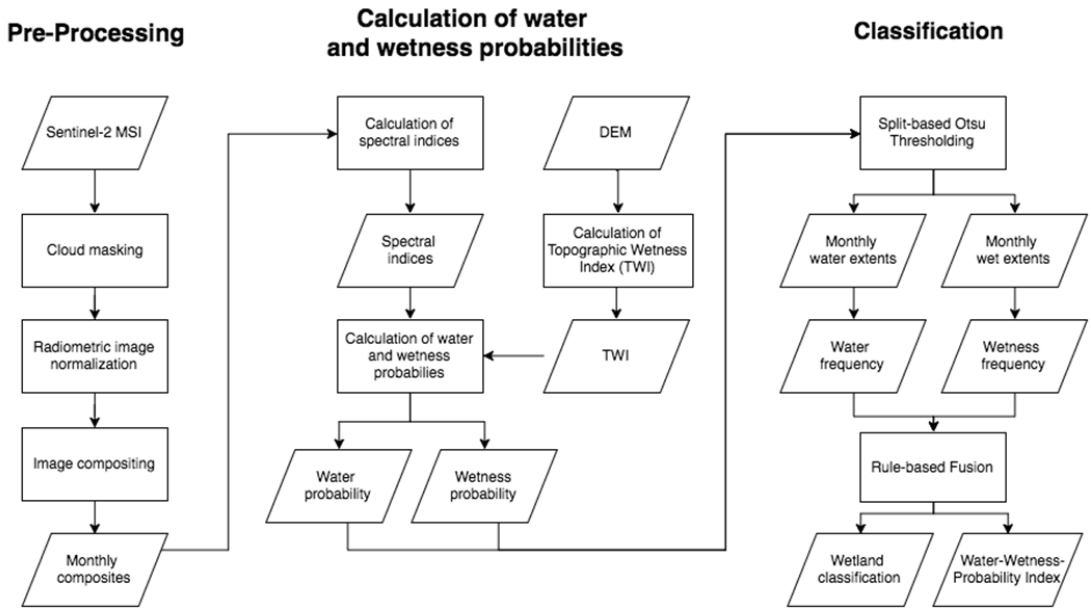


Fig. 1: Graphical workflow diagram of the algorithm for wetland detection

The proposed algorithm consists of three main components: data pre-processing, the calculation of water and wetness probabilities and the wetland classification (Fig.1). During pre-processing cloud and cloud shadow removal is performed for all Sentinel-2 scenes. Subsequently, monthly image composites are produced with all Sentinel-2 scenes belonging to the same composite being radiometrically normalized to each other prior to compositing. For all composites, a

selection of spectral indices sensitive to water and wetness are calculated. Suitable index combinations for water and wetness mapping were identified using a feature selection technique called stability selection. The selected indices and the TWI are subsequently aggregated using different image enhancement techniques to increase the contrast between open water, wetlands and other land cover types. This yields water and wetness probability images for each month. By applying split-based Otsu thresholding to these probability maps monthly water and wetness extents are derived. Aggregating all monthly extents yields the water and wetness frequencies which indicate how often a region is flooded or wet during the observation period. Based on these frequencies the final wetland classification and the Water Wetness Probability Index (WWPI) are derived. To get an estimation of the overall classification accuracy, single water and wetness masks derived from two monthly composites were validated.

3 Results and Discussion

Overall accuracy values for both study sites are very satisfying reaching values above 88%. User's and producer's accuracy were mostly higher for water than for wetness, but mostly reaching more than 80% for both classes. Water was mapped very accurately with omission errors only occurring over floating vegetation (Fig.2) and narrow streams. Considering wetness, most commission errors occurred within forests and agricultural fields, which might however be irrigated at times. Comparisons to precipitation data suggest that the final wetland classification product gives a plausible representation of the presence of water and wetness throughout the year.

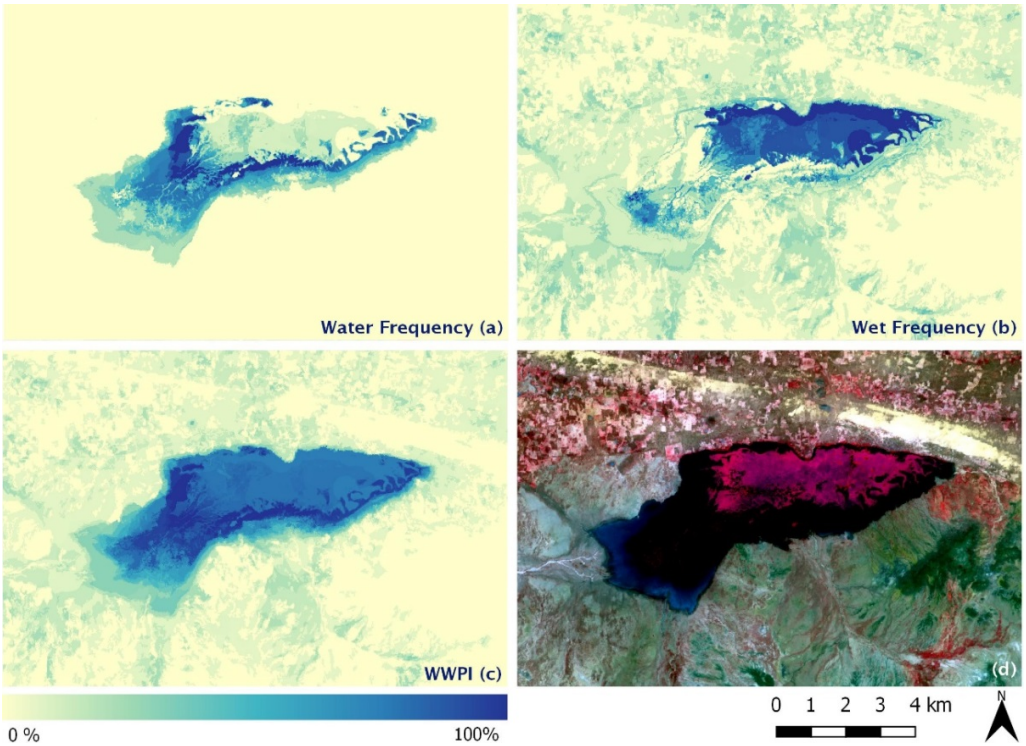


Fig. 2: Water frequency (a), wetness frequency (b) and WWPI (c) of the Kokorou Wetland in the period of December 2015 to October 2016. (d) FCC of the Kokorou wetland in October 2016. Floating vegetation within the lake appearing in red in the FCC image is classified as wet.

4 Conclusion and Outlook

The good validation results and plausible representations of water and wetness occurrence in both study sites prove that the method is applicable in different bio-geographic regions in Africa. Further investigations have to be conducted to increase classification accuracy and to evaluate the method's applicability in other climatic regions. Considering the relation between classification quality and computational effort, the method has the potential to be applicable for large-scale water and wetness mapping and monitoring. In this way, the method has already been successfully applied using additional SAR data within the production of the Copernicus Pan-European High Resolution Layer 2015 and the European Space Agency's GlobWetland Africa project.

5 References

RAMSAR CONVENTION, 2016: An Introduction to the Ramsar Convention on Wetlands. Ramsar Convention Secretariat, Gland, Switzerland, 7th edition.

BWANGOY, J.-R. B., HANSEN, M. C., ROY, D. P., DE GRANDI, G. & JUSTICE, C. O., 2010: Wetland mapping in the congo basin using optical and radar remotely sensed data and derived topographical indices. Remote Sensing of Environment, **114**, 73-86.

CHEN, L., JIN, Z., MICHISHITA, R., CAI, J., YUE, T., CHEN, B. & XU, B., 2014: Dynamic monitoring of wetland cover changes using time-series remote sensing imagery. Ecological Informatics, **24**, 17-26.

CORCORAN, J., KNIGHT, J., BRISCO, B., KAYA, S., CULL, A. & MURNAGHAN, K., 2011: The integration of optical, topographic, and radar data for wetland mapping in northern minnesota. Canadian Journal of Remote Sensing, **37** (5), 564-582.

DOGAN, O. K., AKYUREK, Z. & BEKLIOGLU, M., 2009: Identification and mapping of submerged plants in a shallow lake using quickbird satellite data. Journal of environmental management, **90**(7), 2138-2143.

FLUET-CHOUINARD, E., LEHNER, B., REBELO, L.-M., PAPA, F. & HAMILTON, S. K., 2015: Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. Remote Sensing of Environment, **158**, 348-361.

LANE, C. R., LIU, H., AUTREY, B. C., ANENKHONOV, O. A., CHEPINOGA, V. V. & WU, Q., 2014: Improved wetland classification using eight-band high resolution satellite imagery and a hybrid approach. Remote Sensing, **6** (12), 12187-12216.

MARTINIS, S., TWELE, A. & VOIGT, S., 2009: Towards operational near-real time flood detection using a split-based automatic thresholding procedure on high resolution terrasar-x data. Natural Hazards Earth System Science, **9**, 303-314.

MWITA, E., MENZ, G., MISANA, S. & NIENKEMPER, P., 2012: Detection of small wetlands with multi sensor data in east africa. Advances in Remote Sensing, **1**, 64-73.

A Novel Approach to the Routing Problem of Overhead Transmission Lines

NADINE PIVETEAU[1], JORAM SCHITO[2], MARTIN RAUBAL[2] & ROBERT WEIBEL[1]

Zusammenfassung: Obwohl das Stromnetz in Europa veraltet ist und dringend ausgebaut werden sollte, schreitet dessen Modernisierung nur langsam voran. In dieser Masterarbeit wird ein neuer Ansatz zur Modellierung der optimalen Trasse für Übertragungsleitungen vorgestellt, um den Planungs- und Bewilligungsprozess zu beschleunigen. Im Gegensatz zu bisherigen Forschungsarbeiten werden dabei nicht nur raumplanerische und umwelt- sowie landschaftsbezogene Aspekte, sondern auch technische Kriterien berücksichtigt. Der neue Ansatz wurde in einem gebirgigen Untersuchungsgebiet zwischen zwei Schweizer Ortschaften getestet und einer standardmäßig eingesetzten Methode gegenübergestellt. Die Ergebnisse haben ergeben, dass der neue Ansatz zur Trassenbestimmung Potential hat, den Planungsprozess von Übertragungsleitungen zu verbessern.

1 Einleitung

Der Ausbau des existierenden Stromnetzes ist infolge des steigenden Stromkonsums und der zunehmenden Dezentralisierung der Stromerzeugung durch die Energiewende in den Ländern Europas dringend nötig (RENDIGS 2016). Aufgrund langwieriger Planungs- und Bewilligungsprozesse sowie der geringen Akzeptanz der Bevölkerung gegenüber Übertragungsleitungen schreiten dessen Ausbau und dessen Modernisierung jedoch nur langsam voran. Die im Rahmen dieser Masterarbeit durchgeführte Studie zeigt, dass die Forschung in der Geoinformatik einen wichtigen Beitrag zur Beschleunigung von Netzausbauprojekten leisten kann, indem Produkte zur Vereinfachung von Planungsprozessen und zur Förderung der Akzeptanz entwickelt werden.

2 Problemstellung

Ein konventioneller Ansatz zur Modellierung der optimalen Trasse für Übertragungsleitungen besteht darin, eine Route zu finden, welche die geringsten Kosten verursacht. Dabei sind Kosten in diesem Kontext nicht im monetären Sinn zu verstehen; im Gegensatz dazu beschreiben sie die Auswirkung und die Konflikte, die der Bau einer Übertragungsleitung auf Mensch und Natur auslösen kann. Diese Kosten können in drei Dimensionen zusammengefasst werden:

- Dimension Raumplanung: Diese Dimension beschreibt, wie die menschliche und urbane Umwelt von Übertragungsleitungen hinsichtlich Distanz und Sichtbarkeit betroffen ist.
- Dimension Umwelt und Landschaft: Diese Dimension umfasst Faktoren zum Schutz von Umweltschutzbiotopen und schützenswerter Landschaften.

[1] Universität Zürich, Geographisches Institut, Winterthurerstrasse 190, CH-8057 Zürich, Schweiz
E-Mail: nadine.piv@gmail.com, robert.weibel@geo.uzh.ch
[2] ETH Zürich, Institut für Kartografie und Geoinformation, Stefano-Franscini-Platz 5, CH-8093 Zürich,
E-Mail: [jschito, mraubal]@ethz.ch

- Dimension technische Kriterien: Diese Kriterien sind eng mit ökonomischen Faktoren gekoppelt und beschreiben die bautechnische Umsetzbarkeit der Übertragungsleitungen.

In dieser Masterarbeit wurden die oben aufgelisteten Dimensionen durch die Anwendung eines neuen Verfahrens optimiert. Diese Optimierung führte dazu, dass der Einfluss der modellierten Übertragungsleitung auf die natürliche und bebaute Umwelt minimiert und deren technische Umsetzbarkeit sichergestellt werden konnte.

3 Forschungsstand

Zahlreiche Forschungsarbeiten haben sich mit der Berechnung der optimalen Trasse für Übertragungsleitungen befasst. Die häufigste Vorgehensweise beruht auf einer rasterbasierten Least Cost Path (LCP) Analyse, die aus drei Schritten besteht (STEFANSKI & KAVOURAS 1995; BERRY 2004):

1. Die Definition der diskreten Kostenoberfläche
2. Die Berechnung der akkumulierten Kostenoberfläche
3. Die Bestimmung des LCP

Die diskrete Kostenoberfläche wird aufgrund einer Multikriterienanalyse berechnet und als Raster dargestellt, wobei jeder Rasterzellwert die Kosten für die Durquerung dieser Zelle ausdrückt (BAGLI ET AL. 2011). Aufgrund der Zellwerte der diskreten Kostenoberfläche wird die akkumulierte Kostenoberfläche generiert, indem, vom Ursprung zum Ziel und umgekehrt, jeder Rasterzelle die minimale akkumulierte Kostenentfernung zugewiesen wird. Der LCP zwischen zwei Ortschaften entspricht der Abfolge von Zellen mit abnehmenden Werten in der akkumulierten Kostenoberfläche („steepest downhill path“) (BERRY 2004).

Obwohl rasterbasierte Ansätze effizient und einfach sind, weisen sie Nachteile auf. Vor allem bautechnische Faktoren, die den Bau und die Kosten von Übertragungsleitungen prägen, werden vernachlässigt.

4 Methode

Im Gegensatz zu bisherigen Forschungsansätzen basiert diese Masterarbeit auf der Graphentheorie. Die Knoten im Graphen repräsentieren die Masten und die Kanten stellen die Leitungen zwischen zwei Masten dar. Die optimale Trasse entspricht dem Pfad im Graph, der die Dimensionen Raumplanung, Umwelt und Landschaft berücksichtigt und gleichzeitig die bestmögliche bautechnisch umsetzbare Lösung bietet. Folgende technische Aspekte fließen in die Berechnung der Trasse ein:

- Das Gelände: Schwieriges Gelände wird gemieden und es wird überprüft, ob der Bodenabstand der Leiter eingehalten wird.
- Der Abspannwinkel: Tragmasten mit kleinem Abspannwinkel (0°-3°) werden bevorzugt, da sie günstiger und bautechnisch einfacher umsetzbar sind als Abspannmasten.
- Die Bündelung: Übertragungsleitungen werden mit linearen Infrastrukturen wie Straßen oder Bahnlinien gebündelt, um die Verbauung und die Auswirkung auf die Umwelt zu minimieren.

Die Kriterien wurden entweder als Bedingung oder als Optimierungsfaktor in der Trassenberechnung berücksichtigt. Beispielsweise werden ein maximaler Abspannwinkel und ein minimaler Bodenabstand der Leitungen zum Boden als Bedingung im Graphen definiert. Indem Optimierungsfaktoren einkalkuliert wurden, berechnet der Algorithmus eine Route, die zusätzlich zu ökologisch sensiblen und bewohnten Gebiete auch schwieriges Gelände vermeidet.

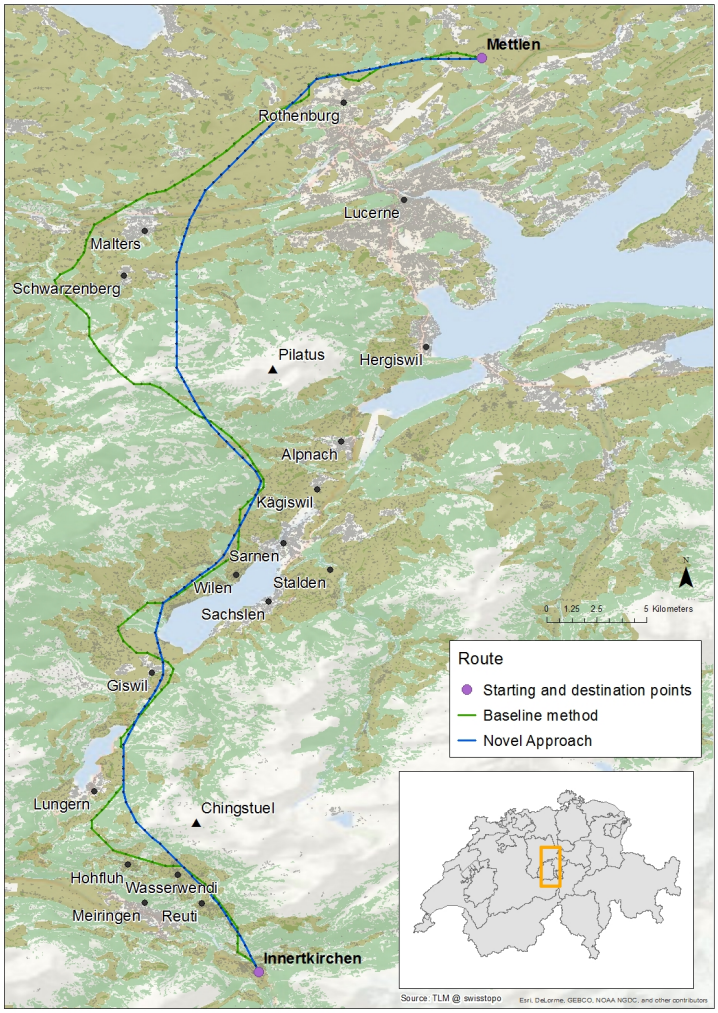


Abb. 1: Die Trasse des neuen Ansatzes (blau) und der herkömmlichen Methode (grün).

5 Ergebnisse

Der neue Ansatz wurde in einem gebirgigen Untersuchungsgebiet zwischen zwei Schweizer Ortschaften getestet. Um den Mehrwert des neuen Algorithmus zu eruieren, wurde die Trasse des neuen Ansatzes mit der der herkömmlichen Methode anhand einer indikatorbasierten Evaluation und durch Expertenbefragungen verglichen.
Sowohl in der quantitativen als auch in der qualitativen Analyse haben beide Ansätze hinsichtlich der Dimension Umwelt und Landschaft sowie der Dimension Raumplanung ähnlich abgeschlossen. In gewissen Situationen berücksichtigt die Trasse des neuen Algorithmus die Kriterien dieser zwei Dimensionen besser als die herkömmliche Methode. In anderen Situationen ist es jedoch umgekehrt. Deshalb waren sich die Experten einig, dass bei beiden Methoden aufgrund von vor Ort durchgeführten Felduntersuchungen punktuelle Anpassungen an der modellierten Trasse vorzunehmen sind.
Die Autoren konnten zudem wesentliche Unterschiede zwischen den beiden Methoden in Bezug auf die technischen Faktoren feststellen. Die Trasse des neuen Ansatzes ist kürzer, günstiger und geradliniger als die, die durch die herkömmliche Methode berechnet wird (siehe Abb. 1). Zudem weist sie einen höheren Anteil an Tragmasten auf, was geringere Kosten verursacht. Außerdem wird schwieriges Gelände automatisch gemieden, was zusätzliche Kosten spart. Diese Eigenschaften wurden ebenfalls von den befragten Experten beobachtet, gutgeheißen und bestätigt.

6 Fazit & Ausblick

Die Resultate dieser Masterarbeit haben ergeben, dass hauptsächlich technische Kriterien den Verlauf der Trasse entscheidend beeinflussen. Experten maßen dem neuen Ansatz ein hohes Potential zu, beim Abwägen unterschiedlicher Szenarien hilfreich zu sein, aber auch bei der Modellierung von Erdkabelleitungen den Planungsprozess zu vereinfachen. Da der Algorithmus in einem gebirgigen Gelände getestet wurde, sind weitere Studien nötig, welche die Anwendbarkeit auf andere Terrainformen überprüfen. Diese Arbeit ist ein weiterer Baustein im Bestreben, das Stromnetz zu modernisieren, und hat das Potential, sich zu einem leistungsfähigen Tool zur Trassenbestimmung zu entwickeln.

7 Literaturverzeichnis

BAGLI, S., GENELETTI, D. & ORSI, F., 2011: Routeing of power lines through least-cost path analysis and multicriteria evaluation to minimise environmental impacts. Environmental Impact Assessment Review, **31**(3), 234-239.

BERRY, J. K., 2004: Topic 19: Optimal paths and routing. In Map Analysis.

RENDIGS, S., 2016: Transformation des Schweizer Stromübertragungsnetzes – Herausforderungen und Aufgaben für die Raumplanung. ETH Zurich (Doctoral dissertation).

STEFANAKIS, E. & KAVOURAS, M., 1995: On the determination of the optimum path in space. Springer Berlin Heidelberg, Berlin, Heidelberg, 241-257.

Kopplung von Verkehrssimulationen und semantischen 3D-Stadtmodellen in CityGML

ROLAND RUHDORFER[1]

Zusammenfassung: Für die Durchführung von Verkehrssimulationen sind detailgetreue Modelle des Straßenraums erforderlich. Die Erstellung derartiger Simulationsgraphen ist derzeitig mit hohem Aufwand verbunden. Semantische 3D-Stadtmodelle im CityGML Format beinhalten eine Modellierung des Verkehrsraumes. Damit stellen sie eine mögliche Datenquelle dar, um den Modellierungsaufwand vor einer Verkehrssimulation zu reduzieren. Durch theoretische Untersuchungen auf Modellebene und Entwicklung von Konzepten wird aufgezeigt, wie eine Kopplung von Verkehrssimulationen und semantischen 3D-Stadtmodellen in CityGML vollzogen werden kann und welche Potentiale und Herausforderungen darin liegen.

1 Einleitung

Im Hinblick auf einen stetigen Wandel im Bereich Mobilität und Zukunftsthemen wie dem autonomen Fahren gilt es, die Nutzung bestehender Verkehrswege zu optimieren und künftige Planungen nachhaltig und zukunftsorientiert zu gestalten. Dabei stellen Verkehrssimulationen ein wichtiges Werkzeug dar. Grundlage derartiger Werkzeuge ist ein möglichst detailgetreues Abbild der Realität. Die Erstellung oder Aufbereitung entsprechender Modelle für Verkehrssimulationen ist jedoch aktuell noch mit großem manuellem Aufwand verbunden. Bestehende Ansätze, diesen Modellierungsaufwand im Vorfeld zu reduzieren, sind vielfältig und reichen von einer Ableitung eines Simulationsgraphen bis hin zu komplexen Kopplungen von Simulationsanwendungen mit Geoinformationssystemen. Durch ihren hohen Informationsgehalt erscheinen semantische 3D-Stadtmodelle im CityGML Format als vielversprechende Grundlage zur Verknüpfung mit Verkehrssimulationen. Die Masterarbeit beschäftigt sich damit, inwieweit Informationen aus CityGML genutzt werden können und welche Prozessschritte zur Ableitung von Simulationsgraphen für Verkehrssimulationen sowie zur anschließenden Visualisierung notwendig sind.

2 Methoden

Im Mittelpunkt des Interesses einer Kopplung von Verkehrssimulation und semantischen 3D-Stadtmodellen stehen Informationen über den Verkehrsraum. Für die Beschreibung des Verkehrsraumes existieren verschiedene Standards und Modelle. Der internationale Standard für semantische 3D-Stadtmodelle CityGML beinhaltet ein Modul „Transportation Objects“ zur Modellierung des Verkehrsraums. Um das Potential einer Verknüpfung von Verkehrssimulation und Stadtmodellen in CityGML bewerten zu können, wird eine Strukturanalyse von

[1] Technische Universität München, Ingenieurfakultät Bau Geo Umwelt, Lehrstuhl für Geoinformatik, Arcisstraße 21, D-80333 München, E-Mail: [roland.ruhdorfer@tum.de]

Verkehrsraummodellen für Verkehrssimulationen durchgeführt. Die Arbeit untersucht Standards und Modelle aus den Bereichen Simulation und Geoinformation entsprechend der analysierten Kriterien und stellt die Ergebnisse einander gegenüber. Auf Modellebene wird eine Verknüpfung von Stadtmodell und Verkehrssimulation vollzogen. Diese demonstriert, welche Modellkomponenten wie miteinander korrespondieren und bilden die Basis für eine Überführung von Informationen zwischen den Systemen. Der Fokus liegt dabei auf zwei wesentlichen Zielen: (a) eine automatische Ableitung von Simulationsgraphen aus CityGML für die Verkehrssimulation Vissim und (b) eine Visualisierung von Simulationsergebnissen in 3D-Stadtmodellen mit Hilfe virtueller Globen.

3 Ergebnisse

Beide Teilkonzepte sind als prototypische Implementierung in Form eines Plugins für den 3DCityDB Importer/ Exporter umgesetzt. Wie die Durchführung an Stadtmodellen von New York City und München zeigen, können damit innerhalb weniger Sekunden auf Knopfdruck Simulationsgraphen ganzer Stadtteile erzeugt werden. Bisher ist die gängige Vorbereitung einer Verkehrssimulation eine manuelle Erstellung von Simulationsgraphen durch Abdigitalisierung von Karten oder Luftbildern. Das neue Konzept stellt demgegenüber einen enormen Mehrwert dar, auch wenn teilweise eine manuelle Nachbearbeitung des Straßennetzwerkes notwendig ist. Abbildung 1 zeigt das Ergebnis einer automatischen Ableitung für einen Bereich in Lower Manhattan (New York City) und stellt es den CityGML Quelldaten in LoD0 gegenüber. Aus einem CityGML Liniendatensatz wird automatisch ein Vissim Netzwerk bestehend aus Links (blaue Linien) und Konnektoren (magenta Linien) abgeleitet.

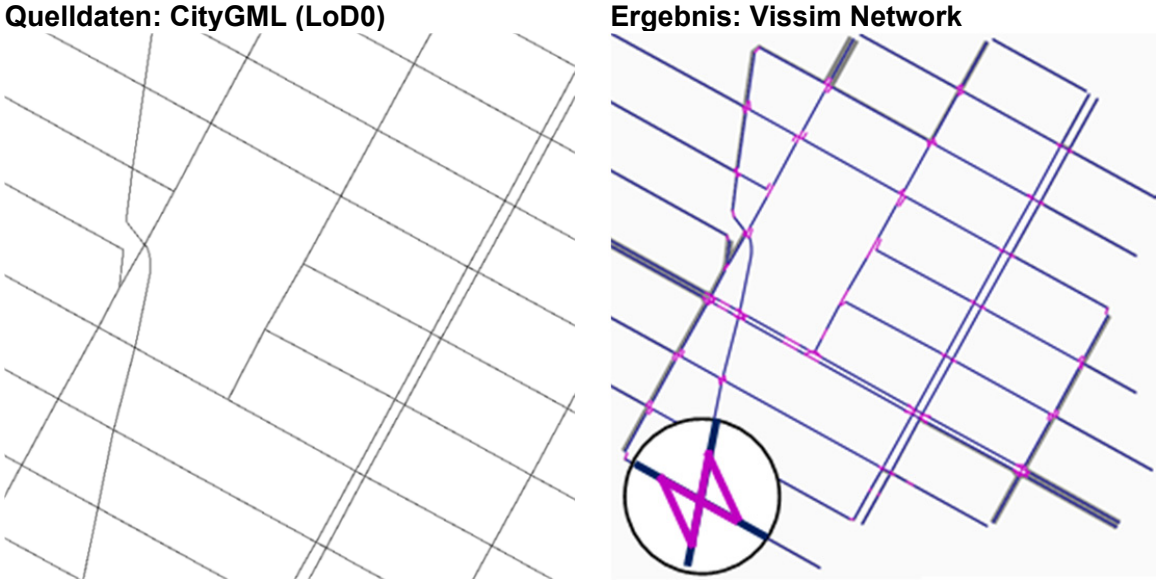


Abb. 1: Testdatensatz aus dem Stadtmodell von New York City - Vergleich der CityGML-Quelldaten mit dem daraus automatisch abgeleiteten Vissim Network mit Links (blau) und Konnektoren (magenta)

Die fotorealistische Visualisierung der aus den Verkehrssimulationen resultierenden Fahrzeugbewegungen ist ein weiteres Ergebnis der Arbeit. Das 3D-Stadtmodell liefert dabei das hochwertige Umgebungsmodell. Für die Darstellung der Verkehrsteilnehmer erlaubt das Konzept eine Einbindung beliebiger 3D-Modelle (Autos, Fußgänger, …) z.B. aus der offenen Online-Bibliothek für 3D-Modelle „SktechUp Warehouse“. Die Implementierung basierend auf dem OGC Standard KML erlaubt eine Visualisierung in Google Earth. Alternativ ist eine Einbindung in den auf Cesium basierenden 3D-Web-Map-Client der 3DCityDB umgesetzt. Abbildung 2 zeigt ein Beispiel für eine Visualisierung von Ergebnissen einer Verkehrssimulation anhand eines Testgebietes in Central Manhattan.

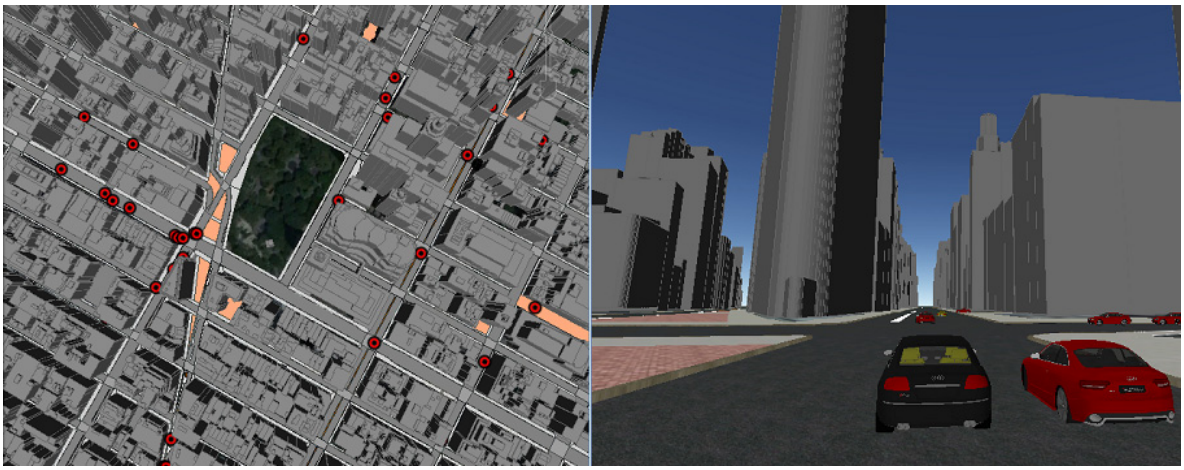


Abb. 2: Visualisierung von Ergebnissen einer Verkehrssimulation in 3D-Stadtmodellen als Punkt-Visualisierung (links) oder mit 3D-Fahrzeugmodellen (rechts)

Ein Resultat der theoretischen Untersuchungen sind Verbesserungsvorschläge für das CityGML Verkehrsmodul, welche in die zukünftige Entwicklung des CityGML Standards einfließen könnten.

4 Fazit & Ausblick

Verkehrssimulationen stellen ein wichtiges Werkzeug dar, den Herausforderungen einer sich wandelnden Mobilität gerecht zu werden. Da die Erstellung von Straßen- und Umgebungsmodellen für Verkehrssimulationen mit großem Aufwand verbunden ist, werden alternative Datenquellen diskutiert. Diese Arbeit überprüft die Potentiale einer Verknüpfung von Verkehrssimulation mit semantischen 3D-Stadtmodellen. Das Ergebnis ist eine weitestgehend automatisierte Prozesskette, die zwei Ziele ermöglicht: (a) eine automatische Ableitung von Simulationsgraphen aus CityGML für die Verkehrssimulation Vissim und (b) eine Visualisierung von Simulationsergebnissen in 3D-Stadtmodellen mit Hilfe virtueller Globen. Zukünftige Arbeiten könnten die Erkenntnisse weiter vertiefen. Sowohl in der Nutzung von Stadtmodelldaten für Verkehrssimulationen, als auch in der Ergebnisrückführung und Visualisierung ist das Potential der Synthese bei weitem noch nicht vollständig ausgeschöpft.

5 Literaturverzeichnis

FELLENDORF, M. & VORTISCH, P., 2010: Microscopic Traffic Flow Simulator VISSIM, Fundamentals of Traffic Simulation, Barceló, J. (Hrsg.), Springer New York

GRÖGER,G., KOLBE, T.H., NAGEL, C. & HÄFELE, K.-H., 2012: OGC City Geography Markup Language (CityGML) Encoding Standard, Version 2.0.0. OGC 12-019, Open Geospatial Consortium

KOLBE, T.H., YAO, Z., NAGEL, C., REDWEIK, R., WILLKOMM, P., HUDRA, G., MÜFTÜOGLU, A. & KUNDE, F., 2016: 3D City Database for CityGML. Version 3.3.0, Lehrstuhl für Geoinformatik, Technische Universität München, virtualcitySYSTEMS GmbH, Berlin & M.O.S.S. Computer Grafik Systeme GmbH, Taufkirchen

RUHDORFER, R., 2017: Kopplung von Verkehrssimulation und semantischen 3D-Stadtmodellen in CityGML. Master Thesis, Lehrstuhl für Geoinformatik, Ingenieurfakultät Bau Geo Umwelt, Technische Universität München.

Fusion of VHR Multispectral and Object Height Data for Urban Environmental Monitoring

CHRISTIAN BERGER[1]

Abstract: Very high resolution (VHR) multispectral and object height data are becoming increasingly available in urban areas. The synergistic utilization of these data holds a large potential for the fine-scale characterization of a city because they are of high descriptive power and non-redundant. However, despite this promising development, detailed and area-wide maps of crucial settlement parameters, like land cover (LC) and urban structure types (USTs), are still lacking in many municipalities. One major reason for this observation is the methodological challenge of turning the wealth of geospatial data into reliable thematic information. My dissertation addresses this problem by (1) developing methods for the fusion of VHR multispectral and object height data, and (2) showcasing the methods' utility in the context of four urban environmental monitoring applications. In this way, both a technical and an applied contribution is made to the fields of urban GIScience and remote sensing.

1 Introduction

Human settlements are complex and dynamic systems having diverse and profound impacts on environmental factors and processes (SCALENGHE & MARSAN 2009). Given the high degree of heterogeneity found in cities, urban environmental monitoring and decision making requires very high resolution (VHR) geospatial data (i.e., ≤5 m spatial resolution) and information products derived thereof (JENSEN & COWEN 1999; KONECNY 1982; WELCH 1982). Thanks to recent technological advancements, reduced production costs, and loosened data policies, this requirement is fulfilled by a growing number of adequate satellite and airborne sensors (BENEDIKTSSON et al. 2012; EHLERS 2009; WENG et al. 2014). VHR multispectral imagery and object height data (e.g., 3D point clouds) are becoming increasingly available in the urbanized regions of the world (LIDAR ONLINE 2014; NOAA 2015; OPENTOPOGRAPHY 2014; SUGARBAKER et al. 2014; USDA 2015). The synergistic utilization of these data sources holds a large potential for the fine-scale characterization of a city because they are of high descriptive power and non-redundant (GAMBA 2014; WENTZ et al. 2014; YAN et al. 2015). However, despite this promising development, detailed, area-wide, and consistent maps of important settlement parameters like urban land cover (LC), urban site characteristics (USCs), and urban structure types (USTs) are still lacking in many municipalities. One reason for this observation is the methodological challenge of turning the wealth of geospatial data into reliable thematic information (LONGLEY et al. 2001; O'NEIL-DUNNE et al. 2013). Accordingly, there is a strong need for accurate and transferable software solutions being able to produce some of the key data sets for human settlement monitoring from VHR multispectral imagery and object height data. My dissertation

[1] Friedrich-Schiller-Universität Jena, Institute of Geography, Department for Earth Observation, Löbdergraben 32, D-07743 Jena, E-Mail: christian.berger@uni-jena.de

(BERGER 2017) aimed at addressing this need. This paper provides a summary of this work for the PFGK18 conference.

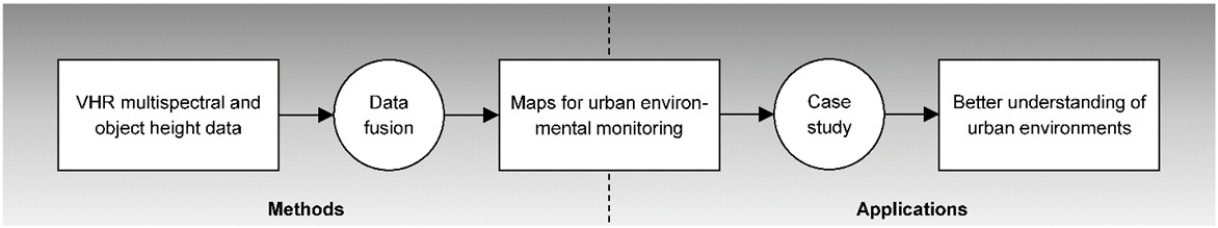


Fig. 1: Scope of my dissertation.

The overall goal of my thesis is to (1) develop methods for the fusion of VHR multispectral imagery and object height data as well as to (2) showcase the methods' utility in the context of four different urban environmental mapping and monitoring applications. The thesis therefore intends to make both a technical ("Methods") and an applied contribution ("Applications") to the fields of urban GIScience and remote sensing (Fig. 1). At the interface of these contributions are urban environmental maps produced by the developed methods and employed within the frame of a predefined case study. More specifically, particular emphasis is put on mapping urban LC, USCs, and USTs, as well as the usage of USCs to study urban land surface temperature (LST) and surface urban heat islands (UHIs; cf. Fig. 3). These settlement parameters were chosen because they are thematically connected, difficult to obtain from other data sources, and of high importance for urban planning and environmental management (CHRYSOULAKIS et al. 2014; JENSEN & COWEN 1999; NICHOL et al. 2007; PAULEIT & BREUSTE 2011; WENTZ et al. 2014). Moreover, they can serve as a starting point to extract further layers of information or to conduct subsequent analyses. From a methodological perspective, much of the research concentrates on geographic object-based image analysis (GEOBIA) (BLASCHKE et al. 2008; BLASCHKE et al. 2014) to achieve the above goal. In addition, the bulk of investigations focuses on test sites located in Germany.

The remainder of this paper is structured as follows. Section 2 defines and differentiates the key terms used in BERGER (2017). Hence, it provides the reader with crucial information for a better understanding of the research conducted. In Section 3, the objectives of my work are deduced from existing research needs and the interconnection between the individual research contributions is introduced by illustrating the practical workflow of the analyses. Section 4 represents the core of this paper. It outlines the research contributions made in BERGER (2017), i.e., the four manuscripts submitted to and accepted by international publication organs with an independent review system. Finally, Section 5 synthesizes the outcomes of my thesis. It summarizes the obtained results and reflects on their implications.

2 Key terms and definitions

In my thesis, each of the four chosen research topics either covers a specific urban environmental phenomenon or a set of settlement parameters belonging together (cf. Fig. 3). The parameters of interest are thematically connected and comprise urban LC, USTs, USCs, and LST. For the sake

of clarity, a definition and differentiation of these and related terms is given in the following. Furthermore, the concept behind the expression of urban environmental monitoring is explained.

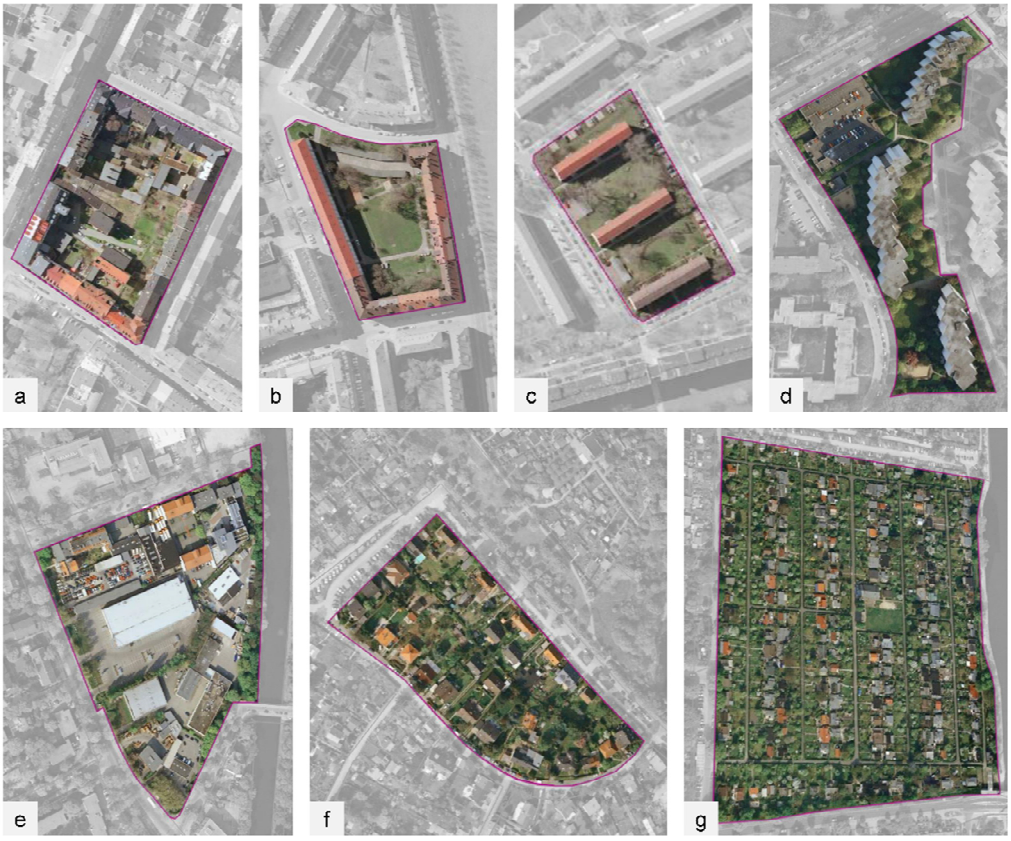


Fig. 2: Visual examples of selected USTs: (a) block development, (b) perimeter block development, (c) row development, (d) high-rise districts, (e) industrial and commercial parks, (f) detached and semi-detached homes, and (g) allotment gardens. Block geometries are highlighted in purple (modified after BOCHOW 2010).

Urban environmental monitoring is conceived here as an umbrella term for mapping, analyzing, and understanding the diverse environmental aspects of a city. As these include, but are not limited to, urban LC, USTs, USCs, and LST, the expression provides an overarching framework to unify the four research subjects selected in my dissertation.

Land cover (LC) is the observed (bio-) physical cover on the Earth's surface (DI GREGORIO 2005; FISHER et al. 2005). Typical examples of (urban) LC elements are buildings, impervious surfaces, trees, grass and shrubs, bare soil, and water (cf. COMBER et al. 2008). An exact definition of LC is crucial because LC is often erroneously confused with land use (LU) (DI GREGORIO 2005).

Land use (LU) is "characterized by the arrangements, activities and inputs people undertake in a certain land cover type to produce, change or maintain it" (DI GREGORIO 2005:3). Therefore, LU refers to the way the land is utilized by people (i.e., function), while LC can be considered as the "visible evidence" of LU (i.e., form; GIRI 2012:9). Residential, industrial, and commercial areas,

parks, sports grounds, and other recreational sites are all common LU types (COMBER et al. 2008).

Urban structure types (USTs) are an ecological concept to divide cities into spatial units of homogenous environmental conditions. A standardized UST classification system does not exist yet (BANZHAF & HÖFER 2008; WURM & TAUBENBÖCK 2010). Existing mapping schemes vary slightly from city to city and usually incorporate a mixture of LU and LC classes (Fig. 2). The spatial delimitation of individual UST entities is based on block geometries, i.e., the streets, railways, and/or topographic features surrounding a block of buildings (cf. GEISS et al. 2011).

Urban site characteristic (USC) is a hypernym specifically defined within my work. The term refers to all kinds of quantitative descriptors that help to map, monitor, and evaluate human settlements. These may include, among others, biophysical proxies like vegetation fraction (VF) and impervious surface area (ISA) (ESCH et al. 2009; RIDD 1995), landscape patterns (BESUSSI et al. 2010; HUANG et al. 2007), 2D and 3D morphology (FINA et al. 2014; YU et al. 2010), and LST (XIAN & CRANE 2006; YUAN & BAUER 2007).

Land surface temperature (LST) is the "average temperature of an element of the exact surface of the Earth calculated from measured radiance" (GILLESPIE 2014). Therefore, it is often referred to as radiometric or skin temperature (LI et al. 2013). LST can either be measured locally by the use of dedicated thermometers or, like in my work, remotely with data acquired by air- and/or spaceborne thermal infrared sensors.

Urban heat islands (UHIs) refer to the observation that cities, towns, and even villages often experience higher air and surface temperatures than their rural surroundings (OKE 1973). They are arguably the most well-documented example of anthropogenic climate modification (ARNFIELD 2003) and can be observed because human settlements have radiative, thermal, aerodynamic, and hydraulic properties that are much different from those of natural landscapes (LANDSBERG 1981).

3 Scientific rationale

An extensive literature survey forms an integral part of my dissertation. The current state of the art in the field of urban environmental monitoring with VHR multispectral and object height data helps identifying past as well as more recent topics of research. But even more importantly, it provides a means to infer those scientific aspects that have gained only little attention so far. This section describes some of the research needs arising from the review conducted in my work and uses them to formulate specific study objectives.

3.1 Research needs

3.1.1 Topic #1: LC mapping

Many studies have already been successful at applying GEOBIA approaches to characterize fine-scale urban LC on the basis of VHR data. However, in spite of previous mapping efforts, only little research in the field has been devoted to building and evaluating robust methods that are capable of processing different sets of VHR multispectral and object height data. So far, existing GEOBIA systems are frequently reported to work well on small spatial subsets (LU et al. 2010;

SALEHI et al. 2012) or for one entire study area (>100 km²) only (PLATT & RAPOZA 2008; WALKER & BLASCHKE 2008). A closer look at the review statistics presented in Table 1 supports this statement. In addition, GEOBIA algorithms often rely on monosensor imagery (TAUBENBÖCK et al. 2010) or even data acquired by one specific sensor (PLATT & RAPOZA 2008; LU et al. 2010; WALKER & BLASCHKE 2008). In order to exploit the full synergistic potential of those VHR multisource data that become increasingly available, suitable fusion techniques and effective practical solutions still

Tab. 1: Selected research deficits arising from the literature review.

Research topic	Review criterion	Number	Percentage
LC mapping	Studies with at least one test site >100 km²	15/115	13.04
	Studies with more than one test site >100 km²	2/115	1.74
	Studies with more than two test sites >100 km²	1/115	0.87
USC assessment	Studies proposing and/or using 2D USCs	78/96	81.25
	Studies proposing and/or using 3D USCs	31/96	32.29
	Studies proposing integrated 2D/3D USCs	4/96	4.17
UST derivation	Studies defining USTs suitable for area-wide mapping	20/38	52.63
	Studies defining USTs suitable for transferable mapping	12/38	31.58
	Studies defining USTs for gap-free and robust mapping	1/38	2.63
USCs vs LST	Studies comparing LST against 2D USCs	119/125	95.20
	Studies comparing LST against 3D USCs	19/125	15.20
	Studies analyzing integrated 2D/3D USCs	0/125	0.00

need to be developed and rigorously tested. The latter requires an in-depth assessment of their performance with regard to classification accuracy as well as spatial, temporal, and sensor-related transferability. An adequate experimental setup should therefore comprise large volumes of VHR data acquired at varying sensing schemes and covering different urban areas with distinct physical structures.

3.1.2 Topic #2: USC assessment

Due to the 3D nature of human settlements, urban environmental monitoring should rely on information sources that account not only for the horizontal spatial dimensions of a city, but also for its vertical spatial dimension. This enables an integrated and more holistic assessment of the "builtscape" (GAMBA et al. 2004). In order to meet these requirements, various 2D and 3D USCs have been proposed and used to date. This holds especially true for the field of urban density mapping by means of remote sensing data and methods. However, despite these developments, it has to be kept in mind that each of the available USCs covers different and distinct features of a city and therefore addresses only specific aspects of human settlement density. This implies that there is a lack of comprehensive USCs being able to interrelate existing and possibly new spatial indicators for a more holistic assessment of density patterns in urban environments. Exceptions to this observation are limited. As Table 1 shows, only four out of 96 reviewed studies (i.e., ca. 4%) put a focus on developing these advanced metrics. Examples include the fuzzy urban index (GOPAL et al. 2016), the Spacematrix (SALOMONS & BERGHAUSER PONT 2012), and a few indices related to urban green (GUPTA et al. 2012; LIU et al. 2015; TOMPALSKI & WEZYK 2012). Apart from these exceptions, there is still a need for integrated spatial indicators taking full advantage

of a combination of 2D and 3D USCs to estimate urban density in its entirety. Their implementation would be straightforward provided that detailed urban LC information as well as VHR object height data are available.

3.1.3 Topic #3: UST derivation

Besides urban LC, GEOBIA approaches have also proven to be successful in the context of several independent UST studies. However, they are part of a relatively young field of science (starting with STEINNOCHER et al. 2001) and there are still some research aspects deserving more attention.

First, existing class hierarchies and definitions are insufficient. To date, investigations are often restricted to a limited number of map categories (STEINIGER et al. 2008; WALDE et al. 2014; WURM & TAUBENBÖCK 2010) or a small spatial subset of the city under consideration (HEIDEN et al. 2012; HU & WANG 2013; LACKNER & CONWAY 2008). Furthermore, even though some of the earlier works deal with an extended set of classes, they rely on site-specific UST legends (e.g., by referring to historic functional use) that are not applicable to other urban areas (cf. BANZHAF & HÖFER 2008; GEISS et al. 2011). Table 1 reflects the above conceptual issues. More specifically, it highlights that there is only one relevant publication (WURM & TAUBENBÖCK 2010) in which UST classes are defined and derived that meet the requirements of map completeness and standardization (BANZHAF & HÖFER 2008; BREUSTE 2010). In order to enable the generation of area-wide (gap-free) and comparable (uniform) UST maps, there is a need for improved class hierarchies with greater depth, enhanced transferability, and, thus, broader applicability. The development of a comprehensive, yet generic UST map key is a prerequisite for the operational monitoring of cities and should therefore be an integral preparatory step within any future study in the field.

Second, the potential of many USCs remains unknown. Based on the form and function of the desired UST categories, a qualitative description and parametrization of all target classes is frequently performed in advance. As a consequence, the majority of existing classification schemes focuses on the extraction of features that have been subjectively predefined and are assumed to be effective according to the conceptual model for each UST of interest (BANZHAF & HÖFER 2008; WURM & TAUBENBÖCK 2010). Such an approach suffers from several drawbacks. One of them is that only a small fraction of potentially useful USCs is considered. Those site properties with no obvious link to any of the anticipated USTs are neglected. Another disadvantage is that the individual and combined descriptive power of all theoretically available USCs is not systematically assessed. Rather, an isolated examination and utilization of apparent physiognomic urban block characteristics takes place. This is one of the reasons why a lot of methods are fairly successful at deriving morphologically distinct USTs but less appropriate for an accurate differentiation of the more common and structurally similar USTs (WURM & TAUBENBÖCK 2010).

To overcome these limitations, the proposed strategies for extracting and selecting USCs need to be reconsidered. For this purpose, a promising approach is to systematically retrieve all kinds of land surface parameters known from the urban remote sensing literature without making specific assumptions about their avail. While some of these parameters may already be established, others, like landscape metrics (e.g., patch density) and several 3D USCs (e.g., VV2BV), may

have never been used before in the context of UST mapping. After the derivation of these variables, it is crucial to statistically evaluate their suitability in the course of a dedicated, importance-based assessment. The latter facilitates the identification of their individual potential as well as their synergies.
In summary, the above suggestions ensure a comprehensive analysis and objective comparison of USCs for a better characterization of USTs. Their implementation would help compiling a new set of UST features, developing more sophisticated class descriptions, and improving previously achieved mapping accuracies.

3.1.4 Topic #4: USCs vs LST

A considerable amount of research has already been directed towards understanding the links between USCs and LST. Most notably, a lot of emphasis has been placed on the 2D features of a city. Table 1 confirms this tendency by indicating that 119 out of 125 publications (i.e., ca. 95%) deal with 2D metrics. In contrast to that, 3D (15%) and integrated 2D/3D USCs (0%) still remain largely unexplored although their impact on the urban climate is unquestioned (OKE 1981; UNGER 2004) and their area-wide calculation has become feasible thanks to the increased availability of appropriate remote sensing data (cf. Section 1). Except for a few cases (e.g., CHUN & GULDMANN 2014; NASSAR et al. 2016; SCARANO & SOBRINO 2015), there is still a lack of studies relating 2D, 3D, and integrated 2D/3D USCs to LST (cf. MALLICK et al. 2013; NGIE et al. 2014; VOOGT & OKE 2003). Moreover, previous research findings are often based on the analysis of a single study area, a limited number of USCs, and/or only a few LST scenes acquired within specific seasons (e.g., LIU & ZHANG 2011; MAIMAITIYIMING et al. 2014; REN et al. 2013; WENG et al. 2004; XIAO et al. 2008; ZHANG et al. 2009). In order to obtain a more complete picture of the connection between USCs and LST, more comprehensive investigations still need to be undertaken.

3.2 Study objectives

In accordance with the overall goal defined above (cf. Section 1) and motivated by the four research needs, the specific objectives of my thesis are defined as follows:

1. *Robust extraction of urban LC information*, including the
   - development and application of a robust GEOBIA approach to extract detailed and accurate urban LC information from VHR multispectral and object height data;
   - critical evaluation of the proposed method with regard to its performance for and transferability to different study areas and various sets of VHR multisensor data.
2. *Integrated assessment of 2D and 3D USCs*, including the
   - design and implementation of a new spatial indicator to integratively assess urban density by taking into account the horizontal and vertical key characteristics of a city;
   - critical evaluation of the proposed metric's plausibility and qualification to estimate human settlement density and its distinct spatial patterns for different urban LU types.
3. *Automated derivation of UST maps*, including the

- development and application of a generic GEOBIA approach for an area-wide and automated derivation of USTs from VHR multispectral and object height data;
- critical evaluation of the proposed method regarding classification accuracy, workflow automation, and preparedness for future UST mapping and monitoring tasks.

4. *Detailed analysis of USC–LST links*, including the
   - calculation and analysis of 2D, 3D, and integrated 2D/3D USCs as potential predictors of urban LST as well as potential indicators of the surface UHI phenomenon;
   - detailed inspection of the observed relationships as well as their spatio-temporal dependencies by designing and implementing a comprehensive experimental setup.

Figure 3 provides an overview of the four chosen research topics and shows their thematic interconnection. Moreover, it illustrates the conceptual workflow of the practical investigations conducted within my work and outlines the general research agenda.

4 Research contributions

This section represents the core of my dissertation. It outlines the research contributions made in BERGER (2017), i.e., the four manuscripts submitted to and accepted by international publication organs with an independent review system.

4.1 Robust extraction of urban LC information

In BERGER et al. (2013b), a robust GEOBIA approach was developed to extract detailed and accurate urban LC information from VHR multispectral imagery and object height data. Special attention was paid to the critical evaluation of the proposed method with regard its performance for and applicability to different test sites and various sets of VHR multisensor data. For this purpose, the experimental setup of the study included three cities featuring different physical structures, four sets of VHR optical and light detection and ranging (LiDAR) data, as well as statistical measures to enable the assessment of classification accuracies and methodological transferability. The study results highlight the great potential of the described approach for accurate, robust, and large-area mapping of urban environments. User's and producer's accuracies observed for all maps were almost consistently above 80%, and in many cases even above 90%. Only few larger class-specific errors occurred mainly due to the rather simple (yet effective) assumptions on which the method is based. The presented feature extraction workflow can therefore be used as a template for or starting point of future urban LC mapping efforts.

4.2 Integrated assessment of 2D and 3D USCs

In BERGER et al. (2013a), a new spatial indicator was designed, implemented, and successfully demonstrated within the framework of the 2012 IEEE GRSS Data Fusion Contest. The aim was to integratively assess urban density by taking into account some of the 2D and 3D key characteristics of a city. To this end, fusion of WorldView-2 and LiDAR data was performed. In

a first step, basic urban LC information was extracted from the preprocessed input data sets using the GEOBIA approach developed in BERGER et al. (2013b). The LC map was then utilized in combination with the object height information provided by the LiDAR data to infer urban density. The resulting map and statistics underscore the plausibility and qualification of the proposed metric as a useful measure to evaluate human settlement density and its distinct spatial patterns for different types of urban LU. By taking into account all three spatial dimensions of the urban environment, an integrated and more holistic assessment of the "builtscape" is enabled. To further enhance the transparency of the study, a dedicated geoportal was set up that visualizes all input data, final results, by-products, and ancillary information (EBERLE & BERGER 2012).

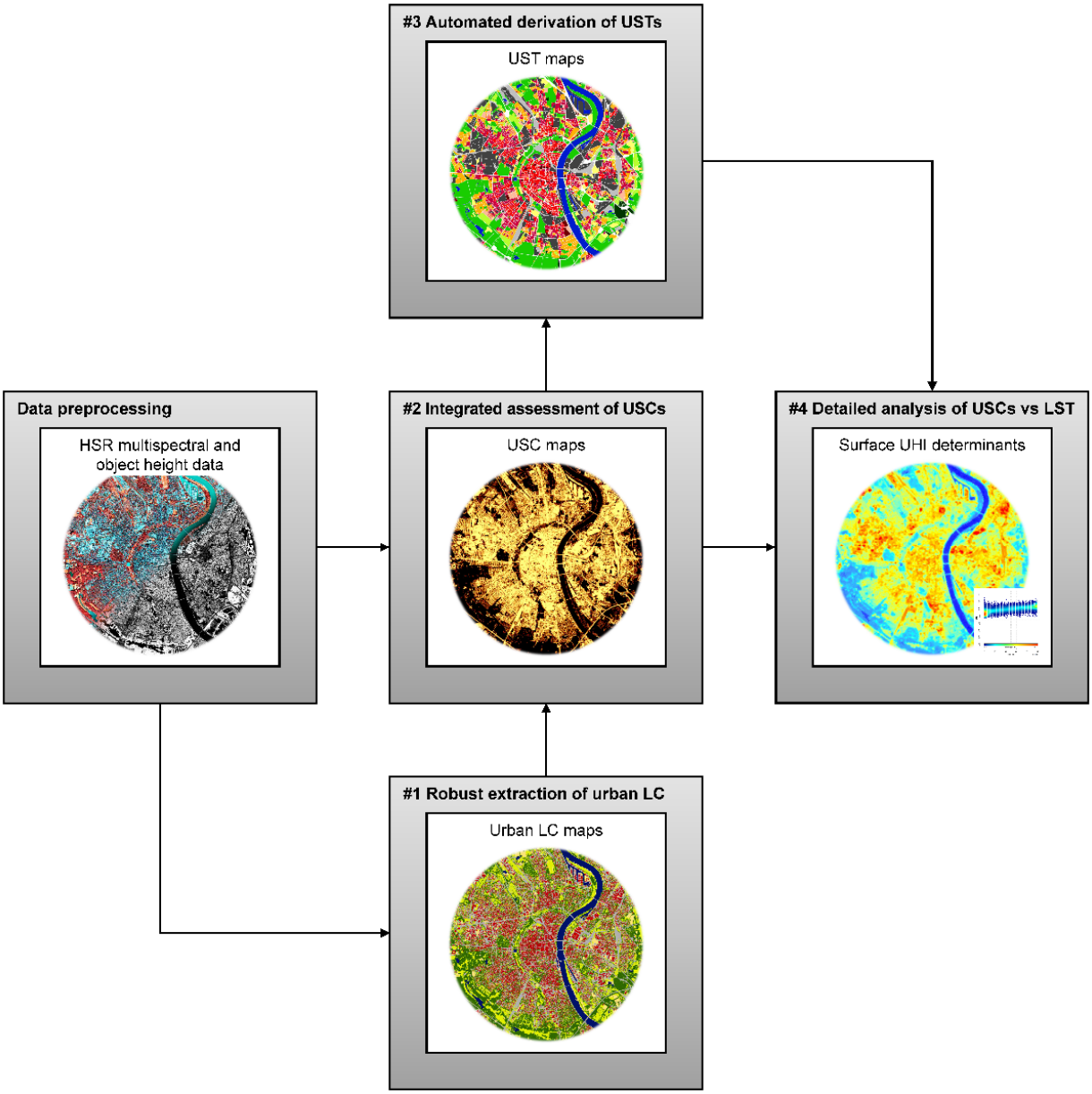


Fig. 3: Practical workflow of my thesis and links between the four chosen research topics.

4.3 Automated derivation of UST maps

In VOLTERSEN et al. (2014), a generic GEOBIA system was developed for an area-wide, automated derivation of USTs. Synergistic use was made of VHR multispectral and object height data acquired by the airborne UltraCamX sensor over the city of Berlin, Germany. Initially, urban LC information was extracted from the preprocessed input data by means of the classification scheme presented in BERGER et al. (2013b). The LC map was then employed together with the UltraCamX object heights to infer a comprehensive set of 2D, 3D, and integrated 2D/3D USCs. In a third step, the importance of the generated image features was evaluated by the random forests machine learning algorithm. Finally, the experiences gained from the assessment were used to build improved UST class descriptions. The latter were based on both well-established as well as more recently introduced USCs, such as the urban density metric proposed in BERGER et al. (2013a). The analysis results and the produced map emphasize the suitability of the described method with regard to classification accuracy, workflow automation, and preparedness in the context of future UST mapping and monitoring tasks. The UST mapping method ranked second in the Student Paper Competition of the JURSE conference in Lausanne, Switzerland (VOLTERSEN et al. 2015).

4.4 Detailed analysis of USC–LST links

In BERGER et al. (2017), the relationship between USCs and LST was studied. Focus was laid on an extensive comparison of 2D, 3D, and integrated 2D/3D USCs as potential predictors of urban LST as well as potential indicators of the surface UHI effect. In addition to the comparison of USCs, the spatio-temporal dependencies of their relation to LST were examined. For this purpose, the experimental setup included two study areas, 26 USCs, and 16 LST scenes covering four seasons. The results of the study demonstrate that the linkage between USCs and LST sensed at small scan angles is not stronger when 3D or integrated 2D/3D parameters are considered. Even though they may offer more holistic representations of the urban landscape, they were consistently outperformed by some of the most widely-used 2D metrics. The analysis of spatial dependencies revealed that the USC–LST interplay does not only differ between, but also within the two test sites. This is due to their distinct geographies, with urban form and compactness, green spaces and street trees, and the structural composition of LC elements being some of the determining factors. The examination of temporal dependencies yielded that the association between USCs and LST is fairly stable over time, but can be subject to larger inter- and intra-season variations for different reasons, including the season of acquisition, vegetation phenology, and meteorological conditions. Since previous research was based on the analysis of a single study area, a limited number of (mainly 2D) USCs, and/or only few LST scenes acquired in specific seasons, the findings of the study provide researchers and practitioners with a more complete picture of the USC–LST relationship.

5 Synthesis

VHR multispectral and object height data are becoming increasingly available in urban areas. The synergistic utilization of these data sets holds a large potential for the fine-scale characterization of a city because they are of high descriptive power and non-redundant.

However, despite this promising development, detailed and area-wide maps of important settlement parameters, like LC, USCs, and USTs, are still lacking in many urban areas. One major reason for this observation is the methodological challenge of turning the wealth of geospatial data into reliable thematic information. Accordingly, there is a strong need for accurate and transferable software solutions being able to produce some of the key data sets for human settlement monitoring from VHR multispectral and object height data. My dissertation (BERGER 2017) aimed at addressing this need. The present paper provided a summary of this work for the PFGK18 conference.

My thesis focused on (1) developing methods for the fusion of VHR multispectral and object height data as well as on (2) showcasing the methods' utility in the context of four different urban environmental mapping and monitoring applications. It therefore aimed at making both a technical and an applied contribution to the fields of urban GIScience and remote sensing. More specifically, particular emphasis was put on mapping urban LC, USCs, and USTs, as well as the usage of USCs to study urban LST and the surface UHI effect. These settlement parameters were chosen because they are thematically connected, difficult to obtain from other data sources, and of high relevance for urban planning and environmental management. To meet the above goal, a comprehensive literature review was conducted in advance. The review helped identifying current deficits within

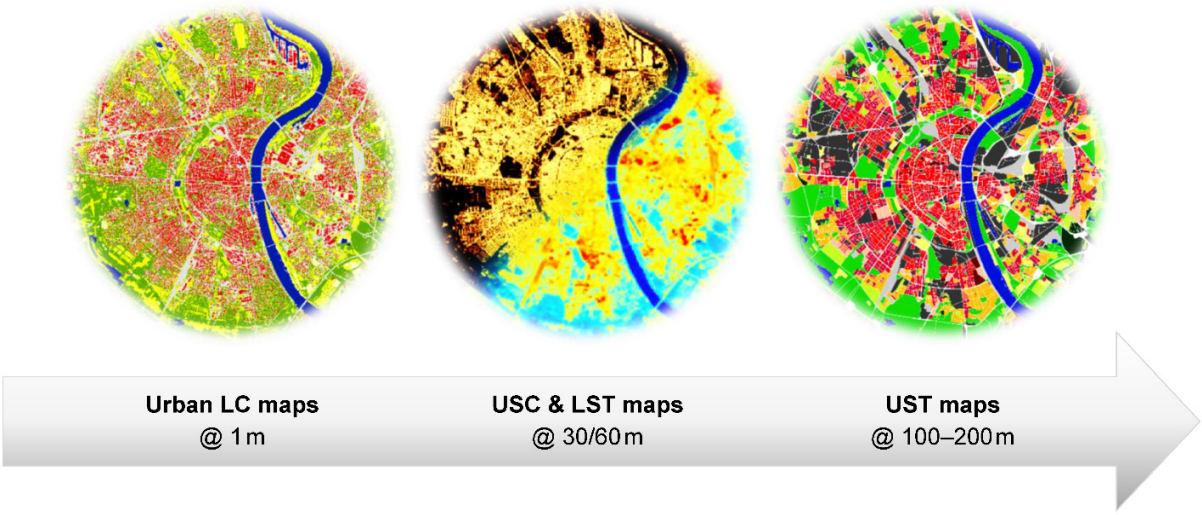


Fig. 4: Spatial resolution of the urban environmental maps produced.

the chosen research fields and led to the formulation of specific objectives. The latter determined the practical agenda of my work, comprising an overall number of four published studies.

With the first three studies, my dissertation has introduced a suite of GEOBIA approaches that build on top of each other. These methods exploit the synergies among VHR multisource data and proved to be suitable for an accurate, yet robust characterization of different settlement parameters, including urban LC, USCs, and USTs. From a geographical perspective, much of my research has concentrated on test sites located in Germany. However, the described classification

and analysis schemes are expected to work well for any other Western culture city as long as the available multispectral and object height data meet some basic requirements (e.g., VHR imagery with four standard spectral bands and 3D data with ≥4 points per m²). Fortunately, these requirements are increasingly fulfilled thanks to a growing number of those VHR multisensor data becoming more and more available in the urbanized regions of the world.
Within the final study of my thesis, the methods developed earlier were effectively employed to produce thematic maps at three spatial scales (Fig. 4). These maps were then used to examine the complex spatio-temporal relation between USCs and LST. This investigation was exemplary for showing that the multilevel information products generated by the fusion of VHR multispectral and object height data are highly versatile and can be reutilized in the context of different follow-up applications. It also illustrates that an individual or joint integration of these maps into other experiments holds the potential for advancing urban research at comparable, coarser, or finer spatial scales. In this sense, the work in BERGER (2017) has provided a continuous and consistent methodological framework for local to regional mapping and monitoring of urban environmental phenomena.

6 Literature

ARNFIELD, A., 2003: Two decades of urban climate research: A review of turbulence, exchanges of energy and water, and the urban heat island. Int J Climatol **23** (1), 1–26.

BANZHAF, E. & HÖFER, R., 2008: Monitoring urban structure types as spatial indicators with CIR aerial photographs for a more effective urban environmental management. IEEE J Sel Top Appl Earth Obs Remote Sens **1** (2), 129–138.

BENEDIKTSSON, J.; CHANUSSOT, J. & MOON, W., 2012: Very high-resolution remote sensing: Challenges and opportunities. P IEEE **100** (6), 1907–1910.

BERGER, C., 2017: Fusion of High Spatial Resolution Multispectral Object Height Data for Urban Environmental Monitoring. <https://www.researchgate.net/publication/318983054_Fusion_of_High_Spatial_Resolution_Multispectral_Object_Height_Data_for_Urban_Environmental_Monitoring> (last accessed on 2018-01-23).

BERGER, C.; ROSENTRETER, J.; VOLTERSEN, M.; BAUMGART, C.; SCHMULLIUS, C. & HESE, S., 2017: Spatio-temporal analysis of the relationship between 2D/3D urban site characteristics and land surface temperature. Remote Sens Environ **193** (C), 225–243.

BERGER, C.; VOLTERSEN, M.; ECKARDT, R.; EBERLE, J.; HEYER, T.; SALEPCI, N.; HESE, S.; SCHMULLIUS, C.; TAO, J.; AUER, S.; BAMLER, R.; EWALD, K; GARTLEY, M.; JACOBSON, J.; BUSWELL, A.; DU, Q. & PACIFICI, F., 2013a: Multi-modal and multi-temporal data fusion: Outcome of the 2012 GRSS Data Fusion Contest. IEEE J Sel Top Appl Earth Obs Remote Sens **6** (3), 1324–1340.

BERGER, C.; VOLTERSEN, M.; HESE, S.; WALDE, I. & SCHMULLIUS, C., 2013b: Robust extraction of urban land cover information from HSR multi-spectral and LiDAR data. IEEE J Sel Top Appl Earth Obs Remote Sens **6** (5), 2196–2211.

BESUSSI, E.; CHIN, N.; BATTY, M. & LONGLEY, P., 2010: The structure and form of urban settlements. In: RASHED, T. & JÜRGENS, C., Eds.: Remote sensing of urban and suburban

areas. Remote sensing and digital image processing 10. Dordrecht, The Netherlands: Springer, 13–31.

BLASCHKE, T.; HAY, G.; KELLY, M.; LANG, S.; HOFMANN, P.; ADDINK, E.; QUEIROZ FEITOSA, R.; VAN DER MEER, F.; VAN DER WERFF, H.; VAN COILLIE, F. & TIEDE, D., 2014: Geographic object-based image analysis – Towards a new paradigm. ISPRS J Photogramm **87**, 180–191.

BLASCHKE, T.; LANG, S. & HAY, G., Eds., 2008: Object-based image analysis for remote sensing. Spatial concepts for knowledge-driven remote sensing applications. Berlin, Germany: Springer.

BOCHOW, M., 2010: Automatisierungspotenzial von Stadtbiotopkartierungen durch Methoden der Fernerkundung. PhD thesis. Osnabrück, Germany: Fachbereich Mathematik/Informatik, Universität Osnabrück.

BREUSTE, J., 2010: Challenges and problems of implementing landscape ecological knowledge in practice – The case of urban development. The Problems of Landscape Ecology **28**, 23–32.

CHRYSOULAKIS, N.; FEIGENWINTER, C.; TRIANTAKONSTANTIS, D.; PENYEVSKIY, I.; TAL, A.; PARLOW, E.; FLEISHMAN, G.; DÜZGÜN, S.; ESCH, T. & MARCONCINI, M., 2014: A conceptual list of indicators for urban planning and management based on Earth observation. ISPRS Int J Geo-Inf **3** (3), 980–1002.

CHUN, B. & GULDMANN, J., 2014: Spatial statistical analysis and simulation of the urban heat island in high-density central cities. Landscape Urban Plan **125**, 76–88.

COMBER, A.; WADSWORTH, R. & FISHER, P., 2008: Using semantics to clarify the conceptual confusion between land cover and land use: The example of 'forest'. J Land Use Sci **3** (2–3), 185–198.

DI GREGORIO, A., 2005: Land cover classification system (LCCS). Classification concepts and user manual. Software version 2. FAO environment and natural resources service series 8. Rome, Italy: FAO.

EBERLE, J. & BERGER, C., 2012: Urban density (UD) mapping results for the City of San Francisco (Geoportal). Contribution to the 2012 IEEE GRSS Data Fusion Contest. <http://sf.maps.essi-blog.org/> (last accessed on 2018-01-23).

EHLERS, M., 2009: Future EO sensors of relevance – Integrated perspective for global urban monitoring. In: GAMBA, P. & HEROLD, M., Eds.: Global mapping of human settlement. Boca Raton, FL: CRC Press, 321–337.

ESCH, T.; HIMMLER, V.; SCHORCHT, G.; THIEL, M.; WEHRMANN, T.; BACHOFER, F.; CONRAD, C.; SCHMIDT, M. & DECH, S., 2009: Large-area assessment of impervious surface based on integrated analysis of single-date Landsat-7 images and geospatial vector data. Remote Sens Environ **113** (8), 1678–1690.

FINA, S.; KREHL, A.; SIEDENTOP, S.; TAUBENBÖCK, H. & WURM, M., 2014: Dichter dran! Neue Möglichkeiten der Vernetzung von Geobasis-, Statistik- und Erdbeobachtungsdaten zur räumlichen Analyse und Visualisierung von Stadtstrukturen mit Dichteoberflächen und -profilen. Raumforsch Raumordn **72** (3), 179–194.

FISHER, P.; COMBER, A. & WADSWORTH, R., 2005: Land use and land cover: Contradiction or complement. In: FISHER, P. & UNWIN, D., Eds.: Re-presenting GIS. Hoboken, NJ: John Wiley & Sons, 85–98.

GAMBA, P., 2014: Image and data fusion in remote sensing of urban areas: Status issues and research trends. Int J Image Data Fusion **5** (1), 2–12.

GAMBA, P.; DELL'ACQUA, F.; CISOTTA, F. & LISINI, G., 2004: High resolution InSAR "builtscape" improvement using LiDAR as ancillary data. Proceedings of IGARSS, 20–24 September, Anchorage, AK, 1808–1811.

GEISS, C.; TAUBENBÖCK, H.; WURM, M.; ESCH, T.; NAST, M.; SCHILLINGS, C. & BLASCHKE, T., 2011: Remote sensing-based characterization of settlement structures for assessing local potential of district heat. Remote Sens **3** (7), 1447–1471.

GILLESPIE, A., 2014: Land surface temperature. In: NJOKU, E., Ed.: Encyclopedia of remote sensing. Encyclopedia of Earth sciences series. New York City, NY: Springer, 314–319.

GIRI, C., 2012: Brief overview of remote sensing of land cover. In: WENG, Q., Ed.: Remote sensing of land use and land cover. Remote sensing applications. Boca Raton, FL: CRC Press, 3–12.

GOPAL, S.; TANG, X.; PHILLIPS, N.; NOMACK, M.; PASQUARELLA, V. & PITTS, J., 2016: Characterizing urban landscapes using fuzzy sets. Comput Environ Urban **57**, 212–223.

GUPTA, K.; KUMAR, P.; PATHAN, S. & SHARMA, K., 2012: Urban Neighborhood Green Index – A measure of green spaces in urban areas. Landscape Urban Plan **105** (3), 325–335.

HEIDEN, U.; HELDENS, W.; ROESSNER, S.; SEGL, K.; ESCH, T. & MUELLER, A., 2012: Urban structure type characterization using hyperspectral remote sensing and height information. Landscape Urban Plan **105** (4), 361–375.

HU, S. & WANG, L., 2013: Automated urban land-use classification with remote sensing. Int J Remote Sens **34** (3), 790–803.

HUANG, J.; LU, X. & SELLERS, J., 2007: A global comparative analysis of urban form: Applying spatial metrics and remote sensing. Landscape Urban Plan **82** (4), 184–197.

JENSEN, J. & COWEN, D., 1999: Remote sensing of urban/suburban infrastructure and socio-economic attributes. Photogramm Eng Rem S **65** (5), 611–622.

KONECNY, G.; SCHUHR, W. & WU, J., 1982: Untersuchungen über die Interpretierbarkeit von Bildern unterschiedlicher Sensoren und Plattformen für die kleinmaßstäbige Kartierung. Bildmessung und Luftbildwesen **50**, 187–200.

LACKNER, M. & CONWAY, T., 2008: Determining land-use information from land cover through an object-oriented classification of Ikonos imagery. Can J Remote Sens **34** (1–2), 77–92.

LANDSBERG, H., 1981: The urban climate. International geophysics series 28. New York, NY: Academic Press.

LI, Z.; TANG, B.; WU, H.; REN, H.; YAN, G.; WAN, Z.; TRIGO, I. & SOBRINO, J., 2013: Satellite-derived land surface temperature: Current status and perspectives. Remote Sens Environ **131**, 14–37.

LIDAR ONLINE, 2014: Lidar Online – Worldwide LiDAR data and geoservices. <https://www.lidar-online.com/> (last accessed on 2018-01-23).

LIU, L. & ZHANG, Y., 2011: Urban heat island analysis using the Landsat TM data and ASTER data: A case study in Hong Kong. Remote Sens **3** (7), 1535–1552.

LIU, Y.; MENG, Q.; ZHANG, J.; ZHANG, L.; JANCSO, T. & VATSEVA, R., 2015: An effective Building Neighborhood Green Index model for measuring urban green space. Int J Dig Earth **9** (4), 387–409.

LONGLEY, P.; BARNSLEY, M. & DONNAY, J., 2001: Remote sensing and urban analysis: A research agenda. In: DONNAY, J.; BARNSLEY, M. & LONGLEY, P., Eds.: Remote sensing and urban analysis. GISDATA 9. Boca Raton, FL: CRC Press, 249–262.

LU, D.; HETRICK, S. & MORAN, E., 2010: Land cover classification in a complex urban-rural landscape with QuickBird imagery. Photogramm Eng Rem S **76** (10), 1159–1168.

MAIMAITIYIMING, M.; GHULAM, A.; TIYIP, T.; PLA, F.; LATORRE-CARMONA, P.; HALIK, Ü.; SAWUT, M. & CAETANO, M., 2014: Effects of green space spatial pattern on land surface temperature: Implications for sustainable urban planning and climate change adaptation. ISPRS J Photogramm **89**, 59–66.

MALLICK, J.; RAHMAN, A. & SINGH, C., 2013: Modeling urban heat islands in heterogeneous land surface and its correlation with impervious surface area by using night-time ASTER satellite data in highly urbanizing city, Delhi- India. Adv Space Res **52** (4), 639–655.

NASSAR, A.; BLACKBURN, G. & WHYATT, J., 2016: Dynamics and controls of urban heat sink and island phenomena in a desert city: Development of a local climate zone scheme using remotely-sensed inputs. Int J Appl Earth Obs **51**, 76–90.

NGIE, A.; ABUTALEB, K.; AHMED, F.; DARWISH, A. & AHMED, M., 2014: Assessment of urban heat island using satellite remotely sensed imagery: A review. S Afr Geogr J **96**, 1–17.

NICHOL, J.; KING, B.; QUATTROCHI, D.; DOWMAN, I.; EHLERS, M. & DING, X., 2007: Earth observation for urban planning and management. State of the art and recommendations for application of Earth observation in urban planning. Photogramm Eng Rem S **73** (9), 973–979.

NOAA (NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION), 2015: Digital coast data access viewer. <http://coast.noaa.gov/dataviewer/#> (last accessed on 2018-01-23).

OKE, T., 1973: City size and the urban heat island. Atmos Environ **7** (8), 769–779.

OKE, T., 1981: Canyon geometry and the nocturnal urban heat island: Comparison of scale model and field observations. J Climatol **1** (3), 237–254.

O'NEIL-DUNNE, J.; MACFADEN, S.; ROYAR, A. & PELLETIER, K., 2013: An object-based system for LiDAR data fusion and feature extraction. Geocarto Int **28** (3), 227–242.

OPENTOPOGRAPHY, 2014: OpenTopography – A portal to high-resolution topography data and tools. <http://www.opentopography.org/> (last accessed on 2018-01-23).

PAULEIT, S. & BREUSTE, J., 2011: Land-use and surface-cover as urban ecological indicators. In: NIEMELÄ, J.; BREUSTE, J.; ELMQVIST, T.; GUNTENSPERGEN, G.; JAMES, P. & MCINTYRE, N., Eds.: Urban ecology: Patterns, processes, and applications. Oxford, UK: Oxford University Press, 19–30.

PLATT, R. & RAPOZA, L., 2008: An evaluation of an object-oriented paradigm for land use/land cover classification. Prof Geogr **60** (1), 87–100.

REN, Z.; HE, X.; ZHENG, H.; ZHANG, D.; YU, X.; SHEN, G. & GUO, R., 2013: Estimation of the relationship between urban park characteristics and park cool island intensity by remote sensing data and field measurement. Forests **4** (4), 868–886.

RIDD, M., 1995: Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: Comparative anatomy for cities. Int J Remote Sens **16** (12), 2165–2185.

SALEHI, B.; ZHANG, Y.; ZHONG, M. & DEY, V., 2012: Object-based classification of urban areas using VHR imagery and height points ancillary data. Remote Sens **4** (8), 2256–2276.

SALOMONS, E. & BERGHAUSER PONT, M., 2012: Urban traffic noise and the relation to urban density, form, and traffic elasticity. Landscape Urban Plan **108** (1), 2–16.

SCALENGHE, R. & MARSAN, F., 2009: The anthropogenic sealing of soils in urban areas. Landscape Urban Plan **90** (1), 1–10.

SCARANO, M. & SOBRINO, J., 2015: On the relationship between the sky view factor and the land surface temperature derived by Landsat-8 images in Bari, Italy. Int J Remote Sens **36** (19–20), 4820–4835.

STEINIGER, S.; LANGE, T.; BURGHARDT, D. & WEIBEL, R., 2008: An approach for the classification of urban building structures based on discriminant analysis techniques. T GIS **12** (1), 31–59.

STEINNOCHER, K.; BAUER, T.; KÖSTL, M. & KRESSLER, F., 2001: Beobachtung von Stadtentwicklung mit Fernerkundung. Applikationen und Innovationen. Österreichische Zeitschrift für Vermessung und Geoinformation **89** (3), 145–148.

SUGARBAKER, L.; CONSTANCE, E.; HEIDEMANN, H.; JASON, A.; LUKAS, V.; SAGHY, D. & STOKER, J., 2014: The 3D Elevation Program initiative – A call for action. Circular 1399. Reston, VA: U.S. Geological Survey.

TAUBENBÖCK, H.; ESCH, T.; WURM, M.; ROTH, A. & DECH, S., 2010: Object-based feature extraction using high spatial resolution satellite data of urban areas. J Spat Sci **55** (1), 117–132.

TOMPALSKI, P. & WEZYK, P., 2012: LiDAR and VHRS data for assessing living quality in cities – An approach based on 3D indices. Int Arch Photogramm Remote Sens Spatial Inf Sci **XXXIX-B6**, 173–176.

UNGER, J., 2004: Intra-urban relationship between surface geometry and urban heat island: Review and new approach. Clim Res **27** (3), 253–264.

USDA (UNITED STATES DEPARTMENT OF AGRICULTURE), 2015: Geospatial data gateway. <https://gdg.sc.egov.usda.gov/GDGOrder.aspx> (last accessed on 2018-01-23).

VOLTERSEN, M.; BERGER, C.; HESE, S. & SCHMULLIUS, C., 2014: Object-based land cover mapping and comprehensive feature calculation for an automated derivation of urban structure types at block level. Remote Sens Environ **154**, 192–201.

VOLTERSEN, M.; BERGER, C.; HESE, S. & SCHMULLIUS, C., 2015: Expanding an urban structure type mapping approach from a subarea to the entire city of Berlin. Proceedings of the 2015 Joint Urban Remote Sensing Event, March 30–April 01, Lausanne, Switzerland, 1–4.

VOOGT, J. & OKE, T., 2003: Thermal remote sensing of urban climates. Remote Sens Environ **86** (3), 370–384.

WALDE, I.; HESE, S.; BERGER, C. & SCHMULLIUS, C., 2014: From land cover-graphs to urban structure types. Int J Geogr Inf Sci **28** (3), 584–609.

WALKER, J. & BLASCHKE, T., 2008: Object-based land-cover classification for the Phoenix metropolitan area: Optimization vs. transportability. Int J Remote Sens **29** (7), 2021–2040.

WELCH, R., 1982: Spatial resolution requirements for urban studies. International Journal of Remote Sensing **3** (2), 139–146.

WENG, Q., Ed., 2014: Global urban monitoring and assessment through Earth observation. Remote sensing applications. Boca Raton, FL: CRC Press.

WENG, Q.; LU, D. & SCHUBRING, J., 2004: Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. Remote Sens Environ **89** (4), 467–483.

WENTZ, E.; ANDERSON, S.; FRAGKIAS, M.; NETZBAND, M.; MESEV, V.; MYINT, S.; QUATTROCHI, D.; RAHMAN, A. & SETO, K., 2014: Supporting global environmental change research: A review of trends and knowledge gaps in urban remote sensing. Remote Sens **6** (5), 3879–3905.

WURM, M. & TAUBENBÖCK, H., 2010: Fernerkundung als Grundlage zur Identifikation von Stadtstrukturtypen. In: TAUBENBÖCK, H. & DECH, S., Eds.: Fernerkundung im urbanen Raum. Erdbeobachtung auf dem Weg zur Planungspraxis. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft, 94–103.

XIAN, G. & CRANE, M., 2006: An analysis of urban thermal characteristics and associated land cover in Tampa Bay and Las Vegas using Landsat satellite data. Remote Sens Environ **104** (2), 147–156.

XIAO, R.; WENG, Q.; OUYANG, Z.; LI, W.; SCHIENKE, E. & ZHANG, Z., 2008: Land surface temperature variation and major factors in Beijing, China. Photogramm Eng Rem S **74** (4), 451–461.

YAN, W.; SHAKER, A. & EL-ASHMAWY, N., 2015: Urban land cover classification using airborne LiDAR data: A review. Remote Sens Environ **158**, 295–310.

YU, B.; LIU, H.; WU, J.; HU, Y. & ZHANG, L., 2010: Automated derivation of urban building density information using airborne LiDAR data and object-based method. Landscape Urban Plan **98** (3–4), 210–219.

YUAN, F. & BAUER, M., 2007: Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in Landsat imagery. Remote Sens Environ **106** (3), 375–386.

ZHANG, Y.; ODEH, I. & HAN, C., 2009: Bi-temporal characterization of land surface temperature in relation to impervious surface area, NDVI and NDBI, using a sub-pixel image analysis. Int J Appl Earth Obs **11** (4), 256–264.

Spatio-Temporal Data Handling for Generic Mobile Geoinformation Systems

PAUL V. KUPER[1]

Abstract: Within this extended abstract, a workflow for an efficient and practical handling of spatio-temporal data is presented. The workflow consists of three layered parts. The first part is the efficient management of spatio-temporal data. The second part focuses on the development of Web services for the dissemination of spatio-temporal data. The third part is a generic mobile GIS for professional users as a typical application for the spatio-temporal data management model and the related Web services.

1 Motivation and Prior Work

The data basis of diverse scientific disciplines consists of data with a spatial reference (TORGE 2003; BARTELME 2005; LI 2015; DAO et al. 2015; WEIHED 2015). Among others PEUQUET (1984), MALLET (2002), WORBOYS (1992), POULIOT et al. (2013), GABRIEL et al. (2015) and BREUNIG et al. (2016) focused on the on the use of simplicial complexes in a 3D space for the modelling of natural surfaces. With a fourth dimension (time) the already huge amount of data that needs to be processed increases. Therefore, the handling of such data is a major challenge.

WORBOYS (1994), GÜTING & SCHNEIDER (2005) and LE (2014) focused on the handling of time within spatio-temporal data. WORBOYS (1994) proposes two separated time dimensions: transaction time and valid time. The transaction time stands for the date on which the data is added into a system, whereas valid time describes the time of the event, which is represented by the data. In some cases additional requirements are needed for the management of time, e.g. for construction projects it is necessary to differentiate between a planning time and a realization time. In such a case the time is also managed in two separated dimensions. The planning time reflects the planning of structures and the realization time reflects the progress of the construction (BREUNIG et al. 2011). When handling time one important aspect is the treatment of time intervals. ALLEN (1984) specified multiple relationships of temporal events. These were used by BREUNIG (2001) in the context of spatio-temporal data management.

Issues that were addressed in former research focus mainly on one of the following subjects: geometry, topology or time. In this abstract a concept is presented that considers all three aspects for an efficient handling of spatio-temporal data and forms the basis for the dissemination of such data for mobile GIS. This extended abstract is a brief overview of the outcome of my dissertation (KUPER 2016).

[1] Karlsruher Institut für Technologie, Geodätisches Institut, Englerstr. 7, D-76131 Karlsruhe, E-Mail: kuper@kit.edu

2 Concept

The workflow for an efficient and practical handling of spatio-temporal data can be divided into three layered parts. The first part focuses on the efficient management of spatio-temporal data itself. The data management model that was developed consists of the five concepts Point Tubes, Delta Storage, Topology Management, Handling of Thematic Data, and Net Components. These concepts were adapted to manage spatio-temporal data which is based on simplicial complexes in a 3D space. The comprehensive model that combines the benefits of these partial concepts was realized within a database management system.
The second part focuses on the development of Web services for the dissemination of spatio-temporal data. For this purpose, two Web services that are based on the OGC standards Web Processing Service and Web Feature Service were developed. The use of OGC standards simplifies the realization of multiple client applications and the combination and integration with existing geodata infrastructures. With the two Web services spatio-temporal data can be distributed and spatio-temporal operations can be processed on a server.
As a third part of the workflow a generic mobile GIS application for professional users was developed as one typical application for the spatio-temporal data management model and the related Web services. First, a universal concept for generic mobile GIS was developed. Afterwards, this concept was extended to include the specific requirements for the handling of spatio-temporal data. An overview of the entire workflow is shown in Figure 1.

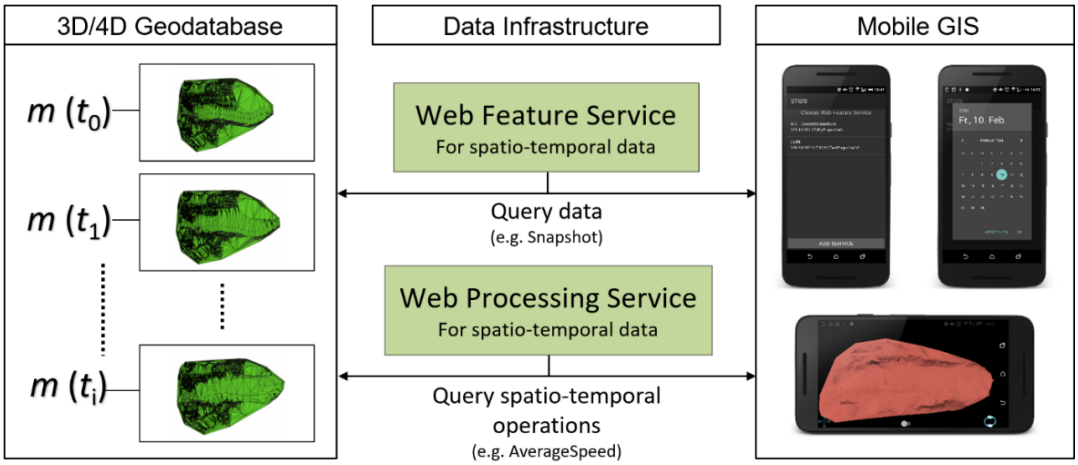


Fig. 1: Overview of the entire workflow based on three layered parts.

Explicit implementations for all three parts of the workflow are provided as free and open source software and are partially integrated into the 3D/4D geodatabase architecture DB4GeO (BREUNIG et al. 2016).

3 Results and Outlook

With the data management model, it is possible to handle various kinds of spatio-temporal applications. Such applications can include massive changes between consecutive time steps and a differentiated handling of thematic and semantic data. Variable temporal discretizations of partial regions of a spatio-temporal model are managed accordingly. It was shown that the data management model reduces the storage requirements for the handling of spatio-temporal models and accelerates computational operations significantly.
In the future the current workflow can be extended and 3D/4D modelling tools such as GOCAD can be integrated. Another promising group of components for the workflow are virtual reality (VR) devices such as Oculus Rift or HTC Vive.

4 References

ALLEN, J. F., 1984: Towards a general theory of action and time. Artificial Intelligence, **23**(2), 123-154.

BARTELME, N., 2005: Geoinformatik: Modelle, Strukturen, Funktionen. 4. vollständig überarb. Auflage, Berlin: Springer, 466p.

BREUNIG, M., 2001: On the Way to Component-Based 3D/4D Geoinformation Systems. Lecture Notes in Earth Sciences, **94**, Springer Berlin Heidelberg, 199p.

BREUNIG, M.; RANK, E.; SCHILCHER, M.; BORRMANN, A.; HINZ, S.; MUNDANI, R.-P.; JI, Y.; MENNINGHAUS, M.; DONAUBAUER, A.; STEUER, H. & VÖGTLE, T., 2011: Towards computer-aided collaborative subway track planning in multi-scale 3D city and building models. Proc. of the 6th International ISPRS Conference on 3D Geoinformation, Wuhan, China, 8p.

BREUNIG, M.; KUPER, P.V.; BUTWILOWSKI, E.; THOMSEN, A.; JAHN, M.; DITTRICH, A.; AL-DOORI, M.; GOLOVKO, D. & MENNINGHAUS M., 2016: The Story of DB4GeO - A Service-Based Geodatabase Architecture to Support Multi-Dimensional Data Analysis and Visualization. ISPRS Journal of Photogrammetry and Remote Sensing, **117**, 187-205.

DAO, T. T.; POULETAUT, P.; CHARLEUX, F.; LAZÁRY, Á.; ELTES, P.; VARGA, P. P. & HO BA THO, M. C., 2015: Multimodal medical imaging (CT and dynamic MRI) data and computer-graphics multi-physical model for the estimation of patient specific lumbar spine muscle forces, Data & Knowledge Engineering, **96-97**, 3-18.

GABRIEL, P.; GIETZEL, J.; LE, H.H. & SCHAEBEN, H., 2015: GST: A Network Based Datastore for Geoscience Data and Geomodels and Its Implementation—ProMine's Contribution Towards Interoperability, 3D, 4D and Predictive Modelling of Major Mineral Belts in Europe. P. Weihed (Ed.), Springer International Publishing, 59-71.

GÜTING, R. H. & SCHNEIDER, M., 2005: Moving Objects Databases. Amsterdam: Boston: Elsevier Science & Technology, 416p.

KUPER, P.V., 2016: Spatio-Temporal Data Handling for Generic Mobile Geoinformation Systems. PhD thesis, 189p.

LE, H.H., 2014: Spatio-temporal Information System for the Geosciences: Concepts, Data models, Software, and Applications, PhD thesis, Technische Universität Bergakademie Freiberg, 169p.

LI, Z., 2015: CAD, 3D Modeling, Engineering Analysis, and Prototype Experimenta-tion: Industrial and Research Applications, Springer, 255p.

MALLET, J.L., 2002: Geomodeling, Oxford University Press, 599p.

PEUQUET, D. J., 1984: A conceptual framework and comparison of spatial data models, Cartographica: The International Journal for Geographic Information and Geo-visualization, vol. 21, no. 4, 66–113.

POULIOT, J.; DANIEL, S.; HUBERT, F. & ZAMYADI, A. EDS., 2013: Progress and New Trends in 3D Geoinformation Sciences, 2013 edition. Heidelberg, New York: Springer, 314p.

TORGE, W., 2003; Geodäsie, de Gruyter, 369p.

WEIHED, P., 2015: 3D, 4D and Predictive Modelling of Major Mineral Belts in Europe, 1st ed. 2015 edition. Springer, 330p.

WORBOYS, M., 1992: A generic model for planar geographical objects, International Journal of Geographical Information Systems, vol. 6, no. 5, 353–372.

WORBOYS, M., 1994: A Unified Model for Spatial and Temporal Information, The Computer Journal, 37(1), 25–34.

Geospatial Data Modelling and Model-driven Transformation of Geospatial Data based on UML Profiles

TATJANA KUTZNER[1]

Extended Abstract

Geospatial data modelling plays an important role in the geospatial domain, not least in the context of current regional, national and international spatial data infrastructure (SDI) initiatives, where more and more conceptual models are being developed using formal modelling languages. In particular, the Unified Modeling Language (UML) is to mention here, which in the geospatial domain is not only popular in research, but also in professional practice and in standardisation. Besides conceptual modelling, the integration of geospatial data is another important task. Especially within the current SDI initiatives the necessity exists to integrate geospatial data based on differing conceptual models into a harmonised data set conforming to one common conceptual model. Provided that the conceptual models are defined using UML, a model-driven transformation approach for integrating the geospatial data can be applied which follows the Model Driven Architecture (MDA) framework.

However, when integrating geospatial data based on different conceptual models, difficulties can arise from the way these models are currently defined within the geospatial domain. Often, the focus is on an informal and graphical representation of the relevant concepts, to the effect that these models contain errors which hinder their successful application in the model-driven transformation approach. Furthermore, the models often adhere to a variety of UML profiles, some of these UML profiles exhibiting deficits which reduce the quality of the models and, thus, also their machine-interpretability. This includes, in particular, UML profiles defined as part of the ISO 191xx series of geographic information standards. For this reason, model-driven transformation of geospatial data is still rarely applied today, in spite of the well-known advantages this approach brings about in general.

The aim of this thesis is to address exactly the problem of the differing and deficient UML profiles and to provide solutions for how to cope with the variety of UML profiles existing in the geospatial domain and with the deficits they exhibit to allow for the creation of high-quality models and the successful integration of geospatial data using the model-driven transformation approach. Starting with a coherent introduction to the most fundamental terms and concepts related to geospatial data modelling and model-driven transformation which also takes into account relevant standards from the standards organisations ISO, OGC and OMG, the state of the art in modelling and model-driven transformation in academia as well as in professional practice is discussed and predominant problems encountered from the way conceptual models are currently defined and used are illustrated.

[1] Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstraße 21, D-80333 München, E-Mail: kutzner@tum.de

Afterwards, a selection of UML profiles currently in use in the geospatial domain is examined. It is discussed, why these UML profiles do not conform to the UML profile definition of the OMG and proposals for how to define these UML profiles in a correct way are provided. Based on these findings, solutions are presented for how to cope with the variety of UML profiles. This includes a generic concept for developing UML profiles in a structured and reusable way, the introduction of a Core UML profile as a universally applicable building block in modelling and model-driven transformation of geospatial data as well as the development of a multi-level information integration framework which allows for transforming between UML models based on differing and deficient UML profiles (cf. figure 1). Finally, the feasibility and applicability of the Core UML profile and of the multi-level information integration framework is demonstrated by applying them to the transformation of geospatial data from Austria, Germany and Switzerland to the European INSPIRE data specifications.

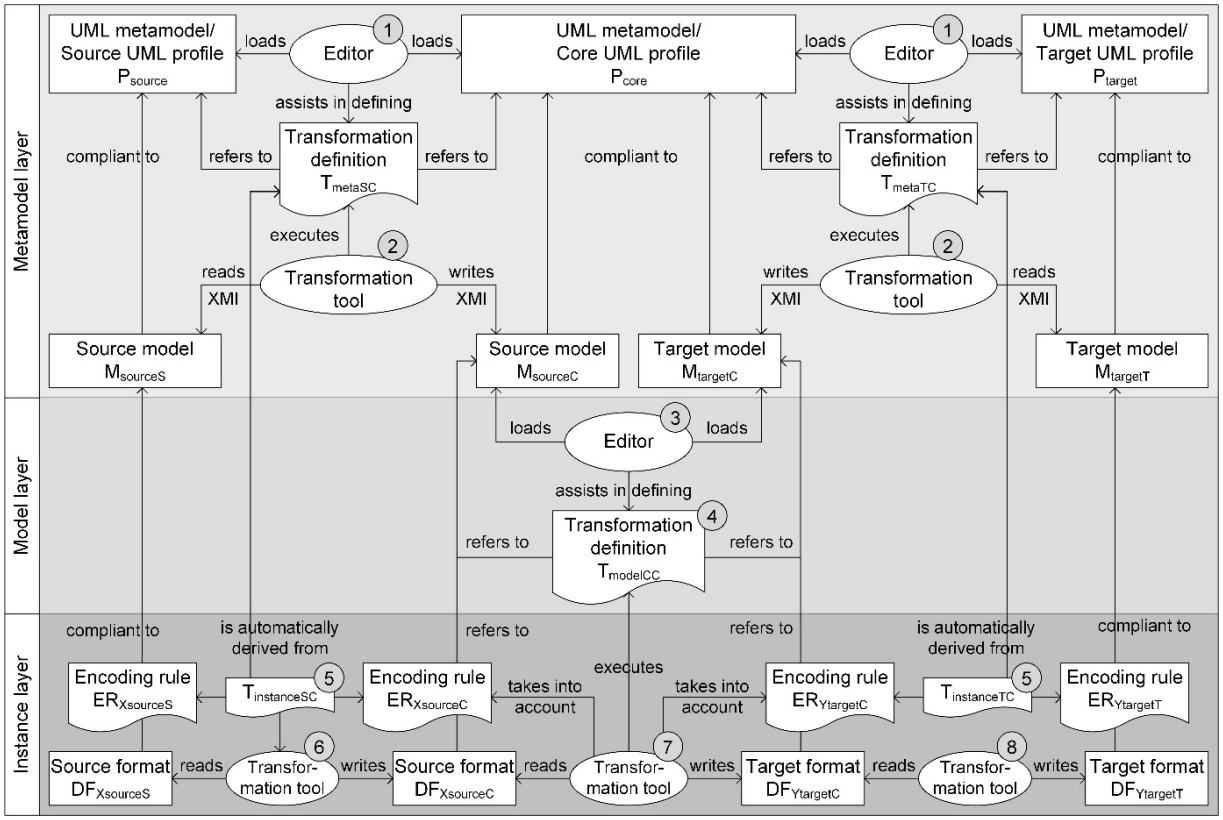


Fig. 1: Framework for multi-level information integration of UML models based on differing and deficient UML profiles

Literature

KUTZNER, T., 2016: Geospatial Data Modelling and Model-driven Transformation of Geospatial Data based on UML Profiles. Dissertation, Technische Universität München.

Reconciling Explanatory and Constructive Modes of Modeling through Deep Modeling

COLIN ATKINSON[1]

Abstract: In the long term, fully exploiting the synergies between ontology engineering and software engineering approaches to modelling is contingent on a complete understanding of their similarities and differences and the development of a unified modeling framework that can support both paradigms in a seamless way. In this talk Colin Atkinson will explain how this can be achieved by characterizing the fundamental modes of modelling the different approaches have evolved to support and by outlining a promising approach for unifying them. This is the potency-based, multi-level modeling paradigm (also known as deep modelling) which allows all ingredients of a comprehensive ontology (e.g. instances, classes and metaclasses) to be modelled in a level-agnostic way, and makes it possible for ontologies to include as many classification levels as needed to best represent the domain in hand (e.g. geographic data). The talk will explain how the basic deep modelling notion of potency can be enhanced to support an "exploratory", ontology-oriented mode of modeling alongside the "constructive", software engineering mode of modeling for which it was originally developed.

1 Overview

As the potential synergy between ontology-based reasoning services (á la OWL) and software-engineering-oriented specification capabilities (á la UML) becomes increasingly apparent, a growing number of researchers have started to explore ways of bringing the two schools of modeling together. In the short to medium term, "bridging" technologies that support mappings between the two paradigms offer the most practical way of jointly leveraging these two forms of modeling. Prime examples include the Ontology Definition Metamodel standardized by the OMG and the TwoUse interoperation tool (PARREIRAS et al. 2007) developed in the MOST project (http://www.mostproject.eu). Nevertheless, in the long term, the cleanest solution is to develop a single, unified modeling paradigm which supports both sets of the capabilities.

Accommodating the different assumptions and interpretations underlying software engineering oriented modeling and ontology (i.e. semantic web) oriented modeling within a single unified framework presents several fundamental challenges, however. Some of these have been identified and widely discussed in the literature. Examples include different underlying logics to define the meaning of classes and taxonomies (e.g. description logic, basic set theory), different interpretation of missing information (e.g. closed-world, open world assumptions) and different architectures used to organize model information (e.g. traditional four level OMG architecture, two level OWL architecture) (ATKINSON et al. 2006; STAAB 2010; JEKJANTUK 2010). Other differences, in contrast, have received relatively little attention and have yet to be fully elaborated. These relate to the fundamental modes by which models are developed and used in the software and ontology engineering communities and the criteria under which models are considered valid. One important difference relates to whether models

[1] Universität Mannheim, Software Engineering Group, Institute für Informatik, B6, 26, Gebäudeteil C, D-68159 Mannheim. E-Mail: atkinson@informatik.uni-mannheim.de

are developed and used in a "constructive" way or in an "exploratory" (aka as "explanatory" (KÜHNE 2016)) way. In constructive (i.e. software engineering oriented) modelling the role of class diagrams is to serve as templates from which populations of instances can be generated at a future point in time, while in exploratory (i.e. semantic web oriented) modeling the role of class models (i.e. ontologies) is to capture classification information wrapped up in an already existing population of instances. This is related to the distinction between "prescriptive" and "descriptive" models (SEIDEWITZ 2003; ASSMANN 2006), but focusses on the conditions under which model content can be considered complete or valid rather than on the purpose for which they are being deployed. Both prescriptive and descriptive models are often developed in exploratory mode and applied in constructive mode.
In this talk, which is based on the paper (ATKINSON et al 2011), Colin Atkinson will fully characterize these different modes of modeling and present a strategy for accommodating them seamlessly within a unified framework. This is based on the multi-level modeling paradigm which allows ontologies to include as many classification levels as needed to best represent the domain in hand and allows their ingredients (e.g. instances, classes and metaclasses) to be treated uniformly as first class citizens (ATKINSON 2009). In the talk Colin Atkinson will – (1) characterize the different modeling modes, (2) introduce and consolidate new terminology to discuss the properties of models in the context of these modes and (3) enhance the notion of potency, the key features that support arbitrary numbers of classification levels, to support the unification of these modes within a single, integrated modeling framework. He will also explain how this unified framework opens up the possibility of achieving a fundamental unification of the ontology engineering and software/database engineering approaches to modeling.

2 References

PARREIRAS, F., STAAB, S. & WINTER, A., 2007: TwoUse - Integrating UML Models and OWL Ontologies. Institut für Informatik, Universität Koblenz-Landau. Nr. 16.

ATKINSON, C., GUTHEIL, M. & KIKO, K., 2006: On the Relationship of Ontologies and Models. 2nd Intl. Workshop on Meta-modeling and Ontologies (WoMM 2006), Karlsruhe.

STAAB, S., WALTER, T., GRÖNER, G. & PARREIRAS, F., 2010: Model Driven Engineering with Ontology Technologies. Reasoning Web. Semantic Technologies for Software Engineering, volume 6325 of LNCS. Springer.

KÜHNE, T., 2016: Unifying Explanatory and Constructive Modeling: Towards Removing the Gulf between Ontologies and Conceptual Models. Proceedings of the ACM/IEEE 19th International Conference on Model Driven Engineering Languages and Systems, 95-102.

SEIDEWITZ, E., 2003: What Models Mean. IEEE Software **20**(5), 26-32.

ASSMANN, U., ZSCHALER, S. & WAGNER, G., 2006: Ontologies, Meta-models, and the Model-Driven Paradigm. Ontologies for Software Engineering and Software Technology, Springer, Berlin, Heidelberg, 249-273.

ATKINSON, C., KENNEL, B. & GÖSS, B., 2011: Supporting Constructive and Exploratory Modes of Modeling in Multi-Level Ontologies, in: 7th International Workshop on Semantic Web Enabled Software Engineering (SWESE 2011), Bonn.

A new Approach to Model Transformation using Graph Transformation System

ZHIHANG YAO[1] & THOMAS H. KOLBE[2]

1 Introduction

Nowadays, many application data models are developed and maintained based on the so-called 'Model Driven Architecture (MDA)' in the fields of software and systems engineering. In general, the MDA stipulates that the data model shall be first defined as an abstract platform-independent model (PIM) and later automatically translated into different platform-specific models (PSM) depending on the target application environments (cf. GASEVIC et al. 2006). A typical case of applying model transformation in real-world applications is the automatic derivation of a relational database schema from an object-oriented data model for realizing the efficient management of large and complex-structured data by making full use of the capabilities of the relational database management systems (RDBMS). According to the literature (cf. NG & LEARMONT 2002), the general conceptual solution is to map the source and target data models onto computer-interpretable formats, such that the model transformation process can be automatically performed by applying a set of user-definable mapping rules. When examining both model representations, it can be seen that they both can be represented by a graph-like structure and can be descriptively formalized as typed and attributed graphs. In addition, the mapping relationships between the different model entities such as classes and tables can also be expressed using directed graph edges connecting the respective graph nodes (cf. Figure 1). Therefore, the model transformation task could be abstracted to a graph transformation problem and accomplished using Graph Transformation Systems (cf. TAENTZER et al. 2005).

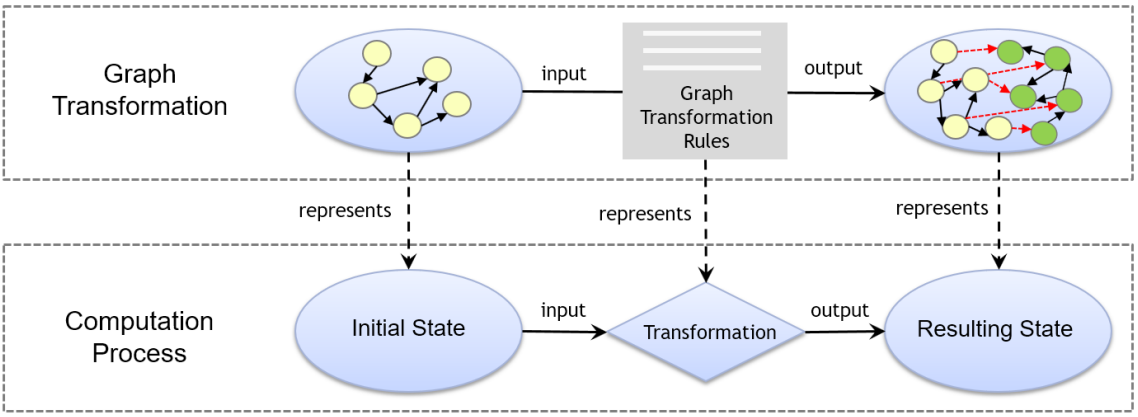


Fig. 1: Conceptual Workflow of using Graph Transformation Systems for performing a model transformation process

[1] virtualcitySYSTEMS GmbH, Marktplatz 2, D-85567 Grafing, E-Mail: zyao@virtualcitysystems.de
[2] Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstr. 21, D-80333 München, E-Mail: thomas.kolbe@tum.de

2 Graph Transformation System

The fundamental concept of using graph transformation systems is an algebraic approach which employs typed and attributed graphs to represent the processing logics at an abstract level for solving specific problems ranging from software engineering up to computation simulations (cf. EHRIG et al. 2005). Basically, the initial state of the computation process shall be first formulated as a typed attributed graph called 'host graph' which is treated as the input for the graph transformation. This transformation process can be carried out by means of a set of user-defined transformation rules called 'graph transformation rules', which uses graph representations to formalize the computation logics and can be denoted as $\{r_1, r_2, \ldots, r_n\}$. Each rule is equivalent to a match morphism denoted as $r: L \rightarrow R$ where L is called left-hand side (LHS) graph, whereas R is called right-hand side (RHS) graph. Both LHS and RHS graphs are also typed and attributed and the LHS graph is considered as a match pattern which could be algebraically isomorphic to every member of those graphs $\{G_1', G_2', \ldots, G_n'\}$ that are subsets of the host graph G_S, where $L \cong G_x'$ and $\{G_1', G_2', \ldots, G_n'\} \subseteq G_S$. While running a graph transformation, each of the given rules will be checked by the graph transformation system. If a match $m(L)$ of LHS graph has been found in the host graph, the respective graph transformation rule will be considered applicable and the matched subgraph in the host graph will be substituted by the corresponding RHS graph. Subsequently, the modified host graph will in turn be treated as the input for the next transformation step. The entire graph transformation process will be continued by successively processing the transformation rules until there is no further match of their LHS graphs can be found in the host graph. After completing the graph transformation process, the computation result can be yielded by interpreting the resulted host graph and can also be graphically displayed to the human operators.

3 Example: Mapping OO-Model onto Relational Model

In this section, an example of using graph transformation systems for model transformations is presented. It is realized using a JAVA-based software tool developed on top of the Open Source graph transformation system engine AGG (cf. TAENTZER 1999) and allows to automatically perform the mapping of an object-oriented model onto a relational database model. The conceptual workflow of the developed approach mainly consists of the following four steps (cf. Figure 2): In the first step, the XML schema definition (XSD) file containing the contents of the input object-oriented model is read and the parsed data model is subsequently mapped onto a graph structure. The respective classes along with their attributes and associations are represented as a set of typed attributed nodes and edges which can fully reflect the semantics of the input data model. In the second step, the mapped graph is passed to the graph transformation system where a number graph transformation rules are predefined to formulate the relevant mapping rules for performing the transformation from the object-oriented model to the desired relational database model. For this step, AGG provides an intuitive graphical editor tool which allows users to conveniently define the transformation rules by dragging and dropping the graph elements in an interactive way. After completing the graph transformation process, a set of new

graph objects representing the derived relational database model as well as the mapping relationships between the source and target model objects are generated.

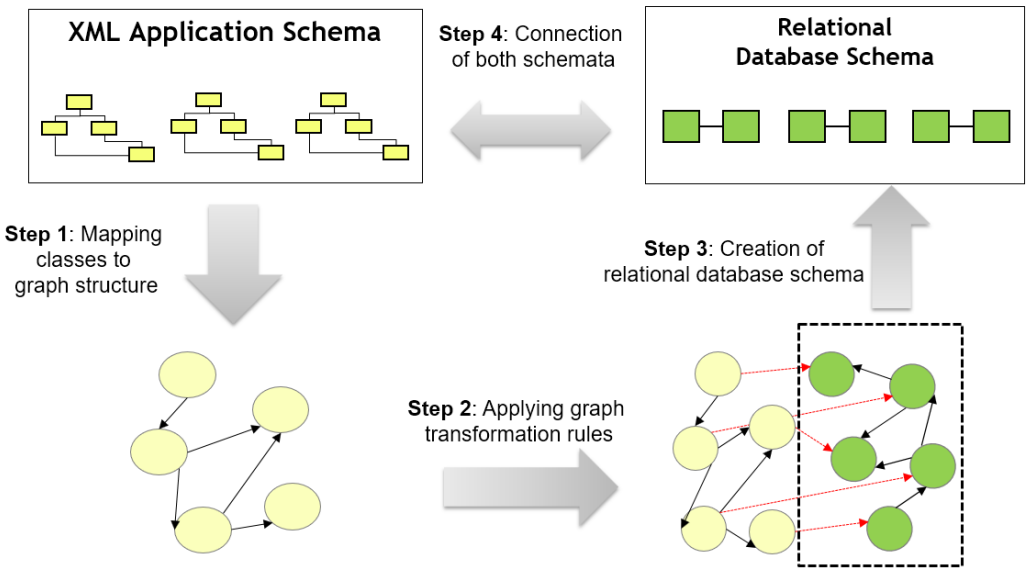


Fig. 2: Workflow of deriving a relational database schema from an XML application schema using graph transformation systems

In the third step, the subgraph representing the newly derived relational database model shall be retrieved from the resulted graph by interpreting the respective graph objects for reconstructing the relational database structure. This can be done by iterating through all those nodes and edges that represent the relational database objects such as database tables, columns, sequences, indexes, and foreign key constraints etc. It is also possible to automatically generate the individual SQL statements for creating different types of database schema according to the specific database products like Oracle and PostgreSQL (cf. YAO & KOLBE 2017). In the last step, the mapping relationships between the XML application schema and the its relational database schema can also be parsed from the output graph and formally expressed using a XML-based document which allows to be easily transmitted between and analyzed in computer applications as well as explored by human users.

4 Conclusions and Outlook

This extended abstract presents a new graph-based approach for realizing model transformations from object-oriented data models to spatially-extended relational database schemas. This approach differs from the classical approaches of OO→relational model transformation, because it allows to express complex and user defined rules (as graph transformation rules), which can transform special configurations of input OO data model entities to special configurations of target relational data model structures. Also optimization rules on the input data model as well as on the resulting relational schema can be expressed using the same language. The concept is currently being employed to automatically derive relational database schemas for CityGML

Application Domain Extensions (ADE) for the Open Source 3D geodatabase 3DCityDB (www.3dcitydb.org).
A brief review of the relevant definitions for graph transformation systems was given first to show their key concepts and mechanism, which allows to facilitate the model transformation task in a systematic way. As a practical example showing how it works, the Open Source graph transformation system tool 'AGG' has been chosen for developing a model transformation tool which allows for automatically deriving a relational database schema from a given XML application schema with an object-oriented data structure. Future work may focus on improving the developed transformation tool by fine-graining or extending the existing transformation rules to make the derived database schema much more efficient in terms of query and processing performance.

5 References

EHRIG, H.; ERMEL, C.; GOLAS, U. & HERMANN, F., 2015: Graph and Model Transformation - General Framework and Applications. Springer Berlin Heidelberg, ISBN: 9783662479797.

GAAEVIC, D.; DJURIC, D.; DEVEDZIC, V. & SELIC, B., 2006: Model Driven Architecture and Ontology Development. Springer New York, Inc., ISBN: 3540321802.

NG, T. C. T. & LEARMONT, T. R., 2002: Rule-based approach to object-relational mapping strategies. U.S. Patent 6,360,223, issued March 19, 2002.

TAENTZER, G.; EHRIG, K.; GUERRA, E.; LARA, J. D.; LENGYEL, L.; LEVENDOVSZKY, T.; PRANGE, U.; VARRÓ, D. & VARRÓ-GYAPAY, S., 2005: Model transformation by graph transformation: A comparative study. Proceedings of Model Transformations in Practice Workshop, 2 -7 October 2005, Models Montego Bay, Jamaica.

TAENTZER, G., 2000: AGG: A Tool Environment for Algebraic Graph Transformation. Applications of Graph Transformations with Industrial Relevance, Lecture Notes in Computer Science, Nagl M., Schürr, A., Münch, M. (eds.), Springer Berlin Heidelberg, 481-488.

YAO, Z. & KOLBE, T.H., 2017: Dynamically Extending Spatial Databases to support CityGML Application Domain Extensions using Graph Transformations. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V. (DGPF), vol 26, Kersten, T. (ed), 37. Wissenschaftlich-Technischen Jahrestagung der DGPF, 7-10 March 2017 in Würzburg, 316-331.

CityGML 3.0: Sneak Preview

TATJANA KUTZNER[1] & THOMAS H. KOLBE[1]

Abstract: The next major version 3.0 of the international OGC standard CityGML is going to be released by the end of 2018. This paper provides an insight into the current developments and presents the most relevant new concepts and modifications which includes new modules as well as revisions of existing modules and concepts.

1 Introduction

Semantic 3D city models represent city objects such as buildings, bridges, tunnels, roads, and vegetation, allowing in this way for an accurate representation of landscapes as well as urban areas. CityGML is the well-established international standard of the Open Geospatial Consortium (OGC) for modelling and exchanging semantic 3D city models (GROEGER et al. 2012). The specification defines both, a data model which can be implemented on different platforms (e.g. Oracle, PostGIS) and an exchange format for transferring data between these platforms.

In order to increase the usability of CityGML for more user groups and areas of application, the OGC CityGML Standards Working Group (SWG) and the Special Interest Group 3D (SIG 3D) of the initiative Geodata Infrastructure Germany (GDI-DE) work since 2014 on the further development of CityGML. This development will result in the next major version *CityGML 3.0* which is to be released by the end of 2018. The new version will bring a number of improvements, extensions, and new functionalities. This paper provides an insight into these developments by presenting the most relevant new concepts and modifications to be included into CityGML 3.0. UML diagrams of the new CityGML 3.0 data model are not provided here, as they are in the current stage of development still subject to changes.

2 New and revised features and modules in CityGML 3.0

The CityGML model is being fully revised to reflect the increasing need for better interoperability with other relevant standards in the field like IndoorGML (LEE et al. 2016), Industry Foundation Classes (IFC) (ISO 2013), Land Administration Domain Model (LADM) (ISO 2012), as well as with Semantic Web Technologies like Resource Description Framework (RDF) (W3C 2014).

All modifications to the new CityGML 3.0 model are carried out in a way to ensure backwards compatibility with CityGML 1.0 and 2.0, with the aim in mind that it should be possible to transform all CityGML 1.0 and 2.0 datasets into the new model by applying syntactical

[1] Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstraße 21, D-80333 München, E-Mail: [kutzner, thomas.kolbe]@tum.de

transformations only. Backwards compatibility is a major requirement for CityGML 3.0 in order to preserve investments by anybody providing CityGML tools, datasets, and extensions.
CityGML 3.0 will apply a model-driven approach in the generation of the data model and exchange formats. The data model is now based directly on the ISO 191xx standards. The data model has been created using Enterprise Architect (EA) and makes use of the EA files provided by ISO and OGC on the underlying standards. The GML encoding (CityGML XML schema) will be fully automatically derived from the data model using the software tool ShapeChange. Also, the feature catalogue with a detailed overview and explanation of all classes, attributes, and relationships will be derived automatically.

2.1 Revised CityGML Level of Detail Concept

CityGML 3.0 will include a revised Level of Detail (LOD) concept. LOD4, which is used for representing the interior of objects (like indoor modelling for buildings and tunnels) has been removed, only the LODs 0-3 will remain. Instead, the interior of objects can be expressed now integrated with the LODs 0-3. It will even be possible to model the outside shell of a building in LOD1 while representing the interior structure in LOD2 or 3. Details on the changes to the CityGML LOD concept are provided in (LÖWNER et al. 2016).

2.2 Revised CityGML Core Model

The CityGML *Core* model has been substantially redesigned as well. A key feature of the new *Core* Model is that all spatial representations are rephrased based on the two pivotal abstract classes *Space* and *SpaceBoundary* from which a number of subclasses are derived. All geometric representations are associated with the semantic concepts of spaces and space boundaries. Furthermore, all feature classes in the thematic modules represent their spatial characteristics almost exclusively using *Space* and *SpaceBoundary* classes and no longer have direct associations with geometry classes. This simplifies geometry handling of CityGML for software developers in general.
The *Core* model implements the new CityGML LOD concept, mentioned above. In addition, a point cloud representation for spaces and space boundaries is being introduced which allows for spatially representing all thematic feature types just by a point cloud. Furthermore, new relationship associations are being introduced to express topologic, geometric, and semantic relations between spaces as well as between space boundaries.
A new class *AbstractToplevelCityObject* will allow to restrict the feature types that are allowed as members of a *CityModel* feature collection. In this way, one can express in the data model that, for instance, from the *Building* Module only the class *Building* can occur as a member of a *CityModel* and that *Door* or *RoofSurface* are not allowed as direct members of *CityModel*. This information is required for software packages that want to filter imports, exports, and visualization according to the general type of a city object (e.g. only show buildings, solitary vegetation objects, and roads).
Another advantage of the space concept is that CityGML objects representing buildings can now directly be mapped onto IndoorGML and vice-versa. Furthermore, spaces which are not bounded by physical objects but are defined according to logical / thematic considerations can now be modelled. This includes building storeys as well as building units like apartments or public

spaces in buildings which are going be represented in CityGML 3.0 based on the new feature type *AbstractBuildingSubdivision* in the *Building* module. The notion of a logical space is also important to express ownership and usage rights as required by the LADM standard.
External references are rephrased and are now better aligned to an RDF representation. Like before, each city object can have an arbitrary number of references to other objects in other datasets / databases, but these can now be additionally qualified by a relation type given by an additional URI.

2.3 New CityGML Construction module

CityGML 3.0 will be complemented by a new *Construction* module which defines concepts that are common to all kinds of man-made constructions like buildings, bridges, tunnels, and other constructions. The module will also comprise the definition of the different kinds of thematic surfaces like *RoofSurface*, *GroundSurface*, *WallSurface* etc., which are passed on to the thematic modules *Building*, *Bridge*, and *Tunnel* by making use of inheritance. This leads to a substantial simplification of the data models of these modules. Man-made structures that are neither buildings, tunnels, nor bridges (e.g. large chimneys, city walls, etc.) can now be modelled by using the new feature type *OtherConstruction*. The interoperability of CityGML with INSPIRE is improved by adopting attributes from the INSPIRE Building data theme (JRC 2013) which allow for specifying multiple elevation levels and measured heights as well as for storing a date of construction, demolition, and multiple dates of renovations. A new feature type *AbstractConstructiveElement* and its concrete building-, bridge-, and tunnel-specific subtypes will allow for mapping constructive elements from BIM datasets given in the IFC standard (e.g. the IFC classes *IfcWall*, *IfcRoof*, *IfcBeam*, *IfcSlab*, etc.) onto CityGML.

2.4 New CityGML Versioning module

Within the new *Versioning* module, CityGML 3.0 will introduce bitemporal timestamps for all objects in alignment with the INSPIRE data specifications. Besides the attributes *creationDate* and *terminationDate* from CityGML 2.0, which refer to the time period a specific version of an object is an integral part of the 3D city model, all objects now can additionally have the attributes *validFrom* and *validTo*, which represent the lifespan a specific version of an object exists in the real world. Furthermore, each geographic feature is being provided with two identifiers: the *identifier* property which is stable along the lifetime of the real-world object, and the *gml:id* attribute which is to mark the respective version of the object. In this way not only the current version of a 3D city model, but its entire history can be represented in CityGML and exchanged. The module will define two new feature types: *Version*, which can be used to explicitly define named states of the 3D city model and denote all the specific versions of objects belonging to such states, and *VersionTransition*, which allows to explicitly link different versions of the 3D city model by describing the reason of change and the modifications applied. Details on the versioning concept are given in (CHATURVEDI et al. 2017).

2.5 New CityGML Dynamizer module

A new *Dynamizer* module has been developed to improve the usability of CityGML for different kinds of simulations as well as to facilitate the integration of sensors with 3D city models. The

integration of sensors with 3D city models is important e.g. in the context of smart cities and digital twins. Both, simulations and sensors provide dynamic variations of some measured or simulated properties like, for example, the electricity consumption of a building. The variations of the value are typically represented using time series data. The data source of the time series data are either sensor observations (e.g. from a smart meter), pre-recorded load profiles (e.g. from an energy company), or the results of some simulation run. A new feature type *Dynamizer* will represent special objects linking the time series data to a specific attribute (e.g. geometry, thematic data, or appearance) of a specific object within the 3D city model. In this way, dynamic variations of city object properties can be injected into an otherwise static representation. For details on the Dynamizer concept and some demonstration applications please refer to (CHATURVEDI & KOLBE 2017).

3 Conclusions

This short paper provided an insight into new features which are to become part of the next major CityGML version 3.0. This new version will provide the new modules *Construction*, *Versioning*, and *Dynamizer*, but will also contain revisions of existing modules and concepts including the LOD concept, the *Core* model, and the *Building* module, and of other modules not mentioned in this paper, such as the *Transportation* module (BEIL & KOLBE 2017). Please note that all changes are subject to the final voting of the OGC CityGML SWG and the Technical Committee.

The CityGML 3.0 standard will consist of at least two parts: 1) the CityGML 3.0 Conceptual Model specification, which is to be released at the end of 2018, and 2) the CityGML 3.0 GML Encoding specification, which is to be published early 2019. Further encoding specifications (e.g. relational database schema, JSON-based representation) may follow in the future.

4 Literature

BEIL, C. & KOLBE, T. H. 2017: CityGML and the streets of New York - A proposal for detailed street space modelling. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., **IV**(4/W5), 9-16.

CHATURVEDI, K., & KOLBE, T. H. (EDS.), 2017: Future City Pilot 1 Engineering Report. OGC Document 16-098. Open Geospatial Consortium.

CHATURVEDI, K., SMYTH, C. S., GESQUIÈRE, G., KUTZNER, T. & KOLBE T. H., 2017: Managing Versions and History Within Semantic 3D City Models for the Next Generation of CityGML. Lecture Notes in Geoinformation and Cartography, Abdul-Rahman, A. (ed.), Advances in 3D Geoinformation, Springer, Cham, 191-206

GRÖGER, G., KOLBE, T. H., NAGEL, C. & HÄFELE, K.-H. (EDS.), 2012: OGC City Geography Markup Language (CityGML) Encoding Standard, Version 2.0.0. OGC Document 12-019. Open Geospatial Consortium.

ISO 2012: ISO/TS 19152:2012 Geographic information – Land Administration Domain Model.

ISO 2013: ISO 16739:2013 Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries.

JRC 2013: D2.8.III.2 Data Specification on Buildings – Technical Guidelines. European Commission Joint Research Centre.

LEE, J., LI, K.-J., ZLATANOVA, S., KOLBE, T. H., NAGEL, C. & BECKER, T. (EDS.), 2016: OGC IndoorGML, Version 1.0.2. OGC Document 14-005r4. Open Geospatial Consortium.

LÖWNER, M.-O., GRÖGER, G., BENNER, J., BILJECKI, F. & NAGEL, C., 2016: Proposal for a new LOD and Multi-Representation Concept for CityGML. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., **IV**(2/W1), 3-12.

W3C 2014: RDF 1.1 specifications. URL: https://www.w3.org/standards/techs/rdf#w3c_all

Einsatz von Multi- und Hyperspektralsensoren in der Landwirtschaft

PATRICK O. NOACK[1]

Zusammenfassung: Multi- und Hyperspektralsensoren werden in der Landwirtschaft für unterschiedlichste Anwendungen eingesetzt. In Erntemaschinen werden aus den Reflektionen des Ernteguts der Trockenmassegehalt und wertbestimmende Parameter abgeleitet. In Güllefässern kommen sie für die Bestimmung des Nährstoffgehalts zum Einsatz. Bei der mineralischen Düngung werden Multispektralsensoren zur Ableitung des Nährstoffbedarfs von Pflanzen genutzt. Die Messungen von satelliten- oder UAS-getragenen Spektralsensoren dienen der Überwachung der Pflanzenentwicklung, der Bestimmung von Anbaukulturen und des Ertragspotenzials. Im Laborbereich werden sie für die Bestimmung des Futterwerts und wertbestimmender Parameter von pflanzlichen Rohstoffen eingesetzt.
Bei allen Anwendungen spielt die Kalibrierung eine zentrale Rolle. Für jede Zielgröße müssen Beziehungen zwischen den Reflektionen in unterschiedlichen Wellenlängenbereichen und der Zielgröße abgleitet werden. Dies erfolgt meist mit Hilfe neuerer mathematischer Methoden wie Künstlichen Neuronalen Netzwerken, Support Vector Machines und Partial Least Squares Regression. Die zentrale Herausforderung ist dabei, dass für eine stabile Kalibrierung eine hohe Anzahl von verlässlichen Referenzwerten erfasst werden muss.

1 Einleitung

Die Reflektion von elektromagnetischer Strahlung steht in Zusammenhang mit den chemischen und physikalischen Eigenschaften der reflektierenden Oberfläche und darunter liegender Schichten. Diese Tatsache machen sich sowohl die Spektroskopie als auch die Augen von Menschen und Tieren zunutze, um Gegenstände zu klassifizieren, also zu bewerten. Dabei wird aus den Reflektionen in unterschiedlichen Wellenlängenbereichen ein Rückschluss auf die Qualität eines Stoffes oder die Quantität einer seiner Komponenten gezogen.
Das menschliche Auge ist hinsichtlich seiner spektralen Bandbreite auf den Bereich des sichtbaren Lichts (400 nm bis 750 nm) und bezüglich der zeitlichen Auflösung auf eine Abtastfrequenz von 60 Hz beschränkt und hat ein Auflösungsvermögen von etwa 60 Bogensekunden. Dies entspricht bei einer Entfernung von einem Meter eine Pixelgröße von 0,3 mm (SPEKTRUM 2017).
Die Signalverarbeitung ist stark vom physischen und psychischen Zustand, der genetischen Disposition sowie der Sozialisation des Betrachters abhängig („Subjektivität“). Somit ist das menschliche Auge als Sensorsystem vor allem aufgrund der geringen spektralen Bandbreite und der geringen Reproduzierbarkeit der Messergebnisse, aber auch wegen der begrenzten Speicherkapazität des menschlichen Gehirns nur eingeschränkt für die Messwerterfassung geeignet.
Spektralsensoren haben im Gegensatz zum menschlichen Auge den Vorteil, dass Reflektionen reproduzierbar und objektiv erfasst werden können. Sie verfügen zudem meist über eine größere spektrale Bandbreite und erfassen dann im länger welligen Bereich auch die Reflektionen im

[1] Hochschule Weihenstephan-Triesdorf, Agrarsystemtechnik, Markgrafenstr. 16, D-91746 Weidenbach, E-Mail: patrick.noack@hswt.de

nahen Infrarot (NIR, 780 nm bis 1400 nm) oder kurzwelligen Infrarot (SWIR, 1400 nm bis 3000 nm). Die Reflektionen in diesem Spektralbereich stehen in engem Zusammenhang zu landwirtschaftlich relevanten Stoffen wie Wasser, Chlorophyll, Carotinoiden, Fetten und Ölen sowie Lignin (Holz, Stroh) und der organischer Substanz („Humus“).
Aus den oben genannten Gründen ist man in unterschiedlichen Bereichen der Tier- und Pflanzenproduktion dazu übergegangen den Menschen bei der Beurteilung von Böden, Pflanzenbeständen, Futter- und Nahrungsmitteln durch den Einsatz von Spektralsensoren zu unterstützen. Die Ziele sind dabei immer die Produktion von Nahrungsmitteln, Futter und Energie hinsichtlich ihrer einzelbetrieblichen Wirtschaftlichkeit, den Anforderungen der abnehmenden Hand (Lebensmittelbranche), den Wünschen der Verbraucher, der Anforderungen der Gesellschaft und nicht zuletzt der Belastung der Umwelt zu optimieren.

2 Spektralsensoren in der Landwirtschaft

Die in der Landwirtschaft eingesetzten Sensoren unterscheiden sich zum Teil erheblich in Größe, Funktionsweise, räumlicher und zeitlicher Auflösung sowie Bandbreite und Auflösung des gemessenen Spektrums.

2.1 Eigenschaften von Spektralsensoren

Hyperspektralsensoren messen ein kontinuierliches Spektrum über den kompletten Messbereich. Die Auflösung der direkt angrenzenden Banden liegt dabei meist im Bereich von wenigen Nanometern oder darunter. Multispektralsensoren messen einzelne Banden mit stark variierender Bandbreite, die nicht direkt angrenzend sind.

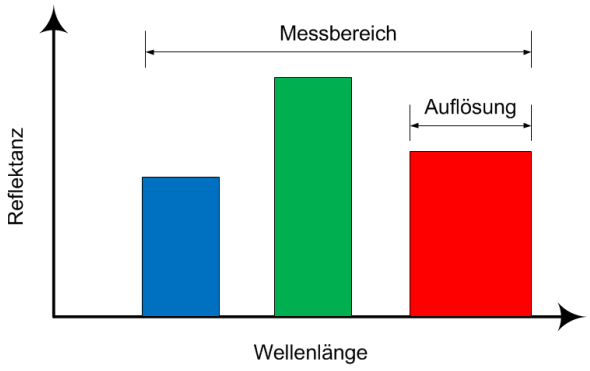


Abb. 1: Multispektralsensor (eigenDarstellung)

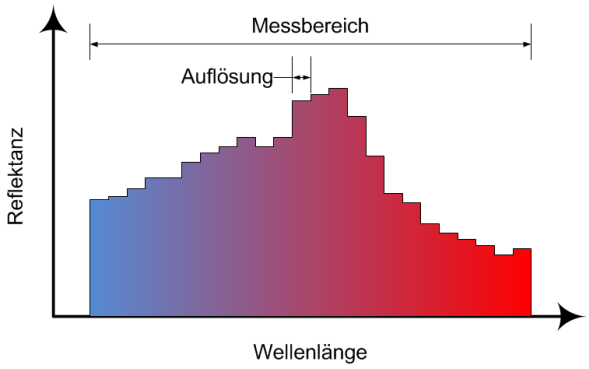


Abb. 2: Hyperspektralsensor (eigenDarstellung)

Spektralsensoren lassen sich zunächst nach der Bildauflösung unterscheiden. Spektrometer sind nicht in der Lage die Rückstrahlung räumlich aufzulösen, fassen also die gesamte Rückstrahlung in einem Messwert zusammen. Die räumliche Abdeckung hängt somit von dem Aufbau der verwendeten Optik ab, die das reflektierte Umgebungslicht bündelt und dem Sensor zuführt.
Bildgebende Verfahren erfassen rasterförmig angeordnete Bereiche, wobei die Bildauflösung erheblich variieren kann (z.B. von 640×480 Pixel bis 6000×4000 Pixel). Bildgebende Spektralsensoren sind bezüglich der Funktionsweise einer Digitalkamera vergleichbar, wobei sie sich bezüglich der erfassten Wellenlängen und der Anzahl der Kanäle unterscheiden. Die räumliche

Auflösung bildgebender Sensoren (in Meter pro Pixel) hängt einerseits von der Bildauflösung und andererseits von der Entfernung zum Untersuchungsobjekt ab. Sensoren an Fahrzeugen oder bodennah eingesetzten Flugobjekten wie UAS liefern auch bei geringer Bildauflösung eine hohe räumliche Auflösung. Satellitenbildaufnahmen weisen trotz hoher Bildauflösung aufgrund der großen Flughöhe eine geringe räumliche Auflösung im Bereich von 1 bis 60 m auf.
Ein weiteres Unterscheidungskriterium ist, ob das Untersuchungsobjekt aktiv beleuchtet wird (aktive Sensoren) oder die Rückstrahlung des Umgebungslichts gemessen wird (passive Sensoren). Aktive Sensoren sind in der Lage aus der eingestrahlten Energie in unterschiedlichen Wellenlängenbereichen und der Reflektion direkt die Reflektanz, also das Verhältnis von Rückstrahlung und Einstrahlung, zu berechnen. Passive Sensorsysteme können lediglich die Reflektion bestimmen oder ermitteln die Einstrahlung mit einem zweiten in Richtung der Lichtquelle ausgerichteten Sensor. Ist keine externe Lichtquelle vorhanden (z.B. bei Nacht), ist die Spektralmessung mit passiven Systemen nicht möglich.
Für den Einsatz in der Landwirtschaft sind je nach Anwendung die spektrale Bandbreite und die spektrale Auflösung von Bedeutung. Entscheidende Parameter von Pflanzenbeständen (Chlorophyllgehalt, Biomasse) lassen sich in der Regel mit drei bis fünf Banden (Multispektral) im sichtbaren Bereich (VIS) und nahinfraroten Bereich (NIR) des Lichts bestimmen. Die Qualität der abgeleiteten Größen ist hier jedoch stark von der spektralen Auflösung des Sensorsystems abhängig, die zwischen 1- 2 nm und mehr als 100 nm variieren kann. Bei der Bestimmung von wertbestimmenden Inhaltsstoffen wie Trockenmassegehalt, Rohprotein oder Rohfaser muss zusätzlich das kurzwellige Infrarot (SWIR) erfasst werden.
Für die Qualität der Messungen ist weiterhin entscheidend, welches Trägersystem bei der Messung eingesetzt wird. In der Landwirtschaft kommen Spektralsensoren handgetragen, an autonomen Kleinfahrzeugen, auf und in landwirtschaftlichen Fahrzeugen wie Traktoren und Erntemaschinen, in landwirtschaftlichen Anbaugeräten (Güllefass) sowie an unbemannten (UAS, Satellit) und bemannten Flugobjekten zum Einsatz.

2.2 Vor- und Nachteile von Sensorsystemen

Die Kosten für Spektralsensoren unterscheiden sich zum Teil erheblich und steigen in der Regel mit der Bildauflösung und der spektralen Auflösung. Die Vor- und Nachteile für unterschiedliche Einsatzzwecke lassen meist lediglich in Zusammenhang mit der Trägerplattform bewerten.
Für die Bestimmung von wertbestimmenden Eigenschaften verschiedener Stoffe (Erntegut, Gülle) werden in der Regel Spektrometer eingesetzt, die einen Messbereich bis in das kurzwellige Infrarot (SWIR, 1400-3000 nm) erfassen. Diese Sensorsysteme sind trotz der geringen Bildauflösung relativ teuer.
Im Bereich der Pflanzenproduktion kann in der Regel auf einen weiten Messbereich verzichtet werden. Die hier eingesetzten Sensoren messen lediglich die Reflektionen ausgewählter Banden im sichtbaren und nahinfraroten Spektrum und sind somit kostengünstiger. Entscheidend ist hierbei, dass die Pflanzenbestände zum richtigen Zeitpunkt und mit ausreichender spektraler und räumlicher Auflösung erfasst werden.
Fahrzeuggetragene Systeme können nur dann eingesetzt werden, wenn die Befahrbarkeit im Feld gegeben ist. Sie wird – meist im Frühjahr und im Herbst - durch die Bodenfeuchte und die hierdurch bedingte Tragfähigkeit des Bodens begrenzt. Zudem ist der Einsatz dieser Sensorsysteme

nur bis zu einer bestimmten Bestandeshöhe möglich. Diese kann, zum Beispiel beim Anbau von Mais, in späten Vegetationsstadien begrenzend wirken. Der Einsatz von fahrzeuggetragenen Sensoren ist mit erheblichem Aufwand und erheblichen Personal- und Maschinenkosten von ca. 100 EUR/h verbunden. Schließlich erfassen die Sensoren bei der Überfahrt lediglich einen Bereich von wenigen Metern neben dem Fahrzeug, messen also die Reflektionen nicht flächendeckend.
An UAS („Drohnen“) montierte Sensoren erzeugen flächendeckende Aufnahmen von Boden und Pflanzenbeständen. Sie können kurzfristig eingesetzt werden und werden kaum durch Bewölkung beeinträchtigt. Allerdings ist die Flächenleistung vor allem bei Koptersystemen aufgrund der geringen Flughöhe und der begrenzten Kapazität der Akkumulatoren stark eingeschränkt (10 bis 100 ha/Tag). Sowohl UAS- als auch fahrzeuggetragenen Systeme können aufgrund der Nähe zum Objekt die Reflektionen mit einer hohen spektralen Auflösung erfassen (1 bis 2 nm) .
Dem hingegen haben Satellitenaufnahmen aufgrund der großen Entfernung meist eine geringere spektrale Auflösung (20 – 200 nm) und sind somit weniger spezifisch mit Eigenschaften der Bestandsoberfläche verknüpft. Ebenso ist die räumliche Auflösung mit meist mehr als 10 m Pixelgröße wesentlich geringer als die von UAS-getragenen Sensoren (0,5 bis 5 cm). Eine Satellitenaufnahme deckt in der Regel ein sehr großes Areal von mehreren hundert Quadratkilometern ab. Die Daten sind allerdings nur dann nutzbar, wenn das Untersuchungsgebiet nicht bewölkt ist. Theoretisch ist die zeitliche Auflösung aktueller Satellitenmissionen (z.B. Sentinel 2) für die Überwachung der Entwicklung von Pflanzenbeständen mit einer Aufnahme in drei Tagen ausreichend. In der Praxis stellt die Bewölkung jedoch oft den begrenzenden Faktor für die Nutzbarkeit dar. Die Aufnahmen der Landsat- und der Sentinel-Mission stehen der Öffentlichkeit kostenlos zur Verfügung. Darüber hinaus vertreiben kommerzielle Anbieter Aufnahmen mit höheren räumlichen und unterschiedlichen spektralen Auflösungen.

2.3 Anwendung von Spektralsensoren in der Landwirtschaft

Spektralsensoren werden in der Landwirtschaft für die Messung von Boden-, Dünger- und Pflanzen- und Ernteguteigenschaften eingesetzt.

2.3.1 Boden

Der Einsatz von Spektrometern in Bodensensoren und Sämaschinen ist relativ neu. Die an Scharaggregaten angebrachten Sensoren gleiten dabei durch den Boden und messen Reflektionen im Bereich des sichtbaren und des nahinfraroten Lichts. Das Ziel ist dabei den Gehalt an organischer Substanz („Humus“) zu bestimmen. Der Gehalt an organischer Substanz korreliert eng mit der Fähigkeit des Bodens Wasser und Nährstoffe zu speichern. Die aus den Spektralmessungen abgeleiteten Humusgehalte werden für die Erstellung von Karten oder die Steuerung der Aussaatdichte genutzt. Untersuchungen in Dänemark (KNADEL et al. 2015) haben gezeigt, dass mit dem Veris® MSP System der Gehalt des Bodens an organischer Substanz bestimmt werden kann. Das Produkt SmartFirmer[2] der amerikanischen Firma Precision Planting funktioniert nach einem vergleichbaren Prinzip und bestimmt zusätzlich die Bodenfeuchte und den Anteil von Ernterückständen im Boden.

[2] http://www.precisionplanting.com/products/smartfirmer/

Tab. 1: Anwendungsbeispiele für Spektralsensoren in der Landwirtschaft

	Anwendung	stationär/ mobil	Träger	Räumliche Auflösung	Spektrale Auflösung
Boden	Organische Substanz	mobil	Bodensensor	-	Multispektral VIS, NIR
	Organische Substanz, Wasser-gehalt	mobil	Satellit	30 bis 60 m	Multispektral NIR, SWIR
Dünger	N-Konzentration Gülle	mobil	Güllefass	-	Hyperspektral VIS, NIR,SWIR
Pflanzenbestand	Chlorophyllgehalt, N-Bedarf Kulturpflanze	mobil	Traktor	6-8 m neben Fahrspur	Multispektral VIS, NIR
	Biomasse, Ausbringmengen Pflanzenschutz	mobil	Traktor	6-8 m neben Fahrspur	Multispektral VIS, NIR
	Biomasse, Chlorophyllgehalt, Kornfeuchte, Ölgehalt	mobil	UAS	0,5 bis 5 cm	Multi- und Hyperspektral VIS, NIR
	Keimung, Wildschäden, Abreife	mobil	UAS	0,5 bis 5 cm	RGB
	Biomasse, Chlorophyllgehalt	mobil	Kleinflugzeug	5 bis 50 cm	Multispektral VIS, NIR
	Biomasse, Chlorophyllgehalt	mobil	Satellit	5 bis 30 m	Multispektral VIS, NIR, (SWIR)
Futtermittel	Trockenmassebestimmung, Rohprotein Futter	stationär	Labor	-	Hyperspektral VIS, NIR, SWIR
	Inhaltsstoffe Körnerfürchte und Leguminosen (Protein, Öl)	stationär	Labor	-	Hyperspektral VIS, NIR, SWIR
	Trockenmassebestimmung Mais/Grashäcksel	mobil	Feldhäcksler	-	Hyperspektral VIS, NIR,SWIR

Der Gehalt an organischer Substanz kann auch aus Fernerkundungsdaten abgeleitet werden GOMEZ et al. 2008 haben dazu Untersuchungen in Australien angestellt. Aus Satellitenaufnahmen können auch Bodenarten bzw. Bodentypen erkannt und kartiert werden. Hierzu liegen unter anderem Ergebnisse von Vibhute vor (VIBHUTE et al. 2015). Oltra Carrió und Kollegen haben Zusammenhänge zwischen Satellitenaufnahmen und der Bodenfeuchte hergestellt (OLTRA-CARRIÓ et al. 2015).

Alle genannten Parameter wie der Gehalt an organischer Substanz, Bodenarten bzw. Bodentypen und die Bodenfeuchte sind für die Pflanzenproduktion entscheidende Faktoren, die kleinräumig stark variieren können. Durch die Wahl standortangepasster Sorten, die teilflächenspezifische Anpassung von Saatdichte, Düngung und Pflanzenschutz können in Abhängigkeit vom Ausmaß der Variabilität der Eigenschaften Mehrerträge erzielt oder Kosten eingespart werden.

2.3.2 Organische Dünger

Bei der tierischen Erzeugung und der Energieerzeugung aus Biogas fallen erhebliche Mengen organischer Dünger in Form von Jauche, Mist und Gülle an. Diese stellen eine wertvolle Ressource dar, da sie mineralische Dünger zumindest teilweise ersetzen können. Organische Dünger tragen im Gegensatz zu mineralischen Düngern auch zum Aufbau der organischen Substanz von Böden bei. Der Nachteil von organischen Düngern liegt in den stark variablen Nährstoffgehalten. Sie werden bei der Ausbringung aktuell lediglich aufgrund der Untersuchung von Stichproben oder Tabellenangaben berücksichtigt.

Untersuchungen in Italien (CABASSI et al. 2015) haben gezeigt, dass mit NIR-Spektrometern der Trockenmasse-, Kohlenstoff- sowie der Gehalt an organischem Stickstoff und seinem Ammoniakanteil in Kälbergüllen bestimmt werden kann. Die Firma Zunhammer bietet seit mehreren Jahren ein System mit der Bezeichnung VAN-Control[3] an, das an Güllefässern montiert fortlaufend den Stickstoffgehalt der Gülle bestimmt und basierend auf den Messwerten die Ausbringmenge regelt. Laut Herstellerangaben kann neben dem Stickstoffgehalt und auch die Konzentration von Kalium und Phosphat mit dem Spektrometer ermittelt werden.

Der Landmaschinenhersteller John Deere nutzt NIR-Spektrometer bereits seit mehreren Jahren um den Trockenmassegehalt des Ernteguts in Feldhäckslern zu bestimmen. Die gleiche Technologie wird nun als Nachrüstlösung für Güllefässer angeboten (JOHN DEERE 2017). Es werden ebenso wie beim VAN-Control System der Trockenmassegehalt sowie die Gehalte an Stickstoff, Kalium und Phosphat ermittelt. Es ist laut Hersteller für Schweinegülle, Rindergülle und Biogasgülle geeignet. Die ausgebrachten Mengen werden auf einem Terminal dargestellt und zum Zwecke der Dokumentation aufgezeichnet.

2.3.3 Düngung und Pflanzenschutz

Fahrzeuggetragene Spektralsensoren werden in der Landwirtschaft oft als N-Sensoren bezeichnet. Dies ist darauf zurückzuführen, dass die Spektralmessungen der Sensoren ursprünglich für die Messung des Stickstoffbedarfs bzw. der Stickstoffaufnahme von Pflanzen entwickelt wurden (REUSCH 1997; SAMBORSKI et al. 2009; ERDLE et al. 2011). Sie werden auf dem Dach der Schlepperkabine oder im Frontanbau montiert und messen dort die Reflektion in ausgewählten Wellenlängenbereichen. Die meisten heute eingesetzten Sensorsysteme erzeugen mit einer Lichtquelle aktiv eine Einstrahlung, so dass sie auch bei Nacht und wechselnden Beleuchtungsbedingungen eingesetzt werden können. Die Reflektionen unterschiedlicher Wellenlängen werden zur Bestimmung der Stickstoffaufnahme zu Indizes verrechnet. Eine Zusammenstellung von in der Landwirtschaft verwendeten Indizes findet sich bei Mulla (MULLA 2013).

Der Stickstoff-Sensor der Firma YARA[4] nutzt den REIP-Index (*Red Edge Inflection Point*)

$$REIP_{[nm]} = R_{700} + 40 * \frac{\frac{R_{670} + R_{780}}{2} - R_{700}}{R_{740} - R_{700}} \quad (1)$$

Dieser Index korreliert eng mit dem Chlorophyllgehalt der Pflanzen und damit mit der Menge Stickstoff, die die Pflanzen bereits aufgenommen haben. Aus diesem sortenspezifischen Wert

[3] http://www.zunhammer.de/de/produkte/elektronik/van-control

[4] http://www.yara.de/pflanzenernaehrung/tools-und-services/n-sensor/

lässt sich die Stickstoffmenge ableiten, die die Pflanzen noch aufnehmen müssen, um den erwarteten Ertrag bilden zu können.
Der GreenSeeker Sensor der Firma Trimble[5] bestimmt den NDVI (*Normalized Difference Vegetation Index*).

(2) $$NDVI_{[nm]} = \frac{R_{NIR} - R_{rot}}{R_{NIR} + R_{rot}}$$

Dieser Index korreliert sowohl mit der Pflanzenmasse als auch mit der Stickstoffaufnahme. Im Gegensatz zum REIP-Index neigt der NDVI dazu bei hohen Chlorophyllgehalten abzusättigen, so dass zwischen Pflanzen mit unterschiedlichen Stickstoffaufnahmen nur noch schlecht unterschieden werden kann. Neben den oben genannten Sensorsystemen wird in der Landwirtschaft ebenfalls der Fritzmeier ISARIA[6] eingesetzt. Er nutzt die beiden Indizes IBI und IRMI um sowohl die Biomasse als auch die Stickstoffaufnahme zu messen.
Die Besonderheit des ISARIA Systems liegt darin, dass es neben den Spektralmessungen des Sensors auch das Ertragspotential bei der Berechnung der Düngermenge berücksichtigt. Das Ertragspotential wird wiederum von Dienstleistern aus mehrjährigen Satellitenaufnahmen und auf Basis von Wetterdaten und Pflanzenwachstumsmodellen abgeleitet.
Die Firmen John Deere, FarmFacts, Vista, Rauch und Sulky haben mit dem als „Connected Nutrient Management“ bezeichneten Konzept verschiedenste Verfahren der Spektralmessung zusammengeführt (VISTA GEO 2015). Basierend auf mehrjährigen satellitenbasierten Multispektralaufnahmen wird das Ertragspotential von Teilflächen abgeschätzt. Während der organischen Düngung wird der Nährstoffgehalt des Düngers fortlaufend mit Spektralsensoren bestimmt und die Ausbringmenge ebenso an den Bedarf angepasst wie bei den nachfolgenden Düngergaben mit mineralischen Stickstoffdüngern.

Abb. 1: ISARIA Stickstoff-Sensor (aus: RECKLEBEN 2014)

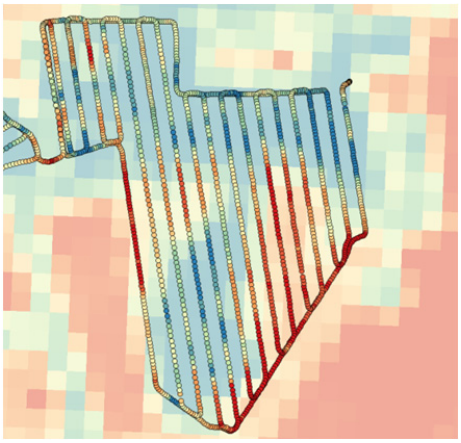


Abb. 2: Messdaten eines Stickstoffsensors (NDVI, Punktdaten) in einem Rapsbestand mit NDVI-Raster aus Landsat 8 Satelliten-aufnahmen (zwei Tage nach Sensorüberfahrt); eigene Darstellung

[5] https://agriculture.trimble.com/precision-ag/products/greenseeker/?lang=de
[6] http://fritzmeier-umwelttechnik.com/isaria/

Grundsätzlich könnten statt der fahrzeuggetragenen Sensoren auch an UAS oder Satelliten montierte Spektralsensoren für die teilflächenspezifische Düngung genutzt werden. Diese Verfahren werden im Gegensatz zu den Online-Sensoren als Offline-Verfahren bezeichnet, da die Daten zunächst erhoben und verarbeitet werden müssen und die Regelung der Stickstoffmenge nicht unmittelbar als Reaktion auf eine Messung erfolgt. Bei den Offline-Verfahren werden in der Regel die gleichen Indizes verwendet wie bei den Online-Verfahren. Der Vorteil der Offline-Verfahren ist, dass die Daten (Landsat-Mission, Sentinel-Mission) kostenlos zur Verfügung stehen und große Gebiete abdecken. Nachteilig wirken sich die geringe zeitliche Verfügbarkeit, der Zeitbedarf für die Verarbeitung und Übertragung von Sollwertkarten auf das Fahrzeug und die teils geringe Auflösung aus. Die auf Basis von Satellitendaten und mit Stickstoff-Sensoren ermittelten Unterschiede im Bestand sind in jedem Fall vergleichbar (s. Abb. 2).
Die Stickstoffdüngung erfolgt über den Verlauf der Vegetation in verschiedenen Gaben. Die Anzahl der Gaben variiert dabei je nach Fruchtart und Bestandesführung zwischen zwei und vier. Die Gesamtmenge orientiert sich dabei am Zielertrag bzw. dem auf einem Feld oder einer Teilfläche zu erwartenden Ertrag. Dieser hängt einerseits von verschiedenen Bodeneigenschaften wie dem Nährstoffgehalt und der Wasserhaltefähigkeit ab, die räumliche Unterschiede aufweisen. Sie bestimmen das (theoretische) Ertragspotential an einem Standort. Die Ertragserwartung ist das Produkt aus dem Ertragspotential und den jährlich variierenden Wetterbedingungen, also dem Verlauf von Niederschlägen, Temperaturen und anderer Parameter. Für die Bestimmung des Ertragspotentials wird aus den spektralen Informationen von Satellitenaufnahmen mehrerer Jahre die Biomasseentwicklung beurteilt. Diese Informationen können zusätzlich mit Wetterdaten und Pflanzenwachstumsmodellen kombiniert werden. Das Produkt schätzt den Gesamtstickstoffbedarf eines Bestandes teilflächenspezifisch ab und kann entweder direkt für die Ableitung von Sollwertkarten für die standortangepasste Düngung genutzt oder als begrenzende Regelgröße bei der sensorgestützten Düngung eingesetzt werden. Ertragspotentialkarten werden von verschiedenen Firmen als Produkt angeboten (Agrosat[7], FarmFacts[8], Geosys[9], GREENSPIN[10], Kleffmann[11], u.a.)
Neben der Ausbringung von Stickstoff können auch andere Betriebsmittel auf Basis von Spektralmessungen in ihrer Aufwandmenge variiert werden. So kann die Ausbringung von systemisch wirkenden Fungiziden an die Bestandsentwicklung angepasst werden. Als Maß für die Bestandsentwicklung gilt hierbei meist die Biomasse, die sich aus dem NDVI oder anderen Spektral-Indices ableiten lässt. Die Anpassung der Aufwandmenge stellt einerseits sicher, dass der gesamte Bestand mit einer für die Bekämpfung von Pilzkrankheiten ausreichenden Wirkstoffkonzentration behandelt wird und schränkt andererseits die Ausbringung auf ein ökonomisches Maß ein und spart so Produktionskosten. Untersuchungen dazu wurden unter anderem in Kanada durchgeführt (PICARD et al. 2011).

[7] http://www.agrosat.de/
[8] https://www.farmfacts.de/
[9] http://www.geosys.com/
[10] https://www.greenspin.de/
[11] www.mydataplant.com

MAHLEIN (2016) betrachtet unterschiedliche Verfahren zur Erkennung von Pflanzenkrankheiten. Neben anderen Verfahren wie der Thermographie und der Chlorophyllfluoreszenzmessung stellen sich dabei auch spektroskopische Verfahren als geeignet heraus. Die Bedingung für den Einsatz von Spektralsensoren bei der Erkennung von Pflanzenkrankheiten ist dabei allerdings eine sehr hohe räumliche Auflösung im Bereich von wenigen Millimetern.

2.3.4 Futtermittel

NIR-Spektrometer werden seit mehr als fünfzehn Jahren in Feldhäckslern für die Bestimmung der Trockenmasse von Gras, Mais und anderen Grünfuttermitteln eingesetzt (KORMANN 2001). Diese Technologie stellt bei fast allen Herstellern von Feldhäckslern den Stand der Technik dar. Ein Spektrometer im Auswurfkrümmer misst fortlaufend die Reflektion des Ernteguts und leitet daraus den Feuchtegehalt ab. Dies dient einerseits der Dokumentation. Andererseits werden die Messwerte genutzt, um die Schnittlänge des Ernteguts der Feuchte anzupassen und die Dosierung von Hilfsmitteln für die Konservierung zu steuern. Zusätzlich können teilweise wertbestimmende Parameter wie der Gehalt des Ernteguts an Rohprotein, Rohfaser und Zucker bestimmt werden (HEINRICH 2005, Herstellerangaben).
Die NIR-Spektroskopie kann auch bei der Getreideernte für die Messung der Feuchte und ausgewählter Qualitätsparameter (Proteingehalt) genutzt werden (RECKLEBEN 2004, REYNS 2002). Diese Technologie hat sich jedoch bisher vornehmlich im Versuchswesen durchgesetzt.
Die NIRS-Technologie wird auch im Laborbereich für die Untersuchung von Futtermitteln und Erntegütern eingesetzt[12]. Dort werden Trockenmasse, Rohproteingehalt, Rohfaser- und Rohfettgehalt, Zuckergehalt und andere wertbestimmende Parameter von Gras- und Mais-Silagen mit Hilfe von Spektrometern bestimmt. Auch bei unterschiedlichen Getreidearten wie Weizen, Gerste, Triticale sowie Ölsaaten (Raps, Sonnenblumen, Leinsaat) können mit Hilfe der Spektroskopie Protein-, Öl- und Trockenmassegehalte ermittelt werden.

3 Fazit & Ausblick

Die Spektroskopie wird in der Landwirtschaft in verschiedensten Bereichen für die Bestimmung von physikalischen und chemischen Stoffeigenschaften genutzt.
Sowohl Böden als auch Dünger, Pflanzenbestände und aus Pflanzen gewonnene Produkte können mit Hilfe der Spektroskopie auf ihren Gehalt an wertbestimmenden Inhaltsstoffen oder produktionsrelevanten Parametern (Stickstoff, Pathogene) untersucht werden. Die Methode ist berührungslos, zerstörungsfrei und ermöglicht die Ermittlung von Messwerten in Echtzeit und mit hoher zeitlicher Frequenz.
Gleichzeitig steht die Genauigkeit und die Zuverlässigkeit der Ableitung von Stoffeigenschaften aus spektralen Signaturen in engem Zusammenhang mit den für die Ableitung der Zusammenhänge verwendeten Methoden und der Anzahl der für die Modellbildung verwendeten Referenzproben. Nachdem bei der Erstellung von Rechenmodellen, die die Zusammenhänge zwischen den Reflektionen und den Stoffeigenschaften beschreiben, mehrere Spektralkanälen betrachtet

[12] http://www.vdlufa-nirs.de

werden, sind einfach zu handhabende mathematische Methoden wie die eindimensionale Regressionsanalyse für die Erstellung von Modellen nicht geeignet. Aus diesem Grund werden für die Modellbildung fortgeschrittene Ansätze wie *partial least squares regression*[13], künstliche neuronale Netzwerke (artifcial neuronal networks, AGATONOVIC-KUSTRIN& BERESFORD 2000) und andere Ansätze (z.B. Support Vector Machine bei DEVOS et al. 2009) für das maschinelle Lernen verfolgt.

Diese neuen Ansätze für die Ableitung von Informationen werden von einigen Wissenschaftlern (MAINZER 2014) kritisiert, da sie das etablierte Vorgehen, das aus einem Prozess- oder Objektverständnis mathematische Beschreibungen ableitet, ersetzt und sich ausschließlich auf die Herstellung eines empirischen Zusammenhangs zwischen Daten und Phänomenen zurückzieht ohne die kausalen Ursachen zu kennen und zu berücksichtigen.

Ungeachtet dessen bieten die genannten Methoden neue Chancen effizient und kostengünstig Stoffeigenschaften zu messen. Dabei gilt, dass insbesondere im landwirtschaftlichen Umfeld die Anzahl der Referenzmessungen einen ausreichend hohen Umfang haben muss und dabei das komplette Spektrum möglicher Variationen abdecken sollte. Sowohl Böden als auch Pflanzen und Pflanzenprodukte weisen aufgrund ihres natürlichen Ursprungs räumlich, zeitlich sowie arten- und sortenbedingt eine große Variabilität auf.

4 Literaturverzeichnis

AGATONOVIC-KUSTRIN, S. & BERESFORD, R., 2000: Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of Pharmaceutical and Biomedical Analysis, **22**(5), 717-727.

CABASSI, G., CAVALLI, D., FUCCELLA, R. & GALLINA, P.M., 2015: Evaluation of four NIR spectrometers in the analysis of cattle slurry. Biosystems Engineering, **133**, 1-13, https://doi.org/10.1016/j.biosystemseng.2015.02.011.

DEVOS, O., RUCKEBUSCH, C, DURAND, A., DUPONCHEL, L. & HUVENNE, J.-P., 2009: Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. Chemometrics and Intelligent Laboratory Systems, **96**(1), 27-33.

ERDLE, K., MISTELE, B. & SCHMIDHALTER, U., 2011: Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. Field Crops Research, **124**(1), 74-84, https://doi.org/10.1016/j.fcr.2011.06.007.

GOMEZ, C., VISCARRA ROSSEL, R.A. & MCBRATNEY, A.B., 2008: Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. Geoderma, **146**(3-4), 403-411, https://doi.org/10.1016/j.geoderma.2008.06.011.

HEINRICH, A., GÜNTHER, A. & BERNHARDT, G., 2005: Bestimmung der Inhaltsstoffe bei der Ernte mit dem Feldhäcksler. Landtechnik, **60**(1), 20-21.

JOHN DEERE, 2017: John Deere Manure Sensing. https://www.deere.com/en_INT/products/equipment/agricultural_management_solutions/precision_farming_solutions/manure_sensing/manure_sensing.page

[13] https://cran.r-project.org/web/packages/pls/vignettes/pls-manual.pdf

KNADEL, M., THOMSEN, A., SCHELDE, K. & GREVE, M., 2015: Soil organic carbon and particle sizes mapping using vis - NIR, EC and temperature mobile sensor platform. Computers and Electronics in Agriculture, **114**, 10.1016/j.compag.2015.03.013.

KORMANN, G., 2001: Untersuchungen zur Integration kontinuierlich arbeitender Feuchtemesssysteme in ausgewählte Futtererntemaschine. Dissertation, Technische Universität München, Fakultät Wissenschaftszentrum Weihenstephan, https://mediatum.ub.tum.de/doc/820167/file.pdf.

MAHLEIN, A.-K., 2016: Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. Plant Disease, **100**(2), 241-251, https://doi.org/10.1094/PDIS-03-15-0340-FE.

MAINZER, K., 2014: Die Berechnung der Welt: Von der Weltformel zu Big Data. C.H.Beck Verlag.

MULLA, D.J., 2013: Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. Biosystems Engineering, **114**(4), 358-371.

OLTRA-CARRIÓ, R., BAUP, F., FABRE, S., FIEUZAL, R. & BRIOTTET, X., 2015: Improvement of Soil Moisture Retrieval from Hyperspectral VNIR-SWIR Data Using Clay Content Information. Remote Sens., **2015**(7), 3184-3205, doi:10.3390/rs70303184.

PICARD, R., MARTENS, G., RABE, N., BISHT, V. & DOBSON, C., 2011: Using Aerial Imagery for Site Specific Fungicide Application. https://doi.org/10.13140/rg.2.2.18921.90727, https://umanitoba.ca/faculties/afs/agronomists_conf/media/Rejean_Picard_fungicide_app_poster.pdf

RECKLEBEN, Y., 2004: Innovative Echtzeitsensorik zur Bestimmung und Regelung der Produktqualität von Getreide während des Mähdruschs. Dissertation, VDI-MEG Nr. 424, Kiel.

RECKLEBEN, Y., 2014: Sensoren für die Stickstoffdüngung – Erfahrungen in 12 Jahren praktischem Einsatz. Journal für Kulturpflanzen, **66**(2), 42-47, DOI: 10.5073/JFK.2014.02.02

REUSCH, S., 1997: Entwicklung eines reflexionsoptischen Sensors zur Erfassung der Stickstoffversorgung landwirtschaftlicher Kulturpflanzen. Dissertation, VDI-MEG Nr. 303, Kiel.

REYNS, P., 2002: Continuous measurement of grain and forage quality during harvest. Diss., Katholieke Universiteit Leuven, Faculteit Landbouwkundige en Toegepaste, Biologsche Wetenschappen

SAMBORSKI,S.M., TREMBLAY, N. & FALLON, E., 2009: Strategies to Make Use of Plant Sensors-Based Diagnostic Information for Nitrogen Recommendations. Agronomy Journal, **101**(4), 800-816.

SPEKTRUM, 2017: Lexikon der Neurowissenschaft, Auflösungsvermögen, http://www.spektrum.de/lexikon/neurowissenschaft/aufloesungsvermoegen/1071, letzter Zugriff 22.1.2017.

VIBHUTE, A., KALE, K., DHUMAL, R. & MEHROTRA, S., 2015: Soil type classification and mapping using hyperspectral remote sensing data. IEEE International Conference on Man and Machine Interfacing (MAMI), 1-4, 10.1109/MAMI.2015.7456607.

VISTA GEO, 2015: Connected Nutrient Management. http://www.vista-geo.de/wp-content/uploads/NutrientManagment_Agritechnica_2015_EN.pdf

Flugroboter für Precision Farming und das Feldversuchswesen

GÖRRES GRENZDÖRFFER[1]

Zusammenfassung: Drohnen bzw. autonome Flugroboter sind für landwirtschaftliche Anwendungen prädestiniert, da sie flexibel einsetzbar sind und vor allem passgenau aktuelle Informationen über den Zustand der landwirtschaftlichen Flächen oder Feldversuche liefern können. Damit aus den Bildern relevante Informationen – möglichst automatisch - abgeleitet werden können, sind je nach Kulturart und Fragestellung (z.B. Düngung, Unkrauterkennung, Pflanzenschutz) Daten mit unterschiedlicher räumlicher, spektraler und zeitlicher Auflösung notwendig. Der Beitrag gibt einen Überblick und zeigt dies anhand verschiedener Beispiele.

1 Einleitung

Dass das Thema Drohnen / Flugroboter (engl. Unmanned Airborne Systems UAS) viel Aufmerksamkeit auf Fachtagungen, Feldtagen und Messen erzeugt, liegt neben Faszination für das Fliegen und dem Spieltrieb, natürlich auch an den vielfältigen zivilen Nutzungsmöglichkeiten von Flugrobotern. Laut einer im Auftrag der AUVSI (Association for Unmanned Vehicle Systems International) entstandenen Studie könnten durch die Zulassung von privaten Drohnen im US-Luftraum alleine in den USA innerhalb von drei Jahren 70.000 neue Arbeitsplätze entstehen (AUVSI, 2013). Mit die meisten Jobs sollen dabei in der Landwirtschaft entstehen. Wie könnte der Einsatz von Flugrobotern in Zukunft für Precision Farming aussehen? Vielleicht hilft ein Blick in die wissenschaftliche Glaskugel:
Ein Landwirt wird über seine Software informiert, dass die Witterungsbedingungen für einen UAS-Flug in den kommenden Stunden gegeben sind. Daraufhin gibt er das Startzeichen für seinen Flugroboter, der wie möglichst jeden Tag, autonom seine vorher programmierte Route abfliegt. Er landet autonom und die Bilddaten werden automatisch georeferenziert und verschiedene Produkte abgeleitet: z.B. die Wuchshöhe, Informationen zur Pflanzenvitalität, der Biomasse und andere biophysikalische und biochemische Parameter. Die qualitätsgeprüften Daten werden anschließend mit den bisherigen Daten verglichen und über Differenzen werden die Veränderungen im Allgemeinen und das Pflanzenwachstum im Speziellen erfasst. Der Verlauf der Veränderungen wird automatisiert in ein dynamisches Pflanzenwachstumsmodell integriert. Dieses aktualisiert auf der Grundlage der tagesaktuellen Fernerkundungsdaten und anderer agrarmeteorologischer Informationen und bisherigen Pflanzenbonituren die Vorhersagen über den weiteren Wachstumsverlauf. Auf dieser Grundlage können dann z.B. Düngeentscheidungen oder der Einsatz von Pflanzenschutzmitteln (PSM) getroffen werden. Darüber hinaus erlaubt das kontinuierliche Monitoring auch die Beobachtung von Schädigungen bzw. Wachstumsdepressionen aller Art in einem frühen Stadium.

[1] Universität Rostock, Professur für Geodäsie und Geoinformatik, J.-v.-Liebig Weg 6, D-18059 Rostock, E-Mail: goerres.grenzdoerffer@uni-rostock.de

Damit dieses Szenario in ein paar Jahren Wirklichkeit werden kann, ist auf alle Fälle einiges an Forschungsaufwand zu betreiben, denn insbesondere die Extraktion von pflanzenbaulich relevanten Daten, die Automation der Datenauswertung und eine standortangepasste Kalibrierung der verschiedenen Sensoren und Daten ist nicht trivial. Bei der Modellierung des Pflanzenwachstums unter Berücksichtigung verschiedener Skalen ist ebenfalls noch viel Forschungsbedarf. Auch die Zuverlässigkeit und der vollautomatische Betrieb eines UAS ist noch verbesserungswürdig. Zudem gibt es zurzeit noch eine Fülle von rechtlichen Rahmenbedingungen / Beschränkungen bei einem vollautomatischen Betrieb eines Flugroboters.
Unbemannte Flugsysteme (UAS) sind also für Precision Farming im Allgemeinen und das Feldversuchswesen im Speziellen in vielerlei Weise ein prädestiniertes Anwendungsfeld, da viele pflanzenbaulich relevante Informationen zu definierten Zeitpunkten, schnell, berührungslos, objektiv, und automatisiert erfasst werden können. Auch das Thema Phenotyping ist im Kontext von großem Interesse, da in Abhängigkeit der Kultur eine Identifikation der Einzelpflanze möglich ist, und deren Wachstumsverlauf individuell erfasst, verglichen und bewertet werden kann.

2 Auf die Auflösung kommt es an!

Der Begriff der Auflösung ist in der Fernerkundung ein sehr zentraler, denn er bestimmt was aus den Luftbildinformationen extrahiert werden kann. Dabei unterscheidet man zwischen räumlicher, radiometrischer, spektraler und temporaler Auflösung.
Die nahezu beliebige Wahl der räumlichen Auflösung, bzw. der Bodenauflösung beim Einsatz von Flugrobotern vervielfachen die Anwendungsmöglichkeiten, denn die freie Skalierbarkeit ermöglicht sowohl den selektiven Blick auf die einzelne Pflanze als auch den großen Überblick des Pflanzenbestandes aus der Vogelperspektive.
In der Praxis gibt es zwei Möglichkeiten die Bodenauflösung zu variieren. Zum einen über die Flughöhe, d.h. bei größerer Flughöhe wird die Bodenauflösung geringer. Oder zum anderen über die Brennweite, d.h. bei gleicher Flughöhe kann mit einer längeren Brennweite (Zoom) eine höhere Bodenauflösung erzielt werden. Mit der Bodenauflösung eng verbunden ist die räumliche Skale, bzw. Betrachtungsebene, die von der Blattebene über die Einzelpflanzenebene bis zur Parzellen- und Bestandsebene reicht. Die Bodenauflösung ist deshalb in vielerlei Hinsicht von großer Bedeutung und hat auch diverse Konsequenzen zur Folge:

1. Die Bodenauflösung bestimmt über die Erkennbarkeit bzw. Trennbarkeit einzelner Objekte und somit auch über die Skalenebene. Bei höchster Bodenauflösung im mm-Bereich können einzelne Blätter, bzw. deren Aufbau und Struktur identifiziert werden. Diese hohen Bodenauflösungen werden in der Regel im Labor realisiert. Bei Bodenauflösungen von 0.5 - 4 cm können, je nach Pflanzenart und Entwicklungsstadium, größere Blätter und einzelne Pflanzen eindeutig separiert werden. Bei Bodenauflösungen von > 4 cm beginnt der Übergang zur Parzellen- bzw. Bestandsebene, d.h. einzelne Pflanzen können nicht mehr separiert werden und der Sensor zeichnet ein Mischsignal auf, dass sich aus mehreren Komponenten (Pflanze, Schatten, Boden) zusammensetzt, wie wir es aus der Flugzeug oder Satellitenfernerkundung kennen.

2. Der Befliegungs- und Speicheraufwand steigt exponentiell mit der Bodenauflösung. Eine Halbierung der Bodenauflösung bedeutet gleichzeitig eine Vervierfachung der Datenmenge und der Flugzeit.
3. Mit den Skalenübergängen sind verschiedene Probleme verbunden. Das betrifft insbesondere die Tatsache, dass sich die spektralen Reflexionseigenschaften stark verändern, je nachdem ob diese an einem einzelnen Blatt(teil) (Leaf Level) gemessen werden oder einen Mittelwert aus mehreren Komponenten, wie etwa Boden, Pflanze und Schattenanteilen (canopy level) darstellen. Da die Sonne als Lichtquelle fungiert, ändern sich die Einstrahlungsbedingungen in Abhängigkeit des Sonnenstandes, der Lichtstärke (Wolkenbedeckung, Jahreszeit) und der Pflanzenarchitektur. Darüber hinaus ändern sich die Reflexionseigenschaften auch kontinuierlich während des Pflanzenwachstums und durch Stress. Um es an dieser Stelle abzukürzen, eine quantitative Ableitung von biophysikalischen Parametern in einer Skale und über Skalengrenzen hinweg ist sehr komplex und Gegenstand vieler Forschungsaktivitäten. Allerdings wird es noch einige Zeit dauern, bis ausgereifte und automatisierte Verfahren zur Verfügung stehen.

Der besondere und einzigartige Vorteil UAS-gestützter Fernerkundung ist die hohe zeitliche (temporale) Auflösung mit der die hochaufgelösten fernerkundlichen Daten erhoben werden können. Da unterhalb der Wolken geflogen werden kann, ist es möglich nahezu an jedem Tag zu fliegen. Man ist also nicht mehr unbedingt auf wolkenlosen Himmel angewiesen, sondern kann auch bei bedecktem Himmel fliegen, was insbesondere die spektrale Kalibrierung stark vereinfachen kann, da die bidirektionalen Reflektionseigenschaften (BRDF) des direkt einfallenden Sonnenlichts nicht berücksichtigt und modelliert werden müssen.

Tab. 1: Vor- und Nachteile von UAS-Flügen bei geschlossener Bewölkung vs. Online-Kalibrierung

Methode	Vorteile	Nachteile
Parallele Messung der einkommenden und reflektierten Strahlung	Flug bei sonnigem oder bewölktem Wetter (mehr Flugtage)	BRDF ist abhängig von Exposition der (Pflanzen)teile zur Sonne. Bei hoher Auflösung Überstrahlung einzelner Pflanzenteile
	Direkte Messung der Reflektanzen	
	Einfacher temporaler Vergleich und multisensoraler Vergleich	
		Kostspieliger Spezialsensor
Flug bei bedecktem Himmel	Sehr einfache Lösung	Ggf. zu wenig Licht
	keine zusätzliche Hardware notwendig	Geringere Anzahl von Flugtagen
	Kein Schatten	Keine absoluten Reflektanzen

Zudem kann man vor Ort flexibel auf wechselnde Witterungsbedingungen reagieren und - wenn möglich - die sonnigen oder voll bedeckten Minuten, bzw. Stunden voll ausnützen. Das setzt allerdings voraus, dass die rasch wechselnden Strahlungsveränderungen zeitgleich erfasst werden. Geeignete Sensoren, wie der Sequoia Sensor von Parrot oder MicaSense sind mittlerweile am Markt verfügbar.
Die damit mögliche kontinuierliche Erfassung des Aufwuchses eröffnet viele neue Anwendungsmöglichkeiten, um z.B. den Verlauf der Pflanzenentwicklung und somit auch Prozesse, wie die Düngeaufnahme und den Verlauf von Pflanzenkrankheiten zu erfassen und zu dokumentieren. Das setzt allerdings voraus, dass die Datenprozessierung und -auswertung mehr oder weniger automatisch geschieht.
Auf der Grundlage eines systematischen Bildflugs, entsprechende Passpunkte bzw. entsprechender GNSS-Technik an Bord des UAS und eine solide photogrammetrische Datenauswertung vorausgesetzt, liefern UAS-Aufnahmen nicht nur zweidimensionale Bilder, sondern dreidimensionale Geodaten, die sich entsprechend (automatisch) auswerten und interpretieren lassen. Zusammengefasst lassen sich aus den UAS-Aufnahmen viele Informationen ableiten, bei denen die geometrischen, spektralen oder temporalen Eigenschaften der Pflanzenentwicklung im Vordergrund stehen. Geometrische Parameter, wie die Wuchshöhe oder auch der Bodenbedeckungsgrad lassen sich mit preisgünstigen Sensoren (Farbdigitalkamera) weitestgehend automatisch berechnen, und sollen deshalb im nachfolgenden genauer vorgestellt werden.

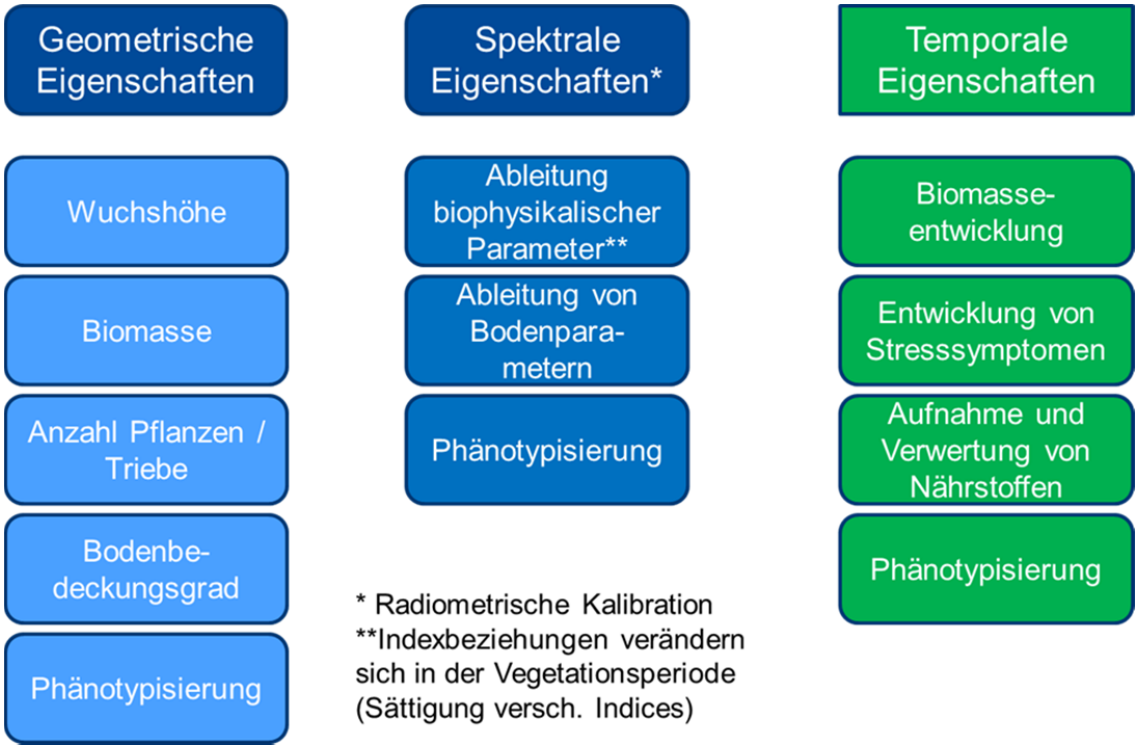


Abb. 1: Übersicht verschiedener Informationen bzw. Pflanzeneigenschaften für Precision Farming und das Feldversuchswesen, die aus UAS-Aufnahmen ableitbar sind

Abbildung 1 gibt einen groben Überblick über die aus UAS-Aufnahmen ableitbaren Informationen in Abhängigkeit ihrer geometrischen, spektralen und temporalen Auflösung. An dieser Stelle sei noch angemerkt, dass manche der in Abbildung 1 aufgeführten Parameter nur zu bestimmten phänologischen Entwicklungsstadien zu beobachten sind.

3 Fazit & Ausblick

In Abhängigkeit der gewählten Bodenauflösung (Bestandsebene oder Einzelpflanzenebene) und der betrachteten Kultur lassen sich verschiedene geometrischen Informationen, wie z.B. die Wuchshöhe (GRENZDÖRFFER 2014), der Bedeckungsgrad (TORRES-SANCHEZ et al. 2014), individuelle Pflanzen und/oder Unkräuter mit einem hohen Automationsgrad bestimmen. Während der Fokus dieses Beitrags auf der Ableitung geometrischer Parameter liegt, ist es natürlich auch möglich die Spektralwerte einer RGB-Kamera zu nutzen, um diese beispielsweise mit biophysikalischen Parametern zu korrelieren, z.B. RASMUSSEN et al. (2016) oder SCHIRRMANN et al. (2016).
Gleichwohl ist es nicht so, dass man einfach mit einem UAS und Kamera über ein Versuchsfeld fliegt und die gewünschten Daten erscheinen automatisch und wohl strukturiert auf dem heimischen Computer. Vielmehr sind an die jeweilige Kultur und Fragestellung angepasste Befliegungs- und Auswertestrategien zu entwickeln. Nachfolgendes Beispiel zeigt dies für Blumenkohlpflanzen, bei denen der Kopfdurchmesser, die Wölbung des Kopfes sowie die Wuchshöhe automatisch erfasst werden kann.

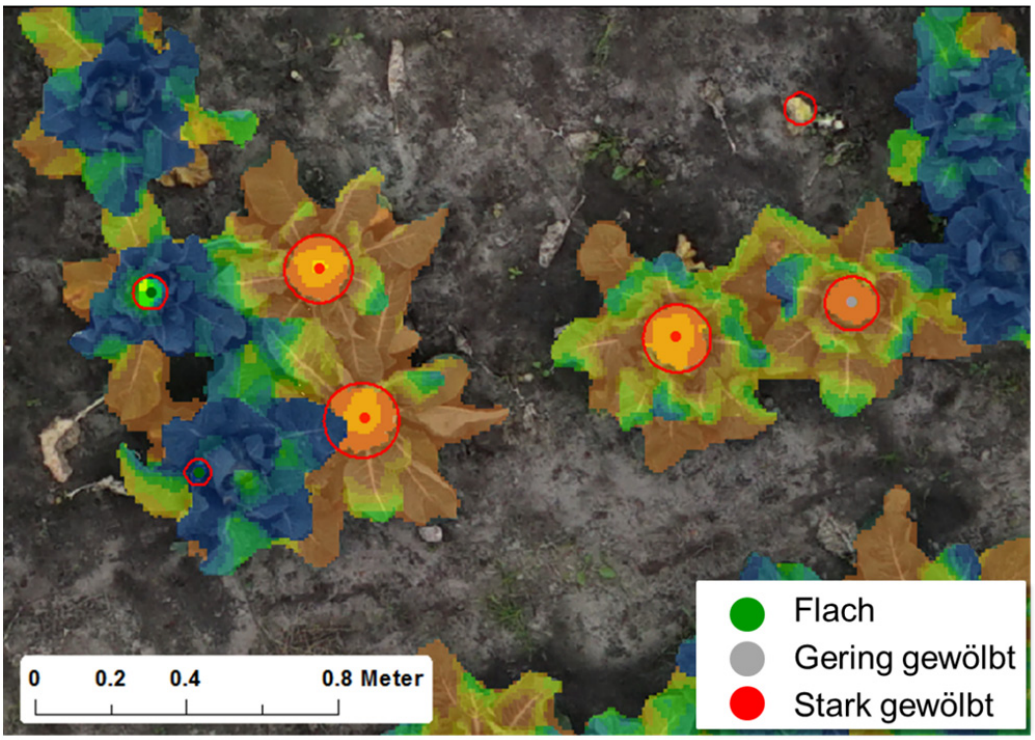


Abb. 2: Automatisch berechnete geometrische Parameter aus einer UAS-Befliegung von Blumenkohlköpfen - Kopfdurchmesser, Kopfwölbung und Wuchshöhe

4 Literaturverzeichnis

AUVSI, 2013: The economic impact of Unmanned Aircraft Systems integration in the United States. 40. (https://higherlogicdownload.s3.amazonaws.com/AUVSI/958c920a-7f9b-4ad2-9807-f9a4e95d1ef1/UploadedImages/New_Economic%20Report%202013%20Full.pdf)

GRENZDÖRFFER, G., 2014: Crop Height Determination with UAS Point Clouds. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **40**(1), 135-140, doi:10.5194/isprsarchives-XL-1-135-2014.

RASMUSSEN, J., NTAKOS, G., NIELSEN, J., SVENSGAARD, J., POULSEN, R.N. & CHRISTENSEN, S., 2016: Are vegetation indices derived from consumergrade cameras mounted on UAVs sufficiently reliable for assessing experimental plots? Eur. J. Agron. 2016, 74, 75–92.

SCHIRRMANN, M., GIEBEL, A., GLEINIGER, F., PFLANZ, M., LENTSCHKE, J. & DAMMER, K.-H., 2016: Monitoring Agronomic Parameters of Winter Wheat Crops with Low-Cost UAV Imagery. Remote Sens. 2016, **8**, 706; doi:10.3390/rs8090706.

TORRES-SÁNCHEZ, J. PEÑA, J.M., DE CASTRO, A.I., LÓPEZ-GRANADOS, F., 2014: Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. Computers and Electronics in Agriculture, **103**, 104-113.

Neue Wege für die Zukunft? Erfahrungen aus dem Einsatz neuer Planungsgrundlagen für die Konzeption ländlicher Kernwegesysteme

THOMAS MACHL[1] & WOLFGANG EWALD[2]

Zusammenfassung: Ein funktionierendes, bedarfsgerecht sowie multifunktional ausgestaltetes Verkehrsnetz bildet eine wesentliche Voraussetzung für die nachhaltige Sicherung der Wettbewerbsfähigkeit land- und forstwirtschaftlicher Betriebe sowie für eine Aufrechterhaltung vitaler ländlicher Räume. Bedingt durch strukturelle Veränderungen in der Landwirtschaft haben sich jedoch die Anforderungen an die landwirtschaftliche Infrastruktur deutlich gewandelt, so dass ländliche Wege vielfach nicht mehr in der Lage sind, den Bedürfnissen der modernen Landwirtschaft und integrierten Gemeindeentwicklung gerecht zu werden.
Ein Modul des Forschungsprojekts LandModell^{TUM} beschäftigt sich vor diesem Hintergrund mit der Entwicklung von Konzepten und Werkzeugen für ein landesweites Monitoring landwirtschaftlicher Hof-Feld-Transportbeziehungen. Die entwickelten Methoden liefern erstmalig flächendeckend detaillierte Aufschlüsse zu landwirtschaftlichen Hof-Feld-Transportfahrten. Erstmals wurden die Ergebnisse der Analysen als Grundlage zur Entscheidungsunterstützung im Zusammenhang mit der Konzeption eines interkommunalen Kernwegekonzepts eingesetzt. Dieser Beitrag behandelt die Erfahrungen aus dem Einsatz dieser neuartigen Planungsgrundlagen.

1 Hintergrund

Strukturelle Veränderungsprozesse in der Landwirtschaft haben die Anforderungen an das landwirtschaftliche Wegesystem verändert (stellvertretend für viele DWA, 2016). Landwirtschaftliche Parzellen werden zunehmend gemarkungs- bzw. gemeindeübergreifend (vgl. SOBOTH 2012; GOCKEL 2012) sowie verstärkt auch überbetrieblich (vgl. DWA, 2016) bewirtschaftet. Parzellen liegen vielfach verstreut und Hof-Feld-Transportdistanzen nehmen zu (SOBOTH 2012). Im Zuge einer zunehmenden Rationalisierung der Bewirtschaftung landwirtschaftlicher Parzellen werden die eingesetzten Maschinen mit wachsender Schlagkraft und Leistungsfähigkeit vielfach schwerer, größer und insbesondere im Straßentransport auch schneller (MACHL et al. 2016). Das landwirtschaftliche bzw. ländliche Wegenetz ist vielerorts nicht mehr in der Lage, den veränderten Ansprüchen der modernen Landwirtschaft gerecht zu werden (SOBOTH 2012).
Die bedarfsgerechte Umgestaltung bzw. der Aus- und Aufbau eines interkommunalen Kernwegesystems ist mit hohen Kosten verbunden. Es sind daher sowohl Schwerpunkte zu setzen als auch zeitliche Prioritäten festzulegen (BROMMA 2014). Neben dem baulichen Zustand der Wege, der Wegefunktion und verschiedenen naturräumlichen Aspekten ist insbesondere die Überol-

[1] Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstraße 21, D-80333 München, E-Mail: thomas.machl@tum.de
[2] Bayerisches Staatsministerium für Landwirtschaft, Ernährung und Forsten, Referat E1 „Integrierte Ländliche Entwicklung, Flurneuordnung“, Ludwigstraße 2, D-80539 München, E-Mail: wolfgang-guenther.ewald@stmelf.bayern.de

lungshäufigkeit bzw. die erschlossene Fläche durch einzelne Wegsegmente wesentliche Entscheidungsgrundlagen für die Priorisierung des Ausbaus. Die Erhebung solcher Parameter ist aufgrund unzureichender Datenverfügbarkeit mit enormem Aufwand verbunden (BERTLING et al. 2015).

2 Neue Ansätze zur Ableitung der Erschließungsfläche aus Hof-Feld-Transportpfaden

Vor diesem Hintergrund beschäftigte sich ein Modul des Forschungsprojekts LandModell^{TUM}, welches derzeit im Auftrag der Bayerischen Verwaltung für Ländliche Entwicklung am Lehrstuhl für Geoinformatik der Technischen Universität München (TUM) bearbeitet wird, mit der Entwicklung von Konzepten und Werkzeugen für ein landesweit flächendeckendes Monitoring landwirtschaftlicher Hof-Feld-Transportpfade.
Die Analysen stützen sich im Wesentlichen auf anonymisierte Daten des Integrierten Verwaltungs- und Kontrollsystems (InVeKoS) sowie auf Daten des Amtlichen Topographisch-Kartographischen Informationssystems (ATKIS). Während die Daten des InVeKoS Aufschluss zur Lage landwirtschaftlicher Parzellen, den angebauten Kulturen sowie zur Lage der jeweils bewirtschafteten Betriebe zulassen, liefern die Daten zum Verkehrswegenetz aus ATKIS ein geometrisches Liniennetzwerk bestehend aus einzelnen Wegsegmenten, welches sich durch Ableitung von Inzidenz- und Adjazenzbeziehungen aus der impliziten Netztopologie in eine Graphstruktur überführen lässt.
Unter Minimierung der Kostensumme (Algorithmus nach Dijkstra, 1959) lassen sich basierend auf der Lage landwirtschaftlicher Betriebe sowie der zugehörigen Parzellen landesweit flächendeckend landwirtschaftliche Hof-Feld-Transportpfade abschätzen. Zur Berücksichtigung von Präferenzen für einzelne Wegekategorien werden den einzelnen Kanten (= Segmenten des Netzwerks) angenommenen Geschwindigkeiten zugeordnet, so dass sich die Kosten für das Überfahren eines Wegsegments aus der Segmentlänge und der angenommenen Geschwindigkeit ergeben und äquivalent zur benötigten Fahrzeit sind.
Die Abbildung der errechneten Pfade erfolgt kleinteilig in Form von Referenzen auf die beteiligten Elemente des Verkehrsnetzes sowie die zugehörigen Parzellen und Hofstellen. Dadurch lassen sich einerseits detaillierte Angaben zur Zusammensetzung von Einzelpfaden ableiten und andererseits können durch Aggregation Erschließungsflächen für einzelne Segmente abgeleitet werden (eine detaillierte Beschreibung findet sich in MACHL et al., 2016).

3 Modellprojekt „NES-Allianz“

Im Rahmen des Modellprojekts „Kernwegekonzept NES-Allianz“ (Zusammenschluss von 14 Gemeinden im unterfränkischen Landkreis Rhön-Grabfeld) wurden die Analyseergebnisse erstmals im Zusammenhang mit der Konzeption eines interkommunalen Kernwegesystems innerhalb einer Region mit Integrierter Ländlicher Entwicklung (ILE) eingesetzt und einerseits hinsichtlich der Qualität des gewählten Schätzansatzes sowie andererseits bezüglich des möglichen Nutzens als zusätzliche Planungs- bzw. Entscheidungsgrundlage evaluiert.

Zentrale Fragestellungen des Modellprojekts sind daher:

- Wie gut können tatsächliche Hof-Feld-Transportpfade durch wegekategorie-spezifische Kantengewichtung und Algorithmen abgebildet werden? Was sind die Ursachen von Abweichungen zwischen geschätzten und tatsächlichen Pfaden.
- Liefert die geschätzte Erschließungsfläche einzelner Wegesegmente hilfreiche Aufschlüsse für die Konzeption überregionaler ländlicher Kernwegesysteme?
- Wo liegen Grenzen des Ansatzes und worin bestehen Einschränkungen der genutzten Methode?

3.1 Abgleich zwischen geschätzten und tatsächlichen Hof-Feld-Transportpfaden

Der Abgleich zwischen geschätzten und tatsächlichen Hof-Feld-Pfaden erfolgte durch Befragung von 13 landwirtschaftlichen Betrieben sowie einem Biogasanlagenbetreiber vor Ort. Die Betriebsgrößen bewegten sich in einem Bereich von 20 bis deutlich über 1000 ha (25 %-Quantil: 78 ha, Median 122 ha, 75 %-Quantil: 181 ha). Für jede bewirtschaftete Parzelle wurde eine Karte mit den geschätzten Pfaden vorgelegt und mit den tatsächlichen Routen abgeglichen. Abweichungen wurden mit Ursache und Verlauf festgehalten. Insgesamt wurden im Rahmen der Befragung 797 Parzellen abgefragt. Die Schätzqualität wurde mit Hilfe des Indikators „Grad der Übereinstimmung“ (GDÜ) quantifiziert. Dieser ergibt sich aus dem Quotienten der Länge des übereinstimmenden Verlaufs und der Länge des geschätzten Pfades. Für 73 % der abgefragten Pfade ergab sich eine vollständige Übereinstimmung, 76 % der Pfade wiesen eine Übereinstimmung von mindestens 90 % auf und für 81 % der Routen ergab sich eine Übereinstimmung von mehr als 80 % der Pfadlänge.

Neben dem Meiden von Ortschaften, stark frequentierten Hauptstraßen, Verkehrswegen mit schlechter Qualität sowie zu gering dimensionierten Verkehrsbauwerken erwies sich bei der Befragung insbesondere die Steigung einzelner Wege- und Straßensegmente als Hauptursache für Abweichungen zwischen tatsächlicher und errechneter Transportrouten. Während sich die ersten beiden Ursachen mit Hilfe veränderter Kantengewichte für einzelne Wegekategorien in die Schätzung miteinbeziehen lassen, erfordern die übrigen Ursachen die Erhebung zusätzlicher Information bzw. die Ableitung weiterer Parameter aus anderen Datenbeständen.

Insgesamt liefert die Schätzung der Hof-Feld-Transportpfade jedoch sehr zufriedenstellende Ergebnisse und lässt daher eine gute Schätzqualität der errechneten Erschließungsflächen für Einzelsegmente erwarten.

3.2 Nutzen der Analyseergebnisse im Zusammenhang mit der Erstellung interkommunaler Kernwegesysteme

Als Grundlage zur Entscheidungsunterstützung wurden die um die Ergebnisse der geschätzten Erschließungsfläche angereichterten ATKIS-Wegsegmente dem mit der Konzeption des landwirtschaftlichen Kernwegenetzes beauftragten Planungsbüro (BBV-Landsiedlung GmbH Würzburg) bereitgestellt. Den geschätzten Erschließungsflächen je Einzelsegment liegen mehr als 80.000 Einzelpfade zu Grunde, welche eine potentielle Fahrbeziehung innerhalb bzw. über die Grenzen der ILE-Region hinausreichende Fahrbeziehung darstellen. Damit ist die regionale wie auch die überregionale Erschließungsfunktion einzelner Segmente berücksichtigt. Neben

reinen Hof-Feld-Pfaden wurden auch die Fahrbeziehungen zwischen den Parzellen der Region und vier größeren Biogasanlagen der Region in den Schätzungen berücksichtigt.
Nach Angaben des beauftragten Planungsbüros liefern die geschätzten Angaben zur Erschließungsfläche einzelner Wegsegmente einen wesentlichen Anhaltspunkt sowie eine objektive Grundlage zur Erkennung von Haupt- und Nebenachsen des Verkehrswegenetzes. Dadurch wird eine gezieltere Begutachtung einzelner Wegtrassen bei Vor-Ort Befahrungen möglich. Darüber hinaus bietet die Erschließungsfläche als Indikator eine objektive Argumentationsgrundlage für die Diskussion mit den Planungsbeteiligten vor Ort. Durch die über die ILE-Region hinausgehende Betrachtung der Transportbeziehungen und damit auch der überregionalen Erschließungsfunktion einzelner Wege lassen sich zudem Anschlussstellen zu benachbarten Wegenetzen detektieren. Als objektiver Indikator für die Bedeutung einzelner Segmente im Gesamtnetz liefert die aus Einzelrouten abgeleitete Erschließungsfläche einzelner Wegsegmente insgesamt eine wichtige Grundlage zur Entscheidungsunterstützung im Zusammenhang mit der Konzeption interkommunaler Kernwegenetze für die Landwirtschaft.
Neben der Betrachtung des Ist Zustandes können mit Hilfe des Werkzeugs unter dem Aspekt des Geodesigns zudem mögliche künftige Zustände der Realwelt bereits in einem virtuellen Abbild der Realität erfasst, evaluiert und hinsichtlich der möglichen Lenkung des landwirtschaftlichen Verkehrs analysiert werden. So können Wege im digitalen Abbild der Realität virtuell ausgebaut und Effekte der Planung bereits in einer frühen Phase - ceteris paribus - betrachtet und diskutiert werden.

4 Diskussion

Der vorgestellte Ansatz zur Ableitung der Erschließungsfläche aus einzelnen Hof-Feld-Transportrouten stützt sich auf tatsächliche Hof-Feld-Transportbeziehungen und bezieht dabei sowohl die regionale wie auch die überregionale Erschließungsfunktion einzelner Wege mit ein. Dabei lässt der im Rahmen der Untersuchungen angetroffene hohe Grad der Übereinstimmung zwischen geschätzten und tatsächlich gewählten Routen eine realitätsnahe Abbildung der über ein Segment erschlossenen Fläche erwarten.
Aufgrund des Wegfalls von Betrieben, wechselnder Pachtverhältnisse oder aber der Veränderung der landwirtschaftlichen Produktionsrichtung unterliegen die bestehenden Fahrbeziehungen einem steten Wandel, was eine Prognose hinsichtlich künftiger Entwicklungen erschwert. Aus diesem Grund sind neben der Erschließungsfläche zusätzliche, komplementäre Kriterien wie beispielsweise die Rolle einzelner Wegsegmente für die Anbindung von Parzellen an das übergeordnete Verkehrswegenetz heranzuziehen. In diesem Zusammenhang wurden am Lehrstuhl für Geoinformatik der TUM bereits entsprechende Werkzeuge entwickelt.

5 Fazit

Das im Rahmen des Forschungsprojekts LandModell^{TUM} konzipierte Werkzeug zur Analyse landwirtschaftlicher Transportpfade liefert für die gewählte Modellregion eine realitätsnahe Abbildung tatsächlicher Hof-Feld-Routen. Die wegen fehlender Angaben hinsichtlich Zufahrtsbeschränkungen (z.B. Einbahnstraßen) oder Abbiegevorschriften eingeschränkte Routingtauglich-

keit des ATKIS-Verkehrsnetzes zeigte im Modellprojekt keine wesentlichen Effekte auf die Qualität der Routenberechnung.
Durch wegekategorie-spezifische Kantengewichtung konnten die Routenwahl der Landwirte gut abgebildet werden. Zusätzliche, für die Routenwahl entscheidende Einflussparameter wie die Steigung von Wegsegmenten oder die zu gering ausgelegte Dimensionierung von Brücken oder Unterführungen sind durch Einbezug zusätzlicher Datenquellen bzw. Erhebung vor Ort zu berücksichtigen. An dieser Stelle könnte sich ein digitales Wegekataster als sehr hilfreich erweisen.
Das entwickelte Werkzeug ist aufgrund des modularen Aufbaus zudem in der Lage, neben einer Betrachtung der Ist-Situation auch verschiedene Planungsszenarien zu analysieren. Einzelne Wege können in einem digitalen Abbild einer möglichen zukünftigen Realität virtuell ertüchtigt und die Effekte dieses Ausbaus auf die Transportströme bereits in einer frühen Planungsphase sichtbar gemacht werden.

6 Dank

Die Finanzierung des Forschungsprojekts LandModell^{TUM} erfolgt durch die Bayerische Verwaltung für Ländliche Entwicklung

7 Literatur

BERTLING, H., KRIESE, H. & LÜTKEMEIER, H., 2015: Neue Planungsansätze im ländlichen Wegebau. zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement, **140**(5), 320-326.

BROMMA, R., 2014: Konzeption und Umsetzung eines ländlichen Kernwegenetzes in der Allianz »Fränkischer Süden«. Mitteilungen des DVW- Bayern e.V., **66**(4), 379-389.

DWA – DEUTSCHE VEREINIGUNG FÜR WASSERWIRTSCHAFT, ABWASSER UND ABFALL E.V, 2016: Richtlinien für den Ländlichen Wegebau (RLW) – Teil 1: Richtlinien für die Anlage und Dimensionierung Ländlicher Wege. DWA-Regelwerk, Arbeitsblatt DWA-A 904-1.

GOCKEL, R., 2012: Zukunftsfähige ländliche Infrastruktur – Welche Hilfen benötigt die Landwirtschaft? DLKG (Hrsg.): Wege in die Zukunft!? – Neue Anforderungen an ländliche Infrastrukturen. Schriftenreihe der DLKG **9**, 31-37.

MACHL, T., EWALD, W.G., DONAUBAUER, A. & KOLBE, T.H., 2016: Entwicklung eines Werkzeugs zur landesweit flächendeckenden Analyse landwirtschaftlicher Transportbeziehungen in Bayern. zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement, **141**(3), 197-205.

SOBOTH, A., 2012: Notwendigkeiten und Möglichkeiten von Kooperationen gemeindeübergreifender Wegenetze. DLKG (Hrsg.): Wege in die Zukunft!? – Neue Anforderungen an ländliche Infrastrukturen. Schriftenreihe der DLKG, **9,** 39-44.

BGT, the realisation of a Key Register of Large Scale Topography in the Netherlands

RUUD VAN ROSSEM[1]

Extended abstract

In The Netherlands a number of Key Registers are operational. Key Registers are registers of e.g. persons, organisations and buildings. They are used on all levels of government and this use is mandatory. One of these Key Registers is the Key Register of Large Scale Topography (BGT: Basisregistratie Grootschalige Topografie, Fig. 1). It is a detailed base map at a scale of 1: 1000. This Key Register has been realized between 2008 and 2017.

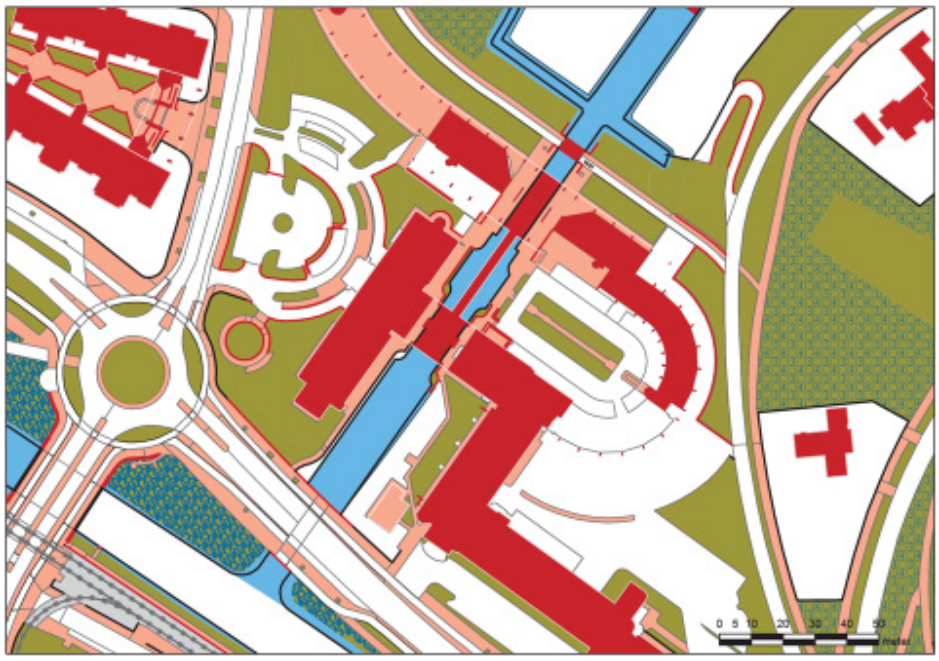


Fig. 1: An example of the BGT

Before the development of this Key Register large scale maps already existed in The Netherlands, but these were not standardized. Earlier attempts to standardize had failed. The biggest challenge at the beginning of the BGT program was to assure the co-operation of all relevant parties in creating this standardized map of the country. The content of the map is made by the municipalities, the provinces, the regional water authorities, Rijkswaterstaat (national organization responsible for motorways and water management), the ministry of Defense, the ministry of Agriculture and ProRail (responsible for the railroads). The Cadastre is responsible for the nationwide database and the publication and distribution of the information.

[1] Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Ruimtelijke Ordening / Beleid Geo-Informatie, Meester Bierensweg 66, 4841AL Prinsenbeek, Niederlande,
E-Mail: Ruud.van.Rossem@kpnmail.nl

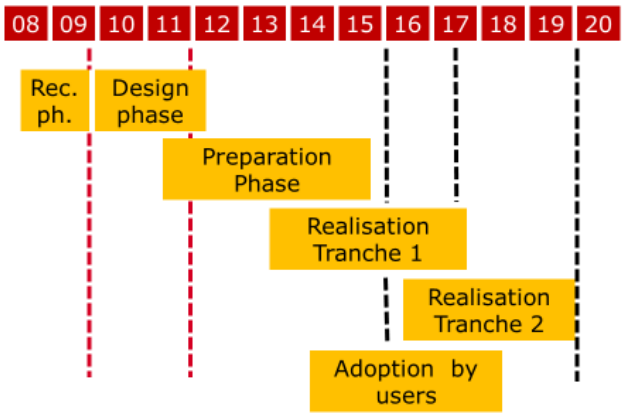


Fig. 2: Time Schedule of the BGT-Program

The design phase started in 2008 and finished in 2011 (Fig. 2). During this phase not only the functional design and the information model were developed, but also the organization model, the financial agreements, the approach for the transition and the proposal for Law.
After the design phase the proposal for the BGT Law went into the legislation process and parallel to this the preparations were made for the realization of the Key Register. Most important in this phase were the creation of an organization to support the transition towards BGT, the development of the central ICT-system, and communication with all parties involved.
The actual realization of the content of the Key Register happened from 2014 to 2017. Here the difficulty was that all the parties involved had to run their own realization projects and this required a lot of attention. After this phase the BGT covers the whole country and is available for use. The BGT is available as Open Data (Fig. 3).
In the presentation the author, who was responsible for this program in the role of general program manager, will give an insight in the way the various challenges have been dealt with.

Fig. 3: Use of the BGT in combination with other information

Sicherheitsaspekte und Best Practices in Geodateninfrastrukturen für den Einsatz in Smart Cities

ANDREAS MATHEUS[1]

Zusammenfassung: Heutige Geodateninfrastrukturen basieren zumeist auf offenen Standards des Open Geospatial Consortiums, die eine Interoperabilität für Daten- und Dienste sicherstellen. Bedenkt man, dass diese Standards oft 10 Jahre oder älter sind und keine Sicherheitsaspekte berücksichtigen, so muss sich die Frage auftun ob diese Standards noch zeitgemäß sind, wenn es darum geht, geschützte oder schützenswerte Daten- und Dienste in Geodateninfrastrukturen mit modernen Web-Applikationen abzurufen.
In diesem Beitrag werden einige relevante Sicherheitsaspekte und Best Practices mit dem Ziel beschrieben, existierende Geodateninfrastrukturen „aufzurüsten" damit branchenübergreifende Nutzung abgesicherter Dienste möglich wird. Finden Sie am Beispiel Smart Cities heraus wie sich moderne grundlegende IT-Security und altgediente OGC Web Services komplementieren.

1 Einleitung und Motivation

INSPIRE war der große Motor in Europa, der die Entwicklung von Geodateninfrastrukturen in den letzten 10 Jahren maßgeblich geprägt hat. Diese Entwicklung ist allerdings zum Stehen gekommen, denn der originäre INSPIRE Use Case war und ist der freie Austausch von offenen Geodaten spezieller Themen um eine gemeinsame Umweltpolitik in Europe zu ermöglichen. Und laut Gesetzgebung soll genau dies Ende 2018 möglich sein. Obwohl es durchaus wünschenswert wäre diese Infrastruktur auch für kommerzielle Nutzung zu verwenden, ist dies so ohne weiteres nicht möglich, denn Sicherheitsaspekte und deren interoperable Umsetzung fehlen seit der ersten Stunde. Überprüft man die aktuell gültigen Implementation Guidelines für z.B. den INSPIRE Download Service (V3.1 von 09.08.2013 oder INSPIRE View Service (V3.11 vom 04.04.2013) so stellt man fest, dass es keine Implementation Guidance für Security gibt[2]. Wenn man zusätzlich die normativen Referenzen untersucht, so muss man zu dem Schluss kommen, dass INSPIRE nur auf HTTP und nicht einmal auf HTTP**S** betrieben werden darf ohne eine Interoperabilitätsgarantie zu verlieren. Denn weder die Implementation Guidelines noch die OGC Web Service Standards erfordern HTTP über TLS (HTTPS). Mehr Informationen zu diesem Thema sind in OGC Testbed Engineering Reports verfügbar (OGC #15-022, OGC #16-048r1 und OGC #17-021r2).
Sind nun Geodateninfrastrukturen wie INSPIRE unbrauchbar und müssen ersetzt werden, wenn es darum geht das Konzept von Smart Cities umzusetzen? Das Konzept von Smart Cities ist eine Evolution in die Vernetzung verschiedenartiger Daten für Städte, um effizienter, produktiver und gleichzeitig attraktiver und grüner zu werden. Die intelligente Vernetzung von Geodaten mit

[1] Secure Dimensions GmbH, Waxensteinstr. 28, D-81377 München, E-Mail: am@secure-dimensions.de
[2] Der View Service hat beispielsweise 92 Implementation Requirements, jedoch kein einziges adressiert Security!

Daten anderer Branchen ist dabei eine der wesentlichen Herausforderungen. So ist eins der Technologieziele die Vernetzung von möglichst vielen Sensoren und anderen Messplattformen, um relevante Umweltdaten zu gewinnen. Damit man nicht bei Null anfangen muss, sollte es ebenfalls möglich sein, existierende Geodateninfrastrukturen so zu erweitern, dass über diese Geodaten, Stadtmodelle, Sensordaten und Daten von kommerziellen Anbietern wie z.B. Energiekonsum und Umweltdaten abgerufen werden können. Dies erfordert allerdings die Integration einer interoperablen und modularen sowie flexiblen Sicherheitslösung.
In diesem Beitrag soll für den speziellen Use Case „Smart Cities" aus dem Projekt Smart Sustainable Districts[3] exemplarisch gezeigt werden, wie eine beispielhafte Geodateninfrastruktur mit OGC Web Services elegant „aufgerüstet" werden kann um die Vernetzung von zugriffsgeschützten Sensorinformationen zu ermöglichen, ohne die eigentliche Basis stark verändern zu müssen.

2 Sicherheitsaspekte und Standards

Das Schlagwort „Sicherheit" im Zusammenhang mit Geodateninfrastrukturen wird oft missverstanden, wenn eine genaue Definition fehlt. Eine umfassende Definition erforderlicher Sicherheitsaspekte, die für Geodateninfrastrukturen relevant sind, findet man in ISO 7498-2 und ISO Multipart Standard 10181. In ISO 10181 werden diese Aspekte dann mittels verschiedener Frameworks standardisiert um eine Umsetzung in verteilten, offenen Architekturen zu ermöglichen: „*The series of Recommendations | International Standards on Security Frameworks for Open Systems addresses the application of security services in an Open Systems environment, where the term "Open Systems" is taken to include areas such as Database, Distributed Applications, Open Distributed Processing and OSI. The Security Frameworks are concerned with defining the means of providing protection for systems and objects within systems, and with the interactions between systems. The Security Frameworks are not concerned with the methodology for constructing systems or mechanisms*." [ISO 10181-1, p.7]
In ISO 10181 sind folgende Frameworks definiert: *Authentication [ISO 10181-2]*, *Access Control [ISO 10181-3]*, *Non-Repudiation [ISO 10181-4]*, *Confidentiality [ISO 10181-5]*, *Integrity [ISO 10181-6]* und *Security Audit and Alarms [ISO 10181-7]*. ISO 10181-1 bietet einen Überblick.
Ohne die verschiedenen Frameworks im Detail vorzustellen kann man festhalten, dass *Non-Repudiation* und *Security Audits and Alarms* In Geodateninfrastrukturen üblicherweise nicht relevant sind.
Relevant sind hingegen *Confidentiality* und *Integrity,* die eine wichtige Rolle bei der Umsetzung von Informationsflusskontrollen haben, aber auch vor unerlaubtem Mithören schützen (Confidentiality) sowie das unbemerkte Verändern von Information verhindern (Integrity). Bei der Umsetzung muss man grundsätzlich unterscheiden, ob ein Schutz nur während der Übertragung garantiert werden soll oder ob ein persistenter Schutz, unabhängig vom Übertragungsmedium etabliert werden soll. Zum Schutz bei der Übertragung kann eine Implementierung basierend auf

[3] https://www.gis.bgu.tum.de/projekte/smart-sustainable-districts-ssd/

dem OSI-Referenzmodell auf der Anwendungsschicht (Layer 7) oder Transportschicht (Layer 4) erfolgen. Im ersten Fall ist SOAP plus WS-Security eine standardisierte Möglichkeit, die allerdings zumeist nur für die Absicherung von Kommunikation zwischen Anwendungen im Back-Office eingesetzt wird. Client Anwendungen werden üblicherweise mittels eines verschlüsselten Kanals angebunden; also HTTP über TLS (HTTPS). Weil diese standardbasierte Lösung (IETF RFC 7246) im OSI Referenzmodell zwischen Layer 7 und 4 angesiedelt ist hat sie eine grundlegende Abhängigkeit: sie erlaubt nur die Verschlüsselung eines TCP Kanals zwischen Netzwerkendpunkten; nicht Anwendungen!

Die wichtigsten Frameworks, die eigentlich zur Absicherung einer Geodateninfrastruktur immer genannt werden, sind Authentication und Access Control. Zumeist existiert eine Abhängigkeit zwischen Access Control (also der Zugriffskontrolle) und der Authentication (also der Nachprüfbarkeit von behaupteten Identitäten), denn Zugriffsrechte werden oft mit identifizierten Benutzern verknüpft.

Für die Umsetzung der Authentication gibt es scheinbar endlos viele standardbasierte Möglichkeiten und proprietäre Verfahren, die immer für einen ganz bestimmten Zweck entwickelt wurden. Zwei wichtige Standards im Zusammenhang mit diesem Beitrag sind SAML2 und OpenID Connect.

Die Security Assertion Markup Language (SAML2) wurde 2005 von OASIS standardisiert und definiert verschiedene Protokolle sowie XML Strukturen zum Austausch von Zusicherungen. Hierbei können Zusicherungen die Identität eines Benutzers garantieren oder dessen Zugriffsrechte beschreiben. Es wird SOAP mit WS-Security sowie XML Digital Signature und XML Encryption eingesetzt um den Austausch von Zusicherungen bzgl. Vertraulichkeit und Integrität abzusichern. Zusätzlich werden Schlüsselpaare verwendet um ebenfalls die Authentizität von Zusicherungen überprüfen zu können. SAML2 wird heutzutage beispielsweise in allen produktiven akademischen Föderationen verwendet. Eine Auflistung der weltweit produktiven Föderationen (aktuell 69) findet man unter https://refeds.org/federations.

OpenID Connect ist eine Spezifikation, die nicht von einem bekannten Standardisierungsgremium verabschiedet wurde. Es wird als Community- oder quasi-Standard bezeichnet der eine Abfrage von benutzerbezogenen Informationen als Erweiterung von OAuth2 ermöglicht. OAuth2 ist als RFC 6749 (und RFC 6750) vom IETF 2012 verbschiedet worden. Allerdings sind bei der Nutzung von OpenID Connect die Vorgaben von OAuth2 zu berücksichtigen: Im Wesentlichen geht es darum, dass ein OAuth2 Access Token nur zur Rechtedelegation verwendet wird, aber damit keine personenbezogen Daten abrufbar sind. Bei der Erweiterung von OpenID Connect kann zusätzlich ein sog. ID-Token übertragen werden, das dann der Applikation ermöglicht, bestimmte personalisierte Seiten aufzubauen. Ebenso ist es möglich, dass personenbezogene Informationen mit dem Access Token abgefragt werden können. Im Vergleich zu SAML2 sind die OpenID Connect Zusicherungen als JSON strukturiert, die zur Umsetzung von Integrity oder Confidentiality als JWT oder JWE kodiert werden können. JWT ist in RFC 7519 definiert und erlaubt die Anwendung einer digitalen Signatur auf die JSON Struktur. JWE ist in RFC 7516 definiert und erlaubt die Verschlüsselung von JSON und damit die Umsetzung von Confidentiality. Zur Umsetzung von Confidentiality und Integrity können beide Verfahren nacheinander angewendet werden. In einer Geodateninfrastruktur ist es wichtig, standardisierte Verfahren für die Authentifizierung zu verwenden, um eine grundlegende Interoperabilität zwischen den Ser-

viceanbietern auch auf der Ebene der Sicherheit zu gewährleisten. Dies ist wichtig, damit u.a. ein geeignetes Rechtemanagement durchgeführt werden kann.
Für die Umsetzung von Access Control (also der eigentlichen Zugriffskontrolle) muss nicht notwendigerweise ein Standard eingesetzt werden, wenn man eigenverantwortlich in der eigenen Domäne agiert, denn dieses Framework hat dann keinen Einfluss auf die Interoperabilität der Geodateninfrastruktur. Allerdings sollte berücksichtigt werden, dass eine Abstimmung von Zugriffsrechten über die einzelnen Anbieter hinweg dann zu Schwierigkeiten und Missverständnissen kommen kann, wenn jede Organisation eine eigene Beschreibung der Zugriffsrechte nutzt. Eine „general purpose“ Sprache wurde von OASIS entwickelt: die eXtensible Access Control Markup Language (XACML) ist heute in verteilten Main-Stream IT Lösungen im Einsatz. Damit die Sprache auch zur Definition und Durchsetzung von geographischen Zugriffsbeschränkungen verwendet werden kann wurde vom OGC GeoXACML entwickelt und 2008 standardisiert. Im Wesentlichen definiert GeoXACML den Datentyp Simple Geometry aus ISO 19107 und 19125 und unterstützt das Geometrie-Encoding WKT sowie GML 2 und 3. Durch Anwendung von GeoXACML können interoperable, raum-zeit-basierte Zugriffskontrollregeln definiert und durchgesetzt werden. Als Beispiel kann der Ortsbezug von Benutzern und Ressourcen (Features oder Karten) oder dessen Historie berücksichtigt werden.

3 OGC Web Services in modernen Geodateninfrastrukturen

Der Einsatz von OGC Web Services ermöglicht das Abrufen von Geodaten in verschiedenen Formaten über standardisierte Schnittstellen. Allerdings sind OGC Web Service Spezifikationen autark bzgl. Security was die Frage aufwirft, ob mit diesen Diensten eine moderne GDI überhaupt aufgebaut werden kann. Der Begriff „modern“ zielt darauf ab, dass nicht nur Dienste abgesichert sind, sondern deren Zugriff über moderne Web-Applikationen erfolgt; also nicht mehr vorwiegend durch Desktop Applikationen.

3.1 OGC Web Services und HTTPS

OGC Web Service Spezifikationen sind ca. 10 Jahre oder älter und mit dem Ziel entwickelt, offenen Zugang zu Geodaten interoperabel zu ermöglichen. Heutzutage wird aber durch Policies vorgeschrieben, dass HTTPS verwendet werden muss. Ebenso implementieren Web-Browser, dass zu einer eingegebenen Web-Adresse standardmäßig ein HTTPS Verbindungsaufbau erfolgt. Wie in OGC Testbed Engineering Reports veröffentlicht, ist eine Service-Instanz die auf HTTPS betrieben wird nicht mehr standardkonform, weil keine normativen Anforderungen für dieses Hosting in den OGC Standards existieren. Somit ist man als Service-Anbieter darauf angewiesen zu hoffen, dass Client Anwendungen auch HTTPS unterstützen. Dies ist kein Thema für Webbasierte Anwendungen, weil sie ja im Web-Browser ausgeführt werden. Allerdings wurde in OGC Testbeds immer wieder festgestellt, dass Desktop-Applikationen entweder über keinen HTTPS Support verfügen, oder aber z.B. die Zertifikatprüfung unsachgemäß implementiert wurde.
Damit das Hosting von OGC Web Services auf HTTPS wieder auf interoperablen Grundlagen verankert ist hat das OGC in der Standards-Arbeitsgrupp OWS Common – Security SWG bereits einen Draft Standard entwickelt, der vermutlich im Juni 2018 verabschiedet wird. Der OGC Web

Services Security Standard (OGC #07-007) definiert ein Standardvorgehen um Sicherheits-Meta-Information in die Service-Capabilities zu integrieren. Das vorgestellte Verfahren ist backwards-compatible und ermöglicht somit eine Nutzung von kompatiblen Diensten mit existierenden Client-Anwendungen. Optional kann eine Service-Instanz pro Netzwerkendpunkt ebenso den Support anderer Security Features beschreiben. Es können ebenso Sicherheitsanforderungen wie Authentication oder Access Control, sowie Integrity und Confidentiality für SOAP basierte Service Instanzen beschrieben werden. Ziel ist es hierbei der Client-Anwendung Metainformationen mitzuteilen, damit z.B. die Konfiguration einer SAML2 oder OpenID Connect Authentication automatisch erfolgen kann.

3.2 Authentifizierungs-Föderation

Um eine Geodateninfrastruktur mit geschützten Diensten nutzbar zu machen ist ein GDI übergreifendes Rechtemanagement erforderlich. Damit dies geschehen kann muss es möglich sein, dass ein Benutzer sich nur einmal Anmelden muss und danach seine Identität bei allen Diensteanbietern akzeptiert wird. Dies kann durch die Abtrennung von Nutzermanagement und Service-Provisioning und dem Einsatz von SAML2 geschehen. Es gibt dann die sog. Identity Provider Entitäten die Benutzerkonten verwalten, eine Loginfunktionalität bereitstellen und anfragenden Service Provider Entitäten Zusicherungen über die Identität bereitstellen. Ein wesentlicher Vorteil von SAML2 gegenüber OpenID Connect ist das Trustmanagement: Es ist zwischen Entitäten und nicht anwendungsabhängig. Ein weiterer Vorteil von SAML2 ist die Unterstützung von Single-Sign-On was eine wesentliche Benutzerfreundlichkeit sicherstellt, denn ein Benutzer kann nach einmaligem Login viele geschützte Dienste unterschiedlicher Anbieter in einer Anwendung nutzen, ohne sich bei jedem Anbieter erneut anmelden zu müssen.
Die Trennung von Service und Nutzerverwaltung in verschiedene Entitäten hat unter der neuen Datenschutzrichtlinie in Europa (GDPR) die im 25. Mai 2018 in Kraft treten wird einen weiteren wesentlichen Vorteil. Weil ein Service Provider dann keine personenbezogenen Daten mehr erheben und verwalten muss, entfallen bestimmte Anforderungen die die GDRP verlangt. Speziell zertifizierte Unternehmen könnten nun die Benutzerverwaltung übernehmen und die Funktionalitäten für die Ausweispflicht implementieren. Allerdings muss der Service Provider ein (vereinfachtes) Privacy Statement bereitstellen, in dem vor allem das Gebot der Datensparsamkeit für personenbezogene Daten garantiert wird. Dies macht erforderlich, dass für jeden Zugriff nur so viele personenbezogene Daten genutzt werden wie zur Erbringung des Auftrags erforderlich ist.

3.3 Web-basierte Anwendungen und CORS

Web-basierte Applikationen nutzen JavaScript zur Implementierung und werden im Web-Browser ausgeführt. Bibliotheken wie OpenLayers, AngularJS oder Leaflet ermöglichen die schnelle und funktionsstarke Umsetzung von Geo-Web-Apps. Allerdings unterliegen diese JavaScript-Anwendungen der sog. Same Origin Policy, die 2010 vom W3C veröffentlicht wurde und in jedem Web-Browser durchgesetzt wird. Diese Policy verhindert, dass JavaScript Code, der von einem Web-Server geladen wurde, ohne weiteres auf andere Web-Server zugreifen kann. Das W3C hat in 2014 die sog. Cross-Origin Resource Sharing (CORS) Recommendation herausgegeben, in der genau definiert wird wie dieser Zugriffsschutz funktioniert. Bestimmte Zugriffe auf andere Web-Server die durch JavaScript (z.B. AJAX) ausgeführt werden, werden vom

Web-Browser dahingehend untersucht, ob bestimmte HTTP Antwort Header vorhanden sind. Welche dies sind, hängt vom jeweiligen HTTP Request ab, der durch JavaScript ausgeführt wurde.
Diese Regelung des Zugriffs auf andere Web-Server als jene, von der die JavaScript Anwendung geladen wurde, hat für die Nutzung in einer Geodateninfrastruktur erhebliche Auswirkungen. Wird eine Web-Anwendung beispielsweise von GeoApps.de geladen und sollen dann OGC Web Services von Service-A.net und Service-B.de aufgerufen werden, so fordert der Web-Browser die Existenz von CORS HTTP Antwort Headern. Im einfachsten Fall macht die Web-Anwendung eine HTTP GET Anfrage auf z.B. Service-A.net. Für diese Anfrage setzt der Web-Browser das HTTP Header „Origin: GeoApps.de“. Bevor der Web-Browser die Antwort von Service-A.net an die Web-Applikation weitergibt wird geprüft ob der geforderte HTTP Header „Access-Control-Allow-Origin“ vorhanden ist. Ist der Wert entweder „*“ oder „GeoApps.de“ so wird die Antwort an die Web-Applikation weitergegeben. Es ist also erforderlich, dass jeder Service-Betreiber seinen Web-Server so konfiguriert, dass entsprechende CORS Header gesetzt werden. Dies ist relativ kompliziert, da CORS verschiedene Use Cases vorsieht.
Die Existenz von Same Origin Policy und CORS hat allerdings erhebliche Auswirkungen auf die Performance von Web-Applikationen, denn beispielsweise wird für alle Anfragen die nicht Standard HTTP Header beinhalten, ein sog. Pre-Flight-Request vom Web-Browser verschickt. Nicht Standard HTTP Header sind z.B. der Header Authorization der zur Übertragung von User Informationen verwendet wird. Beispielsweise in HTTP Basic Authentication wird hier der UserName:Password String als Base64 kodiert übertragen. Ebenso kann ein OAuth2 Access Token als HTTP Header Authorization Bearer verschickt werden. Durch diese Anfragen erhöht sich also die Netzwerkauslastung, denn einem HTTP GET oder POST Request wird nun ein OPTIONS Request vorangestellt.
Letztendlich muss jeder Service-Anbieter CORS ordnungsgemäß implementieren, was üblicherweise weit darüber hinaus geht ein „*“ zurückzuliefern. Neben dem obligatorischen Withelisting der vertrauenswürdigen Domains sind auch Überlegungen für HTTP POST, PUT und DELETE erforderlich. Jedenfalls muss die HTTP Methode OPTIONS ohne Zugriffsschutz bereitgestellt werden, damit die Pre-Flight-Request ordnungsgemäß beantwortet werden.

3.4 Session Management und CORS

Sollen Services abgesichert werden, so muss ein Session Management etabliert werden mit dem üblicherweise ein Security-Context verknüpft ist. Grundsätzlich gibt es zwei Modelle: Serverseitiges und Client-seitiges Session Management. Beim Server-seitigem Modell nutzt man üblicherweise HTTP Cookies, wie sie in RFC 6265 definiert sind, damit die Client-Anwendung in einer Anfrage auf eine Session verweisen kann. Dies ist eine elegante Methode für Web-Applikationen und Web-Browser basierten Zugriffen, denn die relevanten HTTP Cookies werden automatisch im Web-Browser verwaltet. Allerdings wird bei Web-Anwendungen die einen Cross-Site-Request machen das Cookie nur mitgeschickt, wenn der Web-Server das entsprechende HTTP CORS Header „Access-Control-Allow-Credentials“ auf „true“ gesetzt hat. Eine weitere negative Eigenschaft von Server-Side Sessions ist deren Laufzeit: Die Session kann jederzeit ungültig werden und somit kann der Request der Client-Anwendung ungültig sein. Es ist

somit nicht ratsam ein Server-seitiges Session Management zu nutzen, wenn Web-Anwendungen Zugriff haben sollen.
Für Web-Anwendungen sollte Client-seitiges Session Management mit zustandslosen Services genutzt werden, wie es z.B. von OAuth2 unterstützt wird. Es ist die Aufgabe der Web-Anwendung immer ein gültiges Access Token zu haben. Somit kann sichergestellt werden, dass keine komplizierten Callback-Funktionen erforderlich sind, um die Session zu erneuern. Allerdings löst die Übertragung von Access Tokens im HTTP Header (von RFC 6750 empfohlene Art und Weise) bei Requests, die nicht an denselben Web-Server gerichtet sind, von dem die Applikation geladen wurde, einen CORS Pre-Flight-Request aus. Wegen dem damit einhergehenden Performanceverlust muss abgewogen werden, ob die Übertragung des Access Token als Teil der URL tragbar ist. Die Übertragung vom Access Token in der URL ist nach RFC 6750 ebenso zulässig, wird jedoch nach RFC 2396 als „unwise“ eingestuft: „*It is clearly unwise to use a URL that contains a password which is intended to be secret. In particular, the use of a password within the 'userinfo' component of a URL is strongly disrecommended except in those rare cases where the 'password' parameter is intended to be public*.” [RFC 2396, p24]. Die größte Gefahr, dass die Access Token im Klartext ausgegeben werden, ergibt sich durch Access Logs auf dem Web-Server, aber auch durch Redirect Angriffe in denen die ursprüngliche URL - die das Access Token enthält - als HTTP Referer Header mitschickt wird.

3.5 SAML2 vs. Oauth2 Session Management und Cloud Deployment

Vergleicht man die Protokolle von SAML2 und OAuth2 mit denen eine Session aufgebaut wird, so stellt man fest dass SAML2 auf einem 2-Way Handshake basiert: Die Client-Anwendung bekommt beim ersten Zugriff ein temporäres Session Cookie ausgestellt, das dann beim zweiten Zugriff (nach erfolgreichem Login) durch ein echtes Session Cookie ausgetauscht wird. Beim Deployment hinter einem Load-Balancer, wie es typischerweise in Cloud-Umgebungen der Fall ist, muss nun darauf geachtet werden, dass die Session-Information bei allen Instanzen hinter dem Load-Balancer verfügbar ist. Dies ist nicht erforderlich, wenn OAuth2 Access Tokens verwendet werden, denn das Session Management obliegt ja dem Client.
Ob nun allerdings SAML2 Sessions oder OAuth2 Access Tokens verwendet werden hängt nicht nur vom Deployment sondern im Wesentlichen davon ab, wodurch auf die geschützte Resource zugegriffen werden soll: Hat man eine Web-Applikation, so können Access Tokens genutzt werden. Steht allerdings nur der Web Browser zur Verfügung, weil beispielsweise HTML href Links gefolgt wird, ist OAuth2 unbrauchbar weil keine Logik vorhanden ist, die das Access Token einschleust. Dies ist der SAML2 Use Case: Öffnen von geschützten Resourcen direkt durch den Web Browser.
Es kommt somit auf den Use Case an, ob SAML2 Sessions, OAuth2 Access Tokens oder etwa beide Mechanismen genutzt werden müssen um den Zugriff auf eine Web-Resource oder einen Web Service abzusichern.

4 Security Best Practice am Beispiel Smart Cities

An der Technischen Universität München wird am Lehrstuhl für Geoinformatik derzeit an dem Projekt Smart Sustainable Districts gearbeitet. Im Rahmen des Projekts wurde die Smart District

Data Infrastructure (SDDI) (MOSHREFZADEH et al. 2017) entwickelt, die zeigt wie eine Realisierung des Smart Cities Konzepts durch intelligente Vernetzung von CityGML-basierten Stadtmodellen mit Sensorinformationen möglich ist. Hierbei werden überwiegend Standards des OGC verwendet.

Im Rahmen dieses Beitrags soll der folgende Use Case betrachtet werden: Benutzer der SDDI können eine Web-Anwendung laden, mit der die Anzeige eines 3D-Stadtmodells ermöglicht wird. Hierbei wird das 3D.Stadtmodell von einem OGC Web Feature Service abgerufen und mittels CityGML kodiert. Jedes Gebäude ist nun mit einem weiteren OGC Web Feature Service verbunden, über den detaillierte Informationen zum jeweiligen Gebäude durch einen Mausklick abgerufen werden können. Um das Smart Cities Konzept zu realisieren enthalten nun die empfangenen detaillierten Gebäudedaten über das Dynamizer-Konzept (CHATURVEDI & KOLBE 2016) weitere Verlinkung mit zusätzlichen Sensordaten, die von anderen Anbietern mittels OGC Sensor Observation Services (SOS) bereitgestellt werden. In diesem Use Case sind dies Umweltdaten und Energiekonsum der Gebäude. Der Benutzer hat nun die Möglichkeit, durch Klick auf einen entsprechenden Link, diese vernetzten Zusatzinformationen abzurufen. Diese Informationen können entweder im Web-Browser oder durch zusätzliche Web-Applikationen dargestellt werden.

Vom 1. September bis 15. Dezember 2017 wurde in Zusammenarbeit mit Secure Dimensions GmbH ein Security-Konzept umgesetzt, das es ermöglicht den Abruf der detaillierten Gebäudeinformation sowie der verlinkten Sensor-Zusatzinformationen zu schützen. Es soll also jeder Benutzer in der Lage sein das 3D-Stadtmodell zu sehen, aber die verlinkte Sensor-Information kann nur durch autorisierte Benutzer abgerufen werden. Insbesondere sollte es möglich sein, dass sich Benutzer mittels TUM oder Google Login anmelden können und sich nur einmal anmelden müssen, um auf die Ressourcen unterschiedlicher Anbieter zugreifen zu können (Single-Sign-On). Die Umweltdaten können alle Benutzer nach erfolgreichem Login sehen; die Smart-Meter Informationen eines Gebäudes – also z.B. der Stromverbrauch – soll nur durch TUM Benutzer möglich sein. Aus technischer Sicht war eine der wesentlichen Herausforderung eine Absicherung der Verlinkung von Zusatzinformationen zu garantieren, obwohl die eigentlichen Links im CityGML Modell oder in der Web Feature Service Antwort nur **ohne** Sicherheitskontext verfügbar sind. Und dies benutzerfreundlich durch Single-Sign-On das anbieterübergreifend akzeptiert wird.

Um eine realitätsnahe Umgebung für die Umsetzung des Security-Demonstrators zu haben, wurden Anwendungen und Services verteilt: die 3D CityGML Web-Anwendung wird von einem Anbieter bereitgestellt (www.3dcitydb.org); der WFS mit den detaillierten Gebäudeinformationen wird von Secure Dimensions bereitgestellt (wfs.sddi.secure-dimensions.de); der SOS für die Umweltdaten sowie der SOS für die Smart-Meter-Daten werden von der TUM bereitgestellt (ssdsos1.gis.bgu.tum.de) und (ssdsos2.gis.bgu.tum.de).

Die folgende Abbildung zeigt die Komponenten des Security-Demonstrators im Überblick.

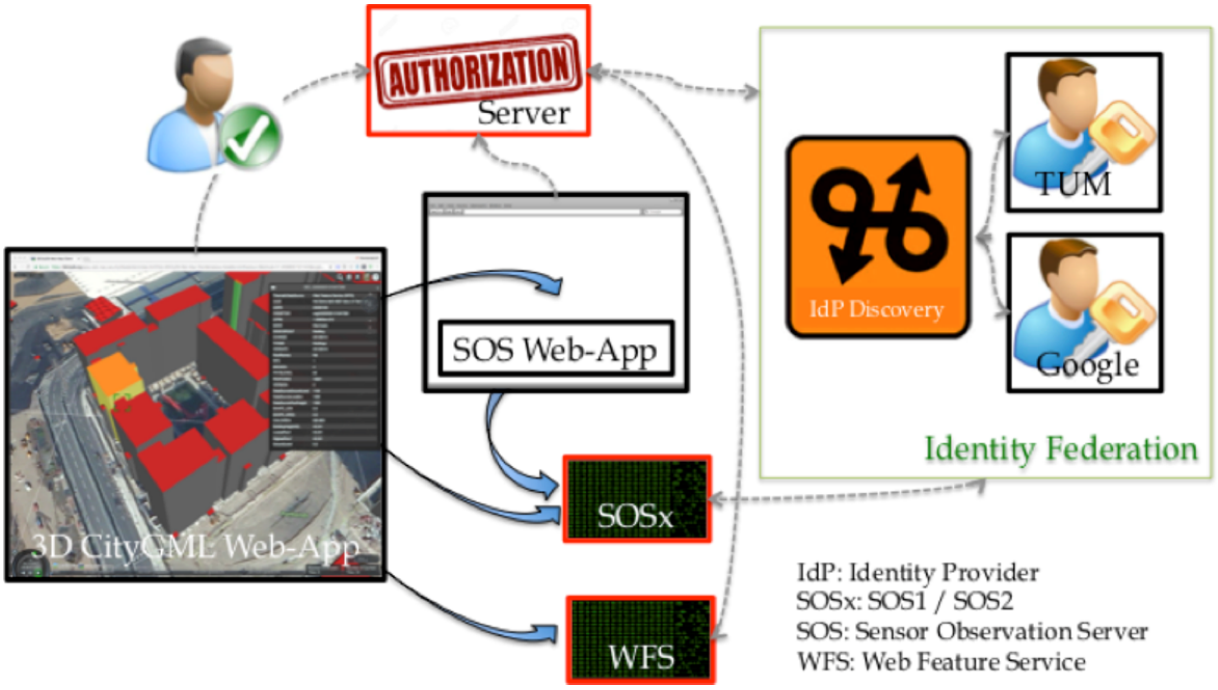


Abb. 1: Architektur Sketch des Security Demonstrator Projektes

Aus diesem Use Case ergeben sich zusammenfassend die folgenden Anforderungen, die bei der Integration einer Security-Lösung zu beachten sind:

1. Single-Sign-On: Wegen Benutzerfreundlichkeit muss ein einmaliges Login ausreichen
2. Föderation: Login via Google oder Technische Universität München
3. Unbeschränkter (anonymer) Zugriff auf die 3D CityGML Web-Applikation und das 3D Stadtmodell
4. Access Control: Benutzer mit Login Google können verlinkte Umwelt-Daten abrufen (im Web-Browser und in spezieller Web-Applikation)
5. Access Control: Benutzer mit Login TUM können verlinkte Umwelt-Daten und Smart-Meter-Daten abrufen (im Web-Browser und in spezieller Web-Applikation)
6. Web-Browser Same Origin Policy und W3C CORS

4.1 Umsetzung Föderation und Single-Sign-On

Die Umsetzung des Logins via Google oder TUM erfolgte durch eine SAML2 Föderation. In dieser Föderation sind die Identity Provider TUM und Google die beiden Entitäten, die das User Login bereitstellen. Beide unterstützen Single-Sign-On durch die Verwendung des SAML2 Session Initiators „PreviousSession“. Die Auswahl der Login-Entität wurde durch einen SAML2 kompatiblen Identity Provider Discovery Service ermöglicht. Dieser IdP DS oder auch WAYF (Where Are You From) bietet dem Benutzer neben der Auswahl des Login-Anbieters auch das persistente Speichern der Auswahl an. Dies erhöht die Benutzerfreundlichkeit, denn der Benutzer sieht die Login-Auswahl erst wieder, wenn die Vorauswahl aufgehoben wurde.

Dass Single-Sign-On via Google unterstützt wird, merkt der Benutzer dann, wenn er bereits bei Google angemeldet ist und die 3D-CityGML-Anwendung ebenfalls im Web-Browser ausgeführt wird. Klickt der Benutzer nun bei der 3D-CityGML-Anwendung auf Login und hat Google voreingestellt, so wird einfach dieses Login verwendet.

4.2 Umsetzung Zugriff auf Gebäude-Zusatzinformationen

Die Gebäude-Zusatzinformationen können per Klick auf das Gebäudeklötzchen von einem WFS abgerufen werden. Hierzu muss der Benutzer angemeldet sein. Die eigentliche WFS Abfrage wird von der 3D CityGML Applikation (einer Web-Applikation) beim Klick durchgeführt. Obwohl der Benutzer-Login durch SAML2 Identity Provider erfolgt, ist – wie zuvor erörtert - für diese Art des Zugriffs durch die Web-Applikation die Nutzung von SAML2 Sessions ungeeignet. Es ist somit erforderlich den Zugriff durch OAuth2 Access Tokens abzusichern.
Zur Nutzung von OAuth2 Access Tokens waren drei Dinge umzusetzen:

1) Zuerst wurde ein OAuth2 kompatibler Authorization Server installiert und die 3D- CityGML-Anwendung registriert. Das Login erfolgt über Google oder den TUM Identity Provider.
2) Der WFS mit den Gebäudezusatzinformationen wird als OAuth2 Resource Server betrieben.
3) Die 3D-City-GML-Anwendung wurde erweitert, um Access Tokens vom Authorization Server abzurufen und bei Anfragen an den WFS mitzuschicken.

4.3 Umsetzung Zugriff auf verlinkte Sensor-Daten

Der Zugriff auf die verlinkten Sensor-Daten zu den Gebäuden kann ebenfalls via Klick auf einen Link erfolgen. Hierbei werden die Ergebnisse der SOS Abfrage entweder direkt im Web-Browser angezeigt oder es wird eine geschützte Applikation geladen, die auf die SOS API zugreift. Hierbei ist die Verwendung von SAML2 Sessions bestens geeignet, denn damit kann der direkte Zugriff durch den Web-Browser erfolgen. Somit wurden bei SOS – für die Umwelt- sowie Smart-Meter Informationen – durch SAML2 Service Provider abgesichert.

4.4 Umsetzung von CORS

Damit die Auswirkung der Same Origin Policy des Web Browsers und dessen Lösung umgesetzt werden konnte, wird ein WFS-Proxy für die Gebäudezusatzinformation von Secure Dimensions betrieben. Somit ist jeder Zugriff von der 3D-CityGML-Anwendung auf den WFS eine Cross-Site Anfrage. Um die Funktionalität bzgl. CORS zu gewährleisten wurde der WFS entsprechend für Simple, wie auch Pre-Flight-Requests konfiguriert.

4.5 Demonstration

Das Ergebnis des Security Demonstrators ist online verfügbar. Als Einstieg muss die 3D-CityGML-Anwendung im Web-Browser geladen werden. Dazu kann die folgende URL genutzt werden: https://3dcitydb.org/qeop-web-map-security/3dwebclient/index.html

5 Fazit & Ausblick

Es ist möglich, eine moderne Geodateninfrastruktur mit interoperabler Sicherheit auf der Basis von existierenden OGC Web Service und Encoding Standards aufzubauen. Es wurde gezeigt wie durch die Kombination entsprechender Sicherheitsstandards – hauptsächlich SAML2 und OAuth2 – eine modulare und interoperable Sicherheitslösung für Authentication in eine existierende Geodateninfrastruktur integriert werden kann. Die Realisierung des Security-Demonstrators basiert im Wesentlichen auf Open Source Mainstream IT Softwarekomponenten mit denen jederzeit ein Produktivsystem aufgebaut werden könnte. Die Implementierung der SAML2 Komponenten basiert auf der Internet2 Software Shibboleth[4]; die Implementierung des OAuth2 Authorization Servers basiert auf der Implementierung von Brent Shaffer[5];
Obwohl die Implementierung auf HTTPS ein Erfolg war muss berücksichtigt werden, dass der Betrieb von OGC Web Services auf HTTPS streng genommen nach den existierenden OGC Standards nicht vorgesehen ist. Aber hierzu wird das OGC in Kürze einen Standard verabschieden, der diese Lücke schließt.
Die Ergebnisse dieses Beitrages haben aber ebenso gezeigt, dass eine moderne GDI neben der Umsetzung von üblichen Sicherheitsaspekten wie Authentication und Access Control vor allem darauf angewiesen ist, dass ebenso Mainstream IT Unterstützung sichergestellt wird. Insbesondere sei hier auf die Anforderungen durch Web-Anwendungen und die sich daraus ergebende Unterstützung wie CORS hingewiesen. Eine einfache Konfiguration, die „*“ für die erlaubte Origin zurückliefert, sollte nie verwendet werden, es sei denn man versteht genau welche Auswirkungen dies hat.
Im Hinblick auf die kommende GDPR gibt sicherlich noch viel zu tun; auch für Geodateninfrastrukturen mit Benutzerverwaltung. Wichtig erscheint jedoch das beschriebene Konzept, die Verwaltung von Benutzern und das Service-Angebot zu trennen, damit Service Anbieter nicht zu sehr mit der GDRP belastet werden. Stattdessen sollten qualifizierte Unternehmen die Benutzerkonten verwalten. Ein föderatives Identity Management System kann auch in GDIs durch die Verwendung von SAML2 erfolgen, wie es in akademischen Föderationen bereits tagtäglich genutzt wird.
Um den Zugriff auf geschützte Dienste durch Web-Applikationen und Desktop-Applikationen zu ermöglichen, kann OAuth2 mit Access Tokens verwendet werden. Beispielsweise hat Secure Dimensions in Kooperation mit der TUM im Rahmen des OGC Testbed 13 SAML2 und OAuth2 Plugins[6] für QGIS 2.18.x als Open Source entwickelt und bietet eine fertig erstellte DLL für Windows 10 an, die von der Testbed 13 Projektseite heruntergeladen[7] werden kann.
Die Umsetzung des Security Demonstrators im Smart Sustainable District Projekt hatte die Integration von Authentication mit einfacher Zugriffskontrolle zum Ziel. Somit ergibt sich als Ausblick die Integration einer Zugriffskontrolle, die nicht nur rollenbasierte sondern auch orts-

[4] https://www.internet2.edu/products-services/trust-identity/shibboleth/
[5] https://github.com/bshaffer/oauth2-server-php
[6] https://github.com/securedimensions
[7] https://www.tb13.secure-dimensions.de/

abhängige Rechte durchsetzt. So könnte beispielsweise der Ort des Benutzers zusätzliche Zugriffsrechte auf Gebäude-Zusatzinformationen ermöglichen.
Zum Schluss stellt sich die Frage in wie weit die Best-Practices-Ergebnisse dieses Beitrags auch auf andere Geodateninfrastrukturen wie z.B. INSPIRE übertagen werden können. Diese Frage kann nicht pauschal beantwortet werden, aber es steht zu hoffen, dass dieser Beitrag eine entsprechende Diskussion in Gang setzt. Dies erscheint insbesondere wünschenswert, da das Ergebnis des JRC Projektes „Authentication, Authorization and Accounting (AAA) for Data and Services in EU Public Administrations“ [JRC, Action Item 1.17, ISA Programme] bereits eine Access Management Federation auf der Basis von SAML2 empfiehlt.

6 Literaturverzeichnis

BRÖRING, A., STASCH, C. & ECHTERHOFF, J., 2012: OGC Sensor Observation Service Interface Standard, OGC #12-006.

CANTOR, S., KEM, J., PHILPOTT, R. & MALER, E., 2005: Assertion and Protocols for the OASIS Security Assertion Markup Language (SAML) V2.0.

CHATURVEDI, K. & KOLBE, T.H., 2016: Integrating Dynamic Data and Sensors with Semantic 3D City Models in the context of Smart Cities. Proceedings of the 11th International 3D Geoinfo Conference (ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences), ISPRS.

GRÖGER, G., KOLBE, T.H., NAGEL, C. & HÄFELE, K.-H., 2012: OGC City Geography Markup Language (CityGML) Encoding Standard, OGC #12-019

HARDT, D., 2012: The OAuth 2.0 Authorization Framework, IETF RFC 6749.

INITIAL OPERATING CAPABILITY TASKFORCE, 2013: Technical Guidance for the implementation of INSPIRE Download Services, Version 3.1.

INITIAL OPERATING CAPABILITY TASKFORCE, 2013: Technical Guidance for the implementation of INSPIRE View Services, Version 3.11.

ISO 7498-2, 1989: Information processing systems – Open Systems Interconnectin – Basic Reference Model – Part 2: Security.

ISO 10181-1, 1996: Information Technology – Open Systems Interconnectin – Security frameworks for open systems: Overview.

ISO 10181-2, 1996: Information Technology – Open Systems Interconnectin – Security frameworks for open systems: Authentication Framework.

ISO 10181-3, 1996: Information Technology – Open Systems Interconnectin – Security frameworks for open systems: Access Control Framework.

ISO 10181-4, 1997: Information Technology – Open Systems Interconnectin – Security frameworks for open systems: Non-Repudiation Framework.

ISO 10181-5, 1996: Information Technology – Open Systems Interconnectin – Security frameworks for open systems: Confidentiality Framework.

ISO 10181-6, 1996: Information Technology – Open Systems Interconnectin – Security frameworks for open systems: Integrity Framework.

ISO 10181-7, 1996: Information Technology – Open Systems Interconnectin – Security frameworks for open systems: Security audits and alarms.

JONES, M. & HARDT, D., 2012: The OAuth 2.0 Authorization Framework: Bearer Token Usage, IETF RFC 6750.

JONES, M. & HILDEBRAND, J., 2015: JSON Web Encryption (JWE), IETF RFC 7516.

JONES, M., BRADLEY, J. & SAKIMURA, N., 2015: JSON Web Token (JWT), IETF RFC 7519.

MATHEUS, A., 2008: Geospatial eXtensible Access Control Markup Language (GeoXACML) Version 1.0, OGC #11-017.

MATHEUS, A., 2015: OGC Testbed-11 Implementing Common Security Across the OGC Suite of Service Standards, OGC #15-022

MATHEUS, A., 2016: OGC Testbed-12 OWS Common Security Extension ER, OGC #16-048r1

MATHEUS, A., 2017: OGC Testbed-13 Security ER, OGC #17-021r2

MATHEUS, A., 2018: OGC Web Services Security Standard, OGC #17-007 (verfügbar voraussichtlich Juni 2018)

MOSHREFZADEH, M., CHATURVEDI, K., HIJAZI, I., DONAUBAUER, A. & KOLBE, T.H., 2017: Integrating and Managing the Information for Smart Sustainable Districts - The Smart District Data Infrastructure (SDDI). In: Kolbe, T.H., Bill, R. & Donaubauer, A. (Hrsg.): Geoinformationssysteme 2017 – Beiträge zur 4. Münchner GI-Runde. Wichmann Verlag.

MOSES, T., 2005: eXtensible Access Control Markup Language (XACML) Version 2.0.

RISSANEN, R., 2013: eXtensible Access Control Markup Language (XACML) Version 3.0.

RESCORLA, E., 2000: HTTP Over TLS, IETF RFC 2818.

SAKIMURA N., BRADLEY, J., JONES, M., DE MEDEIROS, B. & MORTIMORE, C., 2014: OpenID Connect Core 1.0 incorporating errata set 1

VANDENBROUCKE, D., MATHEUS, A., FRIGNE, D., DE GRAEF, P., COPIER, R. & SMITH, R., 2014: Authentication, Authorization & Accounting for Data and Services in EU Public Administrations, D4.1.5 – Final technical report.

VRETANOS, P., 2010: OpenGIS Web Feature Service 2.0 Interface Standard, OGC #09-025r1.

W3C RECOMMENDATION, 2014: Cross-Origin Resource Sharing.

Langzeitspeicherung von Geodaten – eine Lösung in Sicht!

DORIS CARSTENSEN[1]

Zusammenfassung: Geodaten gelten ebenso wie Schriften oder Kunstwerke als Kulturgut, die späteren Generationen zur Verwendung übergeben werden müssen. Diese Vereinbarung wird für Geodaten in analoger Form eingehalten. Das Verfahren ist in den Archivgesetzen der Länder geregelt. Soll dies zukünftig auch für digitale Geodaten gelten, sind einige Anstrengungen notwendig. Im besonderen Fokus stehen hier der Nachweis der Originalität der Geodaten, nachhaltige Dateiformate, zugehörige Datenmodellbeschreibungen, die stetig anwachsende Datenmenge und der physikalische Speicher. Nur wenn alles übergreifend mit einem Verfahren organisiert, gepflegt und zur Benutzung und Abgabe an die Archive verwaltet wird, besteht die Möglichkeit, dass die digitalen Geodaten auf Jahrzehnte für die Nutzer bereitgestellt und auch interpretiert werden können.

1 Geodaten als Kulturgut

Abb.1: LZS-Logo, gesichert für den LGV beim DPMA

Geodaten sind aus der heutigen Welt nicht mehr wegzudenken. Bei Fragestellungen z. B. zur Raumordnung und Planung stehen neben der Nutzung hochaktueller, digitaler oft zusätzlich auch historische Geodaten im Fokus, die seit vielen Jahren zunehmend digital vorliegen. Metadatenkataloge der Geodateninfrastrukturen (GDI) helfen bei der Recherche verwendbarer Daten nicht immer weiter, weil sie vorwiegend aktuelle Geodaten nachweisen und den Zugriff darauf organisieren. Die Archive, als Hort historischer Daten, weisen bisher meist analoge Geodaten nach und dann nicht unbedingt flächendeckend. Diese Defizite sind bei den Geodatenproduzenten bereits seit Jahren bekannt und werden durch die steigende Nachfrage von Nutzern nach historischen, digitalen Geodaten immer präsenter. Der Landesbetrieb Geoinformation und Vermessung Hamburg hat sich seit 2014 mit der Problematik intensiv beschäftigt und war bereit, ein System zur *Langzeitspeicherung von digitalen Geodaten* (LZS) zu beschaffen. Nach Erstellung eines Konzeptes (CARSTENSEN et al. 2015) für dieses Vorhaben mit anschließender Marktanalyse wurde kein marktreifes Produkt für die sichere Speicherung digitaler Geodaten für einen Zeitraum von mehr als 50 Jahren gefunden. Schriftgutverwaltungen sind wegen der Besonderheit von Geodaten keine Alternative. Die Entscheidung, ein passendes Produkt erstellen zu lassen, war folgerichtig, ebenso die Bereitschaft, die Anwendung für andere Geodaten-produzierende Stellen zu öffnen. 2016 folgte ein öffentlicher Teilnahmewettbewerb mit anschließender freihändiger Vergabe, um eine Anwendung zur Langzeitspeicherung von Geodaten erstellen zu lassen. Den Zuschlag erhielt die AED-SICAD AG, die sich für die Erstellung dieser Anwendung mit der SER Group einen star-

[1] Landesbetrieb Geoinformation und Vermessung Hamburg, Fachbereich Geobasisdaten, Neuenfelder Straße 19, D-21109 Hamburg, E-Mail: doris.carstensen@gv.hamburg.de

ken Partner aus der Archivwelt gesucht hat. Der Landesbetrieb Geoinformation und Vermessung Hamburg, die AED-SICAD AG und die SER Group haben zur Entwicklung und Weiterentwicklung der Anwendung eine Implementierungspartnerschaft gegründet.

2 Geodaten im Laufe der Zeit

Digitale Geobasisdaten werden seit Mitte der 70er Jahre in der amtlichen Vermessungsverwaltung Hamburg, wie im ganzen Bundesgebiet, produziert. Speziell bei den alpha-nummerischen Daten des Liegenschaftskatasters gab es seitdem eine Reihe von Verfahren und Datenmodellen. Das letzte in dieser Reihe war das Hamburgische Automatisierte Liegenschaftsbuch (HALB), welches 2010 in das gemeinsame 3A-Modell der Arbeitsgemeinschaft der Vermessungsverwaltungen des Bundes und der Länder (AdV) für Liegenschaftskataster, Topographie und Raumbezug migriert wurde. Auch für den Kartenteil des Liegenschaftskatasters gab es vor dem heutigen ALKIS® bereits einen digitalen Vorläufer, die Digitale Stadtgrundkarte (DSGK). Das aktuelle Portfolio der Geobasisdaten umfasst die amtlichen Daten des Liegenschaftskatasters (ALKIS®), der Geotopografie (ATKIS®) und des Raumbezuges (AFIS®), die Luftbilddaten, die Digitalen Orthofotos (DOP), das Digitale Geländemodell (DGM) und viele weitere digitale Geodaten.
Einige der Geodaten sind „lebende“ Datenbestände, die täglich aktualisiert werden. Hier kommt im Hinblick auf die Langzeitspeicherung noch die Notwendigkeit der Bildung von regelmäßigen Zeitschnitten hinzu. Geodaten, wie die DOP, bilden vom Grundsatz her einen Zeitschnitt ab.
Allen Geodaten gemein ist jedoch der ständige technologische Wandel. Die Daten werden mit regelmäßig neuen, leistungsfähigeren Softwaresystemen in immer komplexeren Dateiformaten und Datenmodellierungen erzeugt. Die Datenmenge wächst kontinuierlich. Bei der Speicherung der Geodaten wird je nach Nutzungsfrequenz und Datenmenge von externen (Offline-) Speichermedien, Fileservern oder Storagesystemen mit eigener Verwaltungslogik Gebrauch gemacht. Den Überblick über die vorhandenen Geodaten zu behalten wird, in Anbetracht der Menge und Zeiträume, immer schwieriger, ebenso wie das Vertrauen in die korrekte Interpretation.

2.1 Geodaten in den Archiven

Erst nach Abschluss der Arbeiten, z. B. bei Einzelprojekten oder nach der Bildung von Zeitschnitten bei lebenden Daten, sind Geodaten für die Archive von Interesse. Die Grundsätze, nach denen zu entscheiden ist, welche digitalen Geodaten in einer historischen Version aus diesem riesigen Vorrat als Kulturgut für die Archive von Bedeutung sind und an diese abgegeben werden müssen, ist sowohl den Archivgesetzen der Länder (HMBARCHG 1991) als auch in den Leitlinien zur bundesweit einheitlichen Archivierung von Geobasisdaten (ADV-KLA-ARBEITSGRUPPE 2015) zu entnehmen.
Der Weg von Geodaten nach Abschluss der Produktion bis hinein ins Archiv wird mit dem Begriff Aussonderung in Verbindung gebracht. Die Archive sind keineswegs verpflichtet, komplette, flächendeckende Datensätze in das Archiv zu übernehmen. Die Entscheidung zur Archivwürdigkeit trifft das Archiv, nach dem Prinzip „Anbietung-Bewertung-Entscheidung“ auf der Grundlage der Gesetze in eigener Verantwortung. Daher können angebotene Geodaten komplett, in Teilen oder gar nicht in das Archiv übernommen werden (Kassationsprinzip). Ein digitales Archiv wird allein aufgrund der Datenmengen die Anforderungen nach einer flächendecken-

den, zeitlich definierten und mit Raumbezug recherchierbaren Geodatenhaltung nicht erfüllen können.
Weil Geodaten-Produzenten aber über die Archivwürdigkeit hinausgehende Anforderungen zum Speicherumfang haben könnten, gilt für alle digitalen Geodaten: Sie können, müssen aber nach einer Aussonderung/Abgabe an das Archiv nicht aus der LZS gelöscht werden. Das wäre z.B. der Fall, wenn die Archive nur eine Stichprobe (einzelne Dateien) übernehmen wollen, die Geodaten-Produzenten aber einen flächendeckenden, historischen Datensatz für sinnvoll halten.
Die Abgabe und Aufnahme aller digitalen Geodaten in die Archive ist bis heute noch nicht möglich. Wenn die Digitalen Magazine der Archive in naher Zukunft betriebsfähig sind, werden die digitalen Geodaten nur noch nach einem festgelegten Verfahren unter Berücksichtigung des OAIS-Referenzmodells (OAIS 2012) entgegengenommen und verwaltet. Im Aussonderungsprozess vom Produzenten hin zum Archiv müssen s.g. Submission Information Packages (SIP) (OAIS 2012) gebildet und gesichert übertragen werden. Bisher bildet keine andere Datenhaltung diesen Vorgang sicher und OAIS-konform ab, die Langzeitspeicherung als Anwendung wird auch diese Lücke schließen.

2.2 Geodateninfrastruktur oder Langzeitspeicherung?

Die Abgrenzung zwischen einer Geodateninfrastruktur (GDI) und einer Langzeitspeicherung ist schwieriger zu klären, als zwischen der Langzeitspeicherung und dem digitalem Archiv. Festgeschrieben ist nach §3(1) HMBARCHG (1991), dass „… alle Unterlagen, die sie zur Erfüllung der Aufgabe nicht mehr benötigen, fortlaufend auszusondern, dem Staatsarchiv anzubieten und ihm nach Feststellung der Archivwürdigkeit abzuliefern. Unterlagen sollen spätestens 30 Jahre nach ihrer endgültigen Entstehung ausgesondert und angeboten werden, soweit sie nicht noch nachweislich im Geschäftsgang erforderlich sind …“.
Damit und den Leitlinien (ADV-KLA-ARBEITSGRUPPE 2015) ist bestimmt, aus welchem Geodatenschatz das Archiv archivwürdige Geodaten auswählen muss. Nicht eindeutig bestimmt ist, wann genau die Geodaten dem Archiv anzubieten sind: „Wann ist die Aufgabe erfüllt? Wann sind sie endgültig entstanden?“. Die Grenzen sind verschwommen, viele digitale Daten werden kontinuierlich aktualisiert und könnten praktisch so lange genutzt werden, wie das Dateiformat noch lesbar ist. Hier können Fragestellungen zu Art der Geodaten (Zeitschnitt oder lebender Datensatz), der Nutzungsfrequenz historischer Geodaten, deren Lesbarkeit, die Interpretationsfähigkeit, der Bestandserhaltung, der Performance im Zugriff und den Speicherkosten wichtige Entscheidungskriterien liefern. Letztlich trifft bei vielen Geodaten der Produzent die Entscheidung, wann ein Geodatensatz aus der GDI in die Langzeitspeicherung verlagert wird, von der LZS in das Archiv ist der Ablauf definiert.
Wichtig bei einer Bereitstellung von Geodaten in zwei Systemen ist, dass dies transparent für den Nutzer erkennbar ist. Hochaktuell und performant über Webdienste wird der Nutzer in einer GDI mit Geodaten versorgt und weniger performant, trotzdem georeferenziert auffindbar und auf jeden Fall in einem lesbaren Format soll das auch viele Jahrzehnte nach der Produktion der Geodaten in der Langzeitspeicherung möglich sein.

3 Anforderungen an die Langzeitspeicherung

Die Anwendung wird geprägt durch die vier Kernforderungen:

- Nachhaltigkeit
- Migration
- Sicherheit
- Schnittstellen

3.1 Nachhaltigkeit

Nachhaltigkeit soll bei dem Verfahren Langzeitspeicherung unter zwei unterschiedlichen Aspekten gewährleistet werden: Erfüllung technischer Anforderungen aber auch organisatorische Unterstützung.

Über einen Zeitraum von mehr als 50 Jahren Nachhaltigkeit in der Nutzung – im Sinne von Lesbarkeit – der Geodaten zu garantieren, ist eine für digitale Informationen extrem anspruchsvolle Forderung. Lesbarkeit umfasst technische Nutzbarkeit, inhaltliche Interpretationsfähigkeit und eine Recherchemöglichkeit.

Mit Lesbarkeit ist nicht unmittelbar die digitale Weiterverarbeitung oder direkte Einbeziehung in Web-Dienste gemeint sondern die Möglichkeit, den Inhalt der Dateien interpretieren und nutzen zu können. Dieser Nutzungslevel ist eine Konzession an die Forderung, nur nachhaltige Dateiformate gemäß den KLA/AdV-Leitlinien (ADV-KLA-ARBEITSGRUPPE 2015) in einer Langzeitspeicherung zuzulassen. Dazu gehören neben PDF/A z. B. XML, NAS, GML, TIFF, JPG2000, CSV. Aktuell begehrte Dateiformate wie z.B. DXF, PNG, GIF, XLSX oder BMP fallen damit aus. Die Feststellung, ob Dateiformate als nachhaltig gelten, mag sich im Laufe der Zeit wandeln aber auch hierauf muss eine Langzeitspeicherung reagieren können, mit rechtzeitig angezeigtem Migrationsbedarf.

Aus Sicht der Interessenten beschränkt die restriktive Haltung bzgl. zulässiger Dateiformate das spontane Nutzen der digitalen Geodaten und erfordert u. U. vermehrte Aufwände, um die Daten in aktuellen Projekten verwenden zu können. Vor dem Hintergrund, dass historische Geodaten ansonsten verloren wären und gar nicht einer weiteren Verwendung zugeführt werden könnten, ist das notgedrungen zu akzeptieren.

Die Interpretationsfähigkeit der Geodaten wird wesentlich durch die gleichzeitige Bereitstellung von zusätzlichen Informationen, wie Datenmodellbeschreibungen, Raumbezügen, Signaturen-Bibliotheken und weiteren Katalogdaten bestimmt. Diese im Archivbereich als Dossiers bekannten Informationen werden als eigene LZS-Objekte betrachtet, in der Langzeitspeicherung abgelegt und mit den eigentlichen Geodaten verlinkt.

Metadaten bilden die Grundlage für erfolgreiche Suche nach relevanten Geodaten in der Langzeitspeicherung. Aktuelle, zu den Produktionsdaten bereits erfasste Metadaten aus der GDI sollen in die Langzeitspeicherung importiert und um Informationen zum jeweiligen Datensatz, z. B. ein Zeitschnitt, ergänzt werden.

Der andere Beitrag zur Nachhaltigkeit wird durch eine organisatorische Unterstützung beim Geodatenproduzenten nach Abschluss der Arbeiten geleistet. Schon bei der Geodaten-Anlieferung, zu Beginn des Langzeitspeicherprozesses, werden vom Produzenten Angaben u.a. zur Anbietungspflicht gegenüber dem Archiv, einer Aufbewahrungszeit und den Aussonde-

rungsbedingungen verlangt und zu den LZS-Geodaten als LZS-Metadaten gespeichert. So können LZS-Aktivitäten Datums-, Anlass- und Datenklassengesteuert definiert und im LZS-Verwaltungsteil als Routine abgespeichert werden. Dies gewährleistet, dass auch bei langen Zeitabständen immer nach dem gleichen Prinzip die Geodaten hinsichtlich anstehender Aussonderungen, Löschmöglichkeiten oder Konvertierungen selektiert, protokolliert und verarbeitet werden.

3.2 Migration

Migration ist in allen Teilen des Verfahrens von zentraler Bedeutung. Sie ist erforderlich, wenn Dateien nicht im zulässigen Dateiformat für die Langzeitspeicherung vom Geodatenproduzenten „angeliefert“ werden. Sie ist erforderlich, wenn im Laufe der Zeit neue Dateiformate als nachhaltiger eingestuft werden und bereits eingelagerte Dateien in diese überführt werden sollen. Sie ist erforderlich für die regelmäßige Anpassung der Storagesysteme, auf denen die Geodaten lagern. Und sie ist auch erforderlich für alle eingesetzten Soft- und Hardwarekomponenten, wie z. B. Betriebssysteme.
Die unterschiedlichen Migrationen sind als Module/Prozesse in der LZS-Anwendung berücksichtigt und wegen der weitreichenden Veränderungsmöglichkeiten im LZS-Datenbestand ausschließlich der Fachadministration vorbehalten. Insbesondere die Migration von Dateiformaten und beim Storage ist parallel zum laufenden LZS-Betrieb vorgesehen und damit ein wesentlicher Anteil der Bestandserhaltungsmaßnahmen.

3.3 Sicherheit

Die Feststellung und Bestätigung der Originalität der langzeitgespeicherten Geodaten ist - besonders zum Produktionsbeginn einer LZS - eine der großen Herausforderungen. Im Nachgang zu einer (erstmaligen) Bestandsaufnahme vorhandener digitaler Geodaten, die in der Verantwortung der Geodatenproduzenten auf diversen Storagesystemen lagern, müssen Kopien und Derivate von den Originaldateien unterschieden werden. Wenn beim Geodatenproduzenten bisher für die Speicherung keine konsequente Kontrolle durch eine zentrale Stelle nach festgelegten Regeln verantwortlich war, ist die Bestimmung der Originaldatei eine schwierige und verantwortungsvolle Aufgabe. Oft ist auch und die ehemals verantwortliche Stelle nicht mehr ermittelbar und Nachfragen damit unmöglich. Vor dem Hintergrund der Datenhygiene und der Speicherkontrolle wird man eine Entscheidung für den wahrscheinlich originalen Datensatz treffen müssen, der dann der Langzeitspeicherung angeboten wird. Kopien und Derivate sind danach unbedingt zu löschen. Ist eine LZS erst einmal in Betrieb, müssen organisatorische Vorgaben dafür sorgen, dass sich kein Stau im Einlagerungsprozess mehr entwickeln kann.
Der Schutz vor physikalischer Zerstörung und ein Zugriffsmanagement sind weitere Anforderungen im Bereich Sicherheit. Das erste wird durch den konsequenten Verzicht auf externe Speichermedien und den Betrieb in einem Rechenzentrum leicht aber durchaus kostenbewusst umgesetzt. Der Zugriffschutz in Form eines Rechtemanagements ist Teil der LZS-Anwendung und im User- and Ressourcemanagement (URM) verankert, konfigurierbar und berücksichtigt die Sicherheitsbestimmungen nach BDSG (2003). Hierbei wird nach Fachadministratoren, den Datenverantwortlichen, die Geodaten anliefern sollen und den Nutzern unterschieden. Die Geodaten

sind unterschiedlichen Schutzklassen, je nach Schutzbedürfnis z. B. bei Personenbezug, zuzuordnen.
Und zudem werden alle Aktivitäten in der LZS protokolliert und wesentliche Aktionen, wie Konvertieren oder Löschen, unterliegen dem 4-Augen-Prinzip. Wenn alle Vorgaben eingehalten wurden, gibt es nur noch das LZS-Original, dem maximalen Schutz zu gewähren ist.

3.4 Schnittstellen

Sowohl zur Sicherheit der Geodaten in einer Langzeitspeicherung als auch zur Schnittstellenthematik gehören Mechanismen der sicheren Datenübertragung bei der Einlagerung (Ingest) und – wenn vereinbart – der Abgabe an die Archive. Grundsätzlich werden mittels LZS-Modul angelieferte Geodaten auf zugelassene und tatsächliche Dateiformate überprüft und ggf. eine Migration empfohlen oder eine Einlagerung abgelehnt. Die dann als „zugelassen“ bewerteten Dateien werden aus einem Anlieferungsbereich in Submission Information Packages (SIP) verpackt (OAIS, 2012) und mit „handshake-Verfahren“ und nach HASH-Wertüberprüfung in die Langzeitspeicherung abgesenkt (Ingest) und auf vergleichbaren Wege an die Archive abgegeben (Aussonderung). Die nach OAIS-Referenzmodell (OAIS 2012) definierten Regelungen werden eingehalten.
Die Fachadministration hat einen besonderen mit vielen Rechten ausgestatteten Client als Zugang zur Langzeitspeicherung (Management). Mit reduziertem Funktions- und Rechteumfang gilt dies auch für die Datenverantwortlichen, die die Verbindung zwischen Geodatenproduktion und Langzeitspeicherung bilden.
Die Langzeitspeicherung kann für die Nutzung der Geodaten (Access) flexibel entweder nur über den eigenen LZS-Recherche-Client, über die Einbindung in andere Web-Auskunftsverfahren oder direkt im 3A-Umfeld eingebunden werden. Insbesondere die Daten des Hamburger Grenznachweises, die anders als die übrigen historischen Geodaten größtenteils noch für die tägliche Arbeit im Liegenschaftskataster benötigt werden, sind in das Auskunftssystem 3A Web ARCHIV von AED-SICAD eingebunden und müssen nach den Regeln des HmbArchG (HMBARCHG 1991) auch nicht ausgesondert werden.

4 Anwendung zur Langzeitspeicherung

4.1 Systemaufbau

Verwaltungskomponente und Archivierungskomponente sind zugleich die Softwareprodukte und die beiden ersten Ebenen der Anwendung, die als dritte Ebene den Storage anspricht. Die gesamte Anwendung verantwortet AED-SICAD als Auftragnehmer und stellt mit der LZS-Verwaltungskomponente (rote Felder in Abb. 2) den organisatorischen Teil und die Verbindung zur georeferenzierten Suche der Anwendung. Sie fügt sich nahtlos in das URM des GIS Portals von AED-SICAD ein und ist zugleich die Grundlage für die Auskunft des Grenznachweises via 3A Web ARCHIV. Zur Administration und Spezialsuche existiert zusätzlich ein eigenständiger LZS-Client. Einige Funktionen der Archivierungskomponente (blau markierte Felder in Abb. 2), als zweites Softwareprodukt, werden über die Verwaltungskomponente angesprochen, wie die Formaterkennung und -validierung oder die Bestandserhaltung.

Die Archivierungskomponente wird durch die Software Doxis4 der SER Group abgebildet. Über Adapter können mehrere Storage-Lösungen wie z.B. Filesystem, Doxis4 SafeLock, NetApp, EMC^2 angesprochen werden. Geodaten können über die Archivierungskomponente je nach ihrer Klassifizierung verschiedenen Storage-Bereichen zugeordnet werden. Das unterstützt nicht nur ein performantes Antwort-Zeitverhalten bei häufig nachgefragten Daten, sondern berücksichtigt im Gegenzug die Nutzung von kostengünstigerem, weniger performantem Speicher für selten abgefragte Geodaten. Gleichzeitig sind unter Betriebslast Umlagerungen von Geodaten in andere Storage-Qualitäten oder zu anderen Storage-Anbietern denkbar.

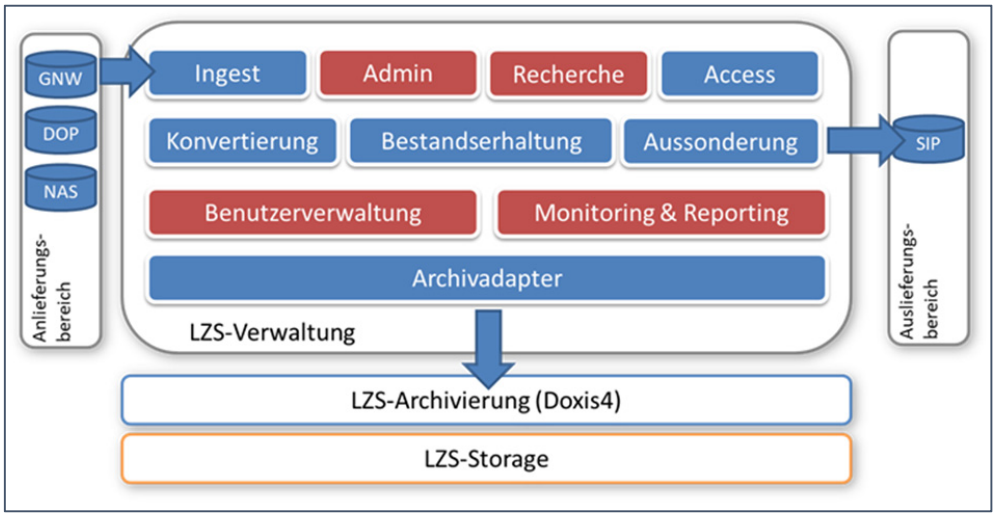


Abb. 2: Anwendungsskizze (LZS-Spezifikation von AED-SICAD, SER GROUP (2017))

4.2 Anwendungsszenarien

Konzeptionell ist die LZS-Anwendung sowohl im 3A-Umfeld des GIS Portals von AED-SICAD, z. B. mit einer Verbindung zum Auskunftssystem 3A Web und 3A Web ARCHIV, als auch als eigenständige Variante einsetzbar. Die erste Variante nutzt eine bestehende 3A-Auskunftslösung vollständig mit und bietet damit gleichzeitig den Vorteil der umfänglichen, auch georeferenzierten Suche in einem aktuellen Geodatenbestand, z. B. des Liegenschaftskatasters oder der Geotopographie, um relevante LZS-Objekte zu ermitteln.

In den Fällen, wo das 3A-Umfeld nicht zur Verfügung steht, ist die Metadaten-basierte Recherche über den LZS-Auskunftsclient möglich, mit oder ohne Einbeziehung eines eigenen Kartenportals für eine georeferenzierte Suche. In diesen Fällen muss ein GIS Portal mit URM (AED-SICAD) installiert werden, jedoch ohne 3A-Daten oder die 3A-Datenhaltung (DHK als 3A Server). Hier sind verschiedene Szenarien denkbar, die im Einzelfall auf Realisierung zu prüfen wären.

Will und kann ein Geodatenproduzent die LZS-Anwendung nicht selber betreiben, weil die Geodatenmenge nicht ausreichend erscheint oder kein Personal für die Aufgabe zur Verfügung steht, so ist es auch denkbar, eigene Geodaten zur Langzeitspeicherung an einen Betreiber, z. B. die Kataster- und Vermessungsverwaltung, Landesamt/-betrieb für Geoinformation, wo die Anwendung einsetzt wird, abzugeben. Das Rechtekonzept der Anwendung ist so konzipiert, dass die

unterschiedlichen Dateneigentümer in der LZS-Verwaltung abgebildet werden können und ihnen eigene Rechte eingeräumt werden können.
Die Verwendung von Storage ist sehr flexibel und im Einzelfall hinsichtlich der Bedingungen zu prüfen. Für alle Anwendungsszenarien wird es je nach Betreiber verschiedene Preis- und Kostenmodelle geben müssen.

4.3 Aktueller Verfahrensstand

Die Implementierung der LZS-Anwendung ist entsprechend des Konzeptes (CARSTENSEN et al. 2015) in Stufen unterteilt und berücksichtigt dabei die verschiedenen Geodatenklassen. Den Start bildet die Stufe 1 mit der komplexen Geodatenklasse des Grenznachweises. Danach folgt die Stufe 2 für die Geodatenklasse der Digitalen Orthofotos. In Stufe 3 sind sehr unterschiedliche Geodatenklassen zusammengefasst. Den Abschluss bildet mit Stufe 4 die Aussonderung und Abgabe an die Archive und berücksichtigt damit gleichzeitig die Umsetzungsaktivitäten der Archive um den Aufbau eines digitalen Archivs.
Nach Erteilung des Zuschlages als Ergebnis der Freihändigen Vergabe wurde Anfang 2017 auf Basis des vorliegenden Konzeptes des Landesbetriebes Geoinformation und Vermessung Hamburg mit den Arbeiten zur Feinspezifikation (AED-SICAD, SER GROUP 2017) begonnen. Diese wurde im Juni 2017 als Stufe 1 für die übergeordneten Funktionen und die fachspezifischen Aspekte der Geodatenklasse des Grenznachweises finalisiert und seitdem laufen die Entwicklungsarbeiten. Die Pilotierungsumgebung wird bei AED-SICAD betrieben und ein erster Prototyp konnte zur Infoveranstaltung am 22.11.2017 in Hamburg gezeigt werden. Die produktive Inbetriebnahme dieser ersten Realisierungsstufe ist beim IT-Dienstleister der Freien und Hansestadt Hamburg bis zur Jahresmitte 2018 geplant.

5 Fazit & Ausblick

Mit der Anwendung zur Langzeitspeicherung von Geodaten kommen die Geodatenproduzierenden Stellen ihrer Verantwortung nach, Geodaten als Kulturgut – analog und digital – nicht nur für sich selbst aufzubewahren sondern an nächste Generationen nachhaltig nutzbar weiterzugeben. Die Anforderungen an die Speicherung und Aussonderung (Abgabe an und Übernahme in das Archiv) werden nach einem sicheren Verfahren (OAIS-konform) bedient (OAIS 2012). Organisatorische Regelungen und definierte Prozesse, die weit vor der eigentlichen Langzeitspeicherung einsetzen bis zur tatsächlichen Ablage der originalen, unverfälschten Geodaten in der Langzeitspeicherung unterstützen alle Beteiligten. Und zuletzt eröffnet das LZS-Verfahren den Geodatenproduzenten langfristig eine sichere aber dennoch insgesamt kostenbewusste Speicherung der Geodaten.
Nach den vielen Kontakten im Zusammenhang mit der Erstellung des Konzeptes, vieler kleiner und einer großen Informationsveranstaltungen, wird immer deutlicher, dass die hier konzipierte Anwendung zur Langzeitspeicherung von Geodaten nicht auf den Geobasisdatensektor beschränkt bleiben muss. Geofachdatenproduzenten, wie z. B. die Geologen oder auch die Verkehrsbetriebe haben sehr ähnliche Anforderungen und vergleichbares Datenmaterial aber bis heute auch noch keine sichere Lösung, das „digitale Vergessen“ zu verhindern.

6 Literaturverzeichnis

AdV-KLA-Arbeitsgruppe, 2015: Leitlinien zur bundesweit einheitlichen Archivierung von Geobasisdaten. Abschlussbericht der gemeinsamen AdV-KLA-Arbeitsgruppe, 2014-2015. http://www.bundesarchiv.de/DE/Content/Downloads/KLA/leitlinien-geobasisdaten.pdf?__blob=publicationFile, letzter Zugriff 13.02.18.

AED-SICAD, SER Group, 2017: Langzeitspeicherung – Feinspezifikation. Unveröffentlicht.

BDSG, 2003: Bundesdatenschutzgesetz, Anlage zu §9 Abs. 1, BGBl, I:88.

Carstensen, D., Linke, J., Redel, T., Hafke, C. & Wriede, T., 2015: Konzept zur Langzeitspeicherung digitaler Geodaten, Landesbetrieb Geoinformation und Vermessung Hamburg. Unveröffentlichter Bericht.

HmbArchG, 1991: Hamburgisches Archivgesetz. 21.01.1991.

Keitel, C., 2010: Digitale Archivierung beim Landesarchiv Baden-Württemberg. Archivar, **63**(1).

Ludwig, J., 2008: Wege ins Archiv - Ein Leitfaden für die Informationsübernahme in das digitale Langzeitarchiv. Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit Digitaler Ressourcen für Deutschland.

OAIS, 2012: Open Archival Information System (Offenes Archiv-Informations-System), Referenzmodell, nach ISO 14721:2012.

Rönsdorf, C., Mason, P., Gerber, U., Bos, M., Shaon, A., Naumann, K., Kirstein, M., Samuelsson, G., Rantala, M., Kvarteig, S. & Stößel, W., 2012: GI+100: Langzeiterhaltung digitaler geographischer Informationen – 16 grundlegende, von staatlichen Vermessungsbehörden und Archiven vereinbarte Prinzipien. www.jiscmail.ac.uk%2Fcgi-bin%2Fwebadmin%3FA3%3Dind1204%26L%3DEUROSDR-GEODATA-ARCHIVING%26E%3Dbase64%26P%3D14733%26B%3D------%253D_NextPart_000_0010_01CD17B6.F6A9EAF0%26T%3Dapplication%252Fpdf%3B%2520name%3D%2522Paper_EuroSDR%2520Archiving%2520Principles_DE.pdf%2522%26N%3DPaper_EuroSDR%2520Archiving%2520Principles_DE.pdf%26attachment%3Dq%26XSS%3D3&usg=AOvVaw1MdcXPO_lTX_rc7-ZQ3moO, letzter Zugriff 13.02.18.

Leitfaden Geodäsie und BIM

ANDREAS DONAUBAUER[1] & THOMAS H. KOLBE[1]

Zusammenfassung: Im Zuge der Digitalisierung des Planens und Bauens wird Building Information Modeling (BIM) gerade als neue Methode etabliert. Der Runde Tisch GIS e.V. und die Gesellschaft für Geodäsie, Geoinformation und Landmanagement (DVW) nehmen sich des Themas aus der Perspektive der Geodäsie und Geoinformatik mit dem Leitfaden „Geodäsie und BIM" an. Mit Bezug zu diesem Leitfaden beleuchtet der Vortrag die Rolle der Geoinformation in BIM. Der Schwerpunkt liegt dabei auf der Integration von digitalen Modellen der bebauten Umwelt aus den Domänen Architektur und Bauwesen (AEC) und Geoinformatik (GIS). Neben einer Darstellung der Gemeinsamkeiten und Unterschiede der Modelle aus den unterschiedlichen Fachdisziplinen werden technische Möglichkeiten der Integration aufgezeigt.

1 Motivation

Building Information Modeling wird im Zuge der Digitalisierung des Planens und Bauens gerade als neue Methode etabliert. Wie in anderen Bereichen der Digitalisierung auch, wird dies zu grundlegenden Änderungen an bestehenden Vorgaben, Prozessen und Arbeitsweisen in der Planung, Architektur und dem Bauwesen führen. Zentraler Bestandteil von BIM ist die Erstellung und Führung einer digitalen Bauakte, die die verschiedenen Entwurfsstände, die Bauprozessplanung und nicht zuletzt die Dokumentation des erstellten Bauwerks in menschen- und maschineninterpretierbarer Form umfasst.

Die Geodäsie spielt beim Thema BIM überall dort eine Rolle, wo das geplante Bauwerk auf die reale Welt trifft, also zum Beispiel dort, wo

- die BIM-Planung durch Absteckung auf die Baustelle übertragen und dort realisiert werden soll;
- die Konstruktion im Sinne einer Baufortschrittskontrolle und der Bau- und Schadensdoku-mentation überprüft werden muss („As-Built"-Dokumentation) bzw. eine kontinuierliche Überwachung bestehender Bauwerke erforderlich ist (Bauwerksmonitoring);
- BIM-Modelle für schon bestehende Bauwerke erzeugt werden sollen, zum Beispiel für das Planen im Bestand („BIM-as-Built");
- Infrastrukturbauten mit großer Ausdehnung wie zum Beispiel Brücken, Tunnel, Talsperren und Straßen modelliert werden müssen, bei denen die Auswirkungen der Erdkrümmung im digitalen Modell nicht mehr vernachlässigt werden können;
- BIM-Modelle in 3D-Stadtmodelle zusammen mit weiteren Geobasis- und Geofachdaten zum Zwecke von Simulationen und Visualisierungen integriert werden sollen.

[1] Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstr. 21, 80333 München, E-Mail: andreas.donaubauer@tum.de

2 Der Leitfaden Geodäsie und BIM

Vor dem Hintergrund der vielfältigen Bezüge zwischen Geodäsie, Geoinformatik und BIM haben sich der DVW – Gesellschaft für Geodäsie, Geoinformation und Landmanagement und der Runde Tisch GIS e.V. dazu entschlossen, gemeinsam den Leitfaden Geodäsie und BIM (KADEN et al. 2017) herauszugeben.
Der im September 2017 online[2] publizierte Leitfaden gliedert sich in folgende Hauptteile:

- Im Grundlagenkapitel werden die wesentlichen Voraussetzungen zum Verständnis der BIM-Methode und der Zusammenhänge zwischen BIM, Geodäsie und Geoinformatik gelegt. Neben einer Beschreibung der BIM-Methode im Hochbau und im Infrastrukturbau werden technische, rechtliche und prozessorientierte Aspekte der BIM-Methode beschrieben und die Bezüge zu Geoinformationssystemen (GIS) und Ingenieurgeodäsie hergestellt.
- Im Kapitel „Praxisbeispiele“ geben Autoren aus der Privatwirtschaft und Verwaltung erste Einblicke in die praktische Anwendung der BIM-Methode auf den Gebieten BIM und Ingenieurvermessung, Integration von BIM und GIS, sowie BIM-Standardisierung.
- Abgerundet wird der Leitfaden durch einen Katalog von Software- und Dienstleistungsangeboten.

Die Beiträge zu diesen Kapiteln wurden von mehr als 50 Autoren aus der Privatwirtschaft, Verwaltung und Wissenschaft verfasst und vom Redaktionsteam des Leitfadens überprüft und redaktionell bearbeitet.
Gestützt auf die im Leitfaden enthaltenen Informationen sowie eigene Forschungsarbeiten auf dem Gebiet der Digitalen Modelle der bebauten Umwelt und der Integration von Geoinformationssystemen in den BIM-Prozess, werden im Vortrag zunächst einige wesentliche Gemeinsamkeiten und Unterschiede digitaler Modelle der bebauten Umwelt aus den Fachdisziplinen Geoinformatik sowie Architektur und Bauingenieurwesen beschrieben und anschließend technische Möglichkeiten zur Integration der Modelle diskutiert.

3 Literaturverzeichnis

KADEN, R., CLEMEN, C., SEUß, R., BLANKENBACH, J., BECKER, R., EICHHORN, A., DONAUBAUER, A., KOLBE, T.H. & GRUBER, U. (Hrsg.), 2017: DVW – Gesellschaft für Geodäsie, Geoinformation und Landmanagement e. V.; Runder Tisch GIS e.V. (Hrsg.): Leitfaden Geodäsie und BIM – Version 1.0, 26. September 2017. DVW – Gesellschaft für Geodäsie, Geoinformation und Landmanagement e. V.; Runder Tisch GIS e.V..

[2] Download unter:
https://rundertischgis.de/images/2_publikationen/leitfaeden/GeoundBIM/Leitfaden%20Geod%C3%A4sie%20und%20BIM_Onlineversion.pdf

Automatisierte Generierung eines digitalen Landschaftsmodells in 3D

GEORG FIUTAK[1], CAROLINE MARX[2], PHILIPP WILLKOMM[1], ANDREAS DONAUBAUER[2] & THOMAS H. KOLBE[2]

Zusammenfassung: Im Forschungsprojekt 3D-DLM, an dem sowohl der Runde Tisch GIS e. V. (in Zusammenarbeit mit der TU München und der Firma M.O.S.S. Computer Grafik Systeme GmbH als Projektbearbeiter), als auch die Länder Bayern und Baden-Württemberg beteiligt sind, erfolgt in der Demonstrationsphase der prototypische Aufbau eines konsistenten 3D Digitalen Landschaftsmodells für ein Testgebiet mit relevanter Größe. Dabei werden die in der Konzeptionsphase (MARX et al. 2017) festgelegten Objektklassen auf das DGM in LoD1 modelliert und um die amtlichen 3D-Gebäudemodelle ergänzt.
Der Beitrag zeigt die notwendigen Schritte bis zur Modellierung eines 3D Digitalen Landschaftsmodells mittels des Open Source Tools „3Dfier“[3]; dabei wird sowohl auf die notwendige Vorverarbeitung der Daten eingegangen, als auch auf die anschließende Abbildung auf das CityGML-Datenmodell. Das Ergebnis wird einer Qualitätsbewertung unterzogen.

1 Einleitung

Dreidimensionale Digitale Landschaftsmodelle nehmen aktuell eine immer größer werdende Bedeutung ein (COORS et al. 2016). Dies zeigt sich sowohl darin, dass es bereits eine Vielzahl an Anwendungsszenarien gibt, bei der die dritte Dimension berücksichtigt werden muss, als auch an den Aktivitäten einiger europäischer Staaten zur Überführung ihrer 2D-Geobasisdaten in die dritte Dimension (STOTER et al 2016).
Um dem zunehmenden Bedarf an Landschaftsmodellen in 3D gerecht zu werden und die Entwicklung dieser Thematik weiter voranzutreiben, wurde im Frühjahr 2016 der Runde Tisch GIS e.V. mit dem Projekt „3D-DLM“ (3D Digitales Landschaftsmodell) beauftragt.
In diesem Zusammenhang sei kurz erwähnt, dass ein 3D-DLM über ein einfaches 2,5-dimensionales DLM hinausgeht. Werden beim 2,5-DLM die Objekte aus dem 2D-DLM lediglich dem Geländeverlauf (DGM bzw. DOM) angepasst, so umfasst ein 3D-Landschaftsmodell zusätzlich auch als dreidimensionale Körper repräsentierte künstliche Strukturen (Gebäude, Brücken, Dämme, u. v. m.) und natürliche Objekte, wie z. B. Vegetationsobjekte. Einen Schritt weiter geht das semantische 3D-Landschaftsmodell: Hierbei sind alle Objekte nach thematischen und funktionalen Aspekten strukturiert; d. h. die Objekte werden nach logischen Kriterien in ihre Bestandteile zerlegt (ein Gebäude besteht z. B. aus einem Dach und Wänden; eine Straße aus der Fahrbahn und den Seitengräben) und nicht etwa nach Visualisierungsaspekten (KOLBE 2009).

[1] M.O.S.S. Computer Grafik Systeme GmbH, Hohenbrunner Weg 13, D-82024 Taufkirchen, E-Mail: [gfiutak, pwillkomm]@moss.de
[2] Lehrstuhl für Geoinformatik, Technische Universität München, Arcisstraße 21, D-80333 München, E-Mail: [caroline.marx, andreas.donabauer]@tum.de
[3] https://github.com/tudelft3d/3dfier

Die Auftraggeber der zweiten Phase des Projekts (Demonstrationsphase) sind die Bodenseeanrainer Bayern (Landesamt für Digitalisierung, Breitband und Vermessung Bayern - LDBV) und Baden-Württemberg (Landesamt für Geoinformation und Landentwicklung Baden-Württemberg - LGL).
Im vorliegenden Beitrag werden die Resultate dieser Demonstrationsphase dargestellt. Der Abschlussbericht zur vorausgegangenen Konzeptionsphase findet sich unter: https://rundertischgis.de/images/2_publikationen/stud_berichte/3D-DLM%20-%20Abschlussbericht.pdf.
Die Demonstrationsphase umfasst die Arbeitspakete „Vorverarbeitung der Daten“, „Anwendung des 3Dfiers auf die Daten“, „Abbildung auf das CityGML-Datenmodell“, „Qualitätsbewertung der erzeugten 3D-Objekte mittels statistischer Methoden“ und „Bereitstellung der Ergebnisdaten“.

1.1 Testgebiet

In Absprache zwischen dem LGL und dem LDBV wurde ein Projekt-Testgebiet im Grenzbereich zwischen Baden-Württemberg und Bayern festgelegt. Das Gebiet nordöstlich vom Bodensee umfasst 254km²; die genaue räumliche Lage und Ausdehnung kann den folgenden Abbildungen entnommen werden.

Abb. 1: Umfang des Testgebiets

1.2 Verwendete Software

Für die Umsetzung der Demonstrationsphase wurde sich für das bestehende Programm 3Dfier entschieden. Dies ist ein Open Source Tool zur Erzeugung von 3D-Objekten aus 2D-DLM und

LIDAR Punktwolken, das an der TU Delft in den Niederlanden entwickelt wurde und für eine ähnlich gelagerte Aufgabe in den Niederlanden eingesetzt wird. Detailinformationen zu diesem Werkzeug finden sich unter https://github.com/tudelft3d/3dfier.
Das Grundprinzip des 3Dfier illustriert die folgende Abbildung:

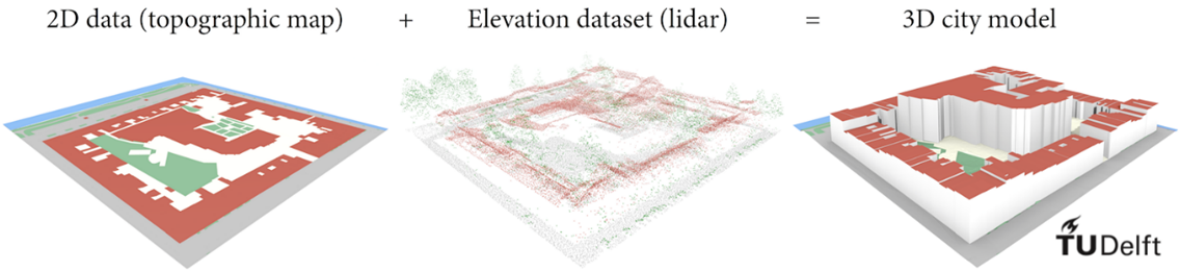


Abb. 2: Funktionsweise 3Dfier

Basierend auf seiner Semantik wird jedes Polygon des zu Grunde liegenden 2D-Datenbestandes mittels angepasster Methoden in die dritte Dimension angehoben. Dabei kommen unterschiedliche Regeln je nach anzuhebender Objektklasse zur Anwendung. Dieses Regelwerk berücksichtigt auch die Semantik der Nachbarschaftsbeziehungen von Polygonen unterschiedlicher Objektklassen.
Ergebnis und Ziel der Anwendung ist ein trianguliertes Digitales Landschaftsmodell in 3D.

2 Vorverarbeitung der 2D Eingangsdaten

Für die Anwendung der Software 3Dfier ist es erforderlich, dass die zweidimensionalen Eingangsdaten lückenlos und überlappungsfrei vorliegen. Eine Ausnahme bilden dabei die Brücken, da sich diese über dem eigentlichen Gelände befinden und dort somit Überlappungen mit den darunter liegenden Objekten auftreten.

2.1 Herausforderungen

Ausgangsdaten für die Erzeugung des 3D Digitalen Landschaftsmodells sind die ALKIS- und ATKIS-Daten der Auftraggeber. Viele dieser Daten liegen lediglich linienhaft vor. Dies erfordert eine entsprechende Pufferung um standardisierte bzw. klassifizierte Werte ausgehend von den ATKIS-Attributen. Eine weitere Herausforderung besteht darin, dass sich die Geometrien der einzelnen Objektarten teilweise gegenseitig überlagern. Dies trifft beispielsweise für den linienhaften Verlauf der Straßen oder von Flüssen innerhalb eines Wald-Polygons zu. Dies verdeutlicht die Gegenüberstellung der beiden unteren Abbildungen. Die Straßen (rot), die Wege (gelb) und die Flüsse (blau) zerschneiden einige der Waldpolygone. Dies macht es einerseits erforderlich, dass die neu entstehenden Teilflächen des ursprünglichen Waldpolygons den Bezug zu diesem nicht verlieren; d.h. nach dem Zerlegen eines Wald-Polygons in mehrere kleine Objekte, müssen diese mittels der ursprünglichen ID zu der Geometrie MultiSurface aggregiert werden. Andererseits müssen für die Pufferung der linienhaft vorliegenden Fluss- und Straßenachsen Teile der Waldflächen entsprechend freigestellt werden.

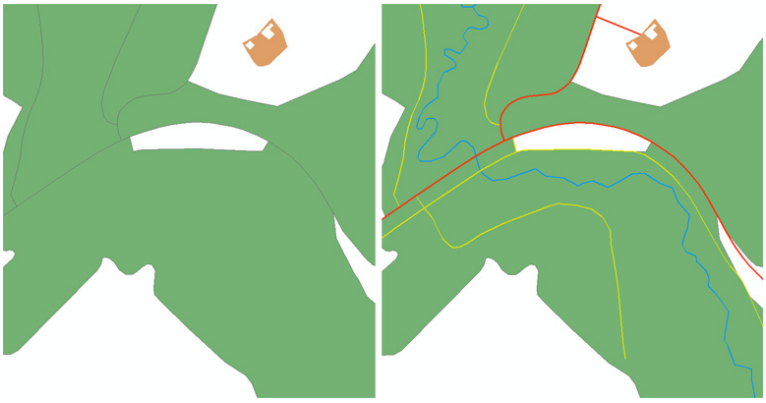


Abb. 3: Auftrennen von Waldflächen

2.2 Allgemeines Vorgehen

Zunächst werden alle relevanten Objektarten einzeln betrachtet und entsprechend aufbereitet. Dies umfasst unter anderem die Filterung, Pufferungen um definierte Werte oder die Erzeugung zusätzlicher Attribute. Da als Resultat der Pufferung von Linienobjekten sich diese neu erzeugten Polygone mit den benachbarten Flächen überlagern, müssen diese in einem nächsten Schritt freigestellt werden. Dadurch liegen in der Ebene alle Objekte überlappungsfrei vor. Schließlich erfolgt die Erzeugung der Terrain-Flächen dadurch, dass die bisher noch vorhandenen Lücken als Terrain-Objekte deklariert werden.

2.3 Betrachtung der einzelnen Objektarten

Im 3Dfier müssen die ALKIS- und ATKIS-Daten auf eine der folgenden Objektklassen abgebildet werden:

- Building
- Terrain
- Road
- Water
- Forest
- Bridge
- Separation

Es folgt die Betrachtung der Aufbereitung der einzelnen Objektarten und ihrer Zuordnung zu den entsprechenden Objektklassen des 3dfier.

2.3.1 Building

Die vom 3Dfier erzeugten Gebäudemodelle sollen im Nachgang durch die vorhandenen LoD2-Gebäudemodelle der Auftraggeber ersetzt werden. Für eine korrekte Triangulation der an die Gebäude angrenzenden Polygone müssen dennoch zunächst die entsprechenden Gebäude-Grundflächen in den Workflow zur automatischen 3D-DLM Erzeugung eingehen. Daher werden für die Umsetzung die Grundrisse der LoD2-Gebäudemodelle verwendet.

2.3.2 Bridge

Als Grundlage für die Aufbereitung der Brücken-Objekte dienen Bauwerke, Anlagen und Einrichtungen für den Verkehr (MARX et al. 2017).
Sowohl die flächenhaft, als auch die linienhaft vorliegenden Bauwerke werden mittels der Bauwerksfunktion BWF gefiltert. Dadurch werden Objekte herausgefiltert, die Tunnel, Schutzgalerien, Schleusenkammern oder Durchfahrten darstellen.
Die Geometrie der flächenhaften Brückenobjekte kann dabei direkt übernommen werden. Die linienhaften Brückenobjekte werden zusätzlich um die Mittelachse gepuffert. Da keine Angaben über die Breite der Brücken vorhanden sind und außerdem diese der Breite der anschließenden Straßenbreiten entsprechen sollen, werden diese Breiten für die Pufferung übernommen.
Nach Rücksprache mit den Auftraggebern wurde sich außerdem darauf verständigt, dass die Straßen an den Brückenobjekten nicht unterbrochen werden sollen, sondern diese auch zusätzlich zur Brücke dargestellt werden sollen. Dadurch ist es beispielsweise leichter möglich, die Gesamtfläche einer Straße zu berechnen.

2.3.3 Road

Straßenverkehr

Da die flächenhaft vorliegenden Objekte des Straßenverkehrs nicht nur die eigentlichen Fahrbahnen darstellen, sondern zusätzlich beispielsweise auch Verkehrsbegleitflächen, werden diese entsprechend gefiltert und nicht weiter als Straßenobjekte behandelt.
Außerdem ist zu beachten, dass sich flächenhaften Straßen und Straßenachsen teilweise überlagern. In diesem Fall werden die bereits vorhandenen Straßenflächen übernommen und lediglich die Achsen gepuffert, die keine Überlappung aufweisen.

Wege

Bei den Wegen erfolgt die Pufferung ebenfalls über die Breite der Fahrbahn und der entsprechenden Klassenkategorie.

Gleise

Um die Gleise zu bearbeiten, erfolgt zunächst eine Filterung über das Attribut BKT (Bahnkategorie). Somit lassen sich zum Beispiel Seilbahnen herausfiltern.

2.3.4 Water

Auch im Bereich der Gewässer-Objekte liegen diese flächenhaft und linienhaft vor. Die flächenhaften Objekte können direkt übernommen werden. Die linienhaft vorliegenden Gewässerobjekte werden mittels der Breite des Gewässers (BRG), die wiederum durch Klassen repräsentiert werden, gepuffert.

2.3.5 Forest

Um unerwünschte Überlagerungen zu vermeiden, müssen die folgenden Objektarten in den Wald-Polygonen freigestellt werden:

- Brücken
- Gewässer
- Straßen
- Wege
- Gleise

Abb. 4: Freistellung Wald

2.3.6 Sonstige Objektarten

Neben den bereits beschriebenen Objektarten, die sich eindeutig einer Klasse des 3Dfiers zuordnen lassen, gibt es in den ALKIS- und ATKIS-Daten noch weitere Objektarten, die bei der Erzeugung des 3D-Landschaftsmodells berücksichtigt werden sollten. Diese werden zunächst einer geeigneten 3Dfier-Klasse zugeordnet und im anschließenden Mapping der passenden CityGML-Klasse zugewiesen.

Tab.1: Zuordnung Sonstiger Objektarten

Objektart	**3Dfier**	**CityGML**
Landwirtschaft	TERRAIN	LandUse
Bauwerke oder Anlagen für Sport, Freizeit und Erholung	BUILDING	LandUse
Bauwerke oder Anlagen für Industrie und Gewerbe	BUILDING	LandUse
Flugverkehr	TERRAIN	LandUse
Übrige Objektarten der Vegetation außer Wald und LWS	TERRAIN	LandUse

2.3.7 Terrain

Die Klasse „Terrain“ wird nicht auf Grundlage vorhandener ATKIS-Daten erzeugt. Vielmehr werden die nach der Daten-Vorverarbeitung noch vorhandenen ‘Lücken‘ durch Flächen aufgefüllt, die der Klasse „Terrain“ zugewiesen werden. Da unter den Brücken-Objekten Gelände dargestellt werden soll, werden die Flächen, die durch die Brücken repräsentiert werden, ebenfalls zusätzlich dem Terrain zugewiesen.

3 Anwenden des 3Dfiers

3.1 Eingangsdaten

Der 3Dfier erfordert zwei Arten von Eingangsdaten. Zunächst sind dies lückenlose, flächenhafte 2D-Daten im ESRI-Shape-Format, welche die Landschaft beschreiben und den Objektklassen des 3Dfiers zugeordnet sind (s. Kapitel 2.3).
Des Weiteren ist eine die Oberflächenstruktur (DOM) wiedergebende Punktwolke (LAS/LAZ) erforderlich, die mindestens die folgenden Klassen aufweist:

- 0-1: Created, never classified and / or unclassified
- 2: Ground
- 3-5: Vegetation

3.2 Ergebnisdaten

Neben CityGML stehen noch die Ausgabeformate OBJ, CSV (nur Gebäude), Shapefile und PostGIS zur Verfügung. Die Tests wurden mit den Ausgabeformaten CityGML (für die semantische Weiterverarbeitung im Projekt) und OBJ (für Testvisualisierung und visuelle Qualitätsbewertungen) durchgeführt.

3.3 Umsetzung

Die Umsetzung und damit die konkrete Anwendung des 3Dfier im Rahmen dieses Projektes wurde mit den im Vorfeld aufbereiteten ALKIS- und ATKIS-Daten, sowie den zur Verfügung gestellten 3D-Punktwolkendaten (LIDAR) der Auftraggeber durchgeführt (MARX et al. 2017).
Die gesamte Steuerung des 3Dfier erfolgt über eine Konfigurationsdatei. Dabei handelt es sich um eine reine ASCII-Datei, die auch einfach über ein Steuerprogramm erzeugt werden kann. Damit eignet sich der 3Dfier sehr gut für die Einbindung in automatisierte rein batchorientierte und auftragsgesteuerte Prozesse, wie sie bei der Verarbeitung von Massendaten in der Praxis zur Anwendung kommen.

3.3.1 Prozessierung

Der Umsetzungsprozess erfolgt in 4 Einzelschritten:

1. Einlesen der Konfigurationsdatei und Prüfung auf korrekte Syntax
2. Analyse der Polygone und ggf. Aufbereitung (auflösen von Multipolygonen, bereinigen von Selbstüberschneidungen)
3. Einlesen der Punktwolke; dies ist der zeitintensivste Prozess und macht ca. 90% der Gesamtzeit aus, abhängig von der Punktdichte
4. Erzeugung des 3D-Modells

Parallelisierung

Vor dem Hintergrund einer landesweiten Umsetzung wurde im Rahmen des Projektes untersucht, in wie weit sich die Anwendung des 3Dfier parallelisieren lässt und welche Einsparungseffekte bzgl. Laufzeit durch eine solche Parallelisierung erreichbar wären.
Die Architektur und Prozesslogik des 3Dfiers lässt eine Parallelprozessierung mit einfachen Mitteln zu. Allerdings sind für das notwendige Vernähen der entstehenden Teilgebiete des 3D-DLM Vorkehrungen im Prä- und Postprocessing zu treffen.
Um eine parallelisierte Produktion zu simulieren, wurde das Testgebiet in annähernd gleich große Gebiete unterteilt und diese wiederum gleichmäßig fünf Parallelprozessen zugeordnet.
Die Durchführung erfolgte auf einem Rechner mit den folgenden Merkmalen:

```
Operating System:          Microsoft Windows 7 Professional (x64)
CPU:                       Intel(R) Core(TM) i7-4712MQ CPU @ 2.30GHz
Number of CPU Cores:       4
Number of Logical CPUs:    8
Total Memory Size [MB]:    8192
```

Vergleich Gesamtumsetzung / Parallelisierung
(254 km², 296 LIDAR-Kacheln á 1 km², ∅ Punktabstand 1,2 m):

```
Komplettumsetzung:              110 Minuten
Gebiete in 5 Parallelprozessen:  27 Minuten
```

Mit der Aufteilung in Gebiete, die hinsichtlich der verfügbaren IT-Infrastruktur (Prozessoren, Hauptspeicher, Plattendurchsatz) optimiert portioniert werden und der daraus resultierenden Parallelisierung der Prozesse kann somit ein deutlicher Geschwindigkeitszuwachs erreicht werden.

3.3.2 Prozessanalyse

Der Umsetzprozess bietet eine Reihe von Konfigurationsmöglichkeiten zur Steuerung der Umsetzung. Je nach Objektart haben diese unterschiedlichen Auswirkungen und auch die Kachelränder erfordern eine intelligente Vorverarbeitung der Daten.

Ausdünnung der Punktwolke
Über den Konfigurationsparameter *thinning* kann die Punktwolke zur Laufzeit ausgedünnt werden. Damit kann die Laufzeit bei besonders hoher Punktdichte erheblich verringert werden.
Allerdings führt eine zu geringe Punktdichte zu Geometriefehlern bzw. semantisch nicht korrekten Ergebnissen wie unruhigen Straßenoberflächen.
Von der Anwendung des Parameters *thinning* ist daher im aktuellen Softwarestand abzuraten.

Triangulation reduzieren (simplification)
Die Oberflächen der Objektklassen *Forest* und *Terrain* werden direkt aus der Punktwolke abgeleitet und trianguliert. Die Oberflächen-Triangulation des Ergebnisobjekts kann über den Parameter *simplification* beeinflusst werden. Durch Anwendung dieses Parameters kann zu einem die Dateigröße reduziert werden (gerade bei der Erzeugung von CityGML), zum anderen wird aber natürlich auch das Ergebnis generalisiert.
Bei der Objektklassen *Forest* kann der Detailgrad der Umsetzung je nach Anwendungsfall von niederrangiger Bedeutung sein, dient er doch im Wesentlichen ‚nur' zur groben Visualisierung.
Umso mehr hat dieser Wert aber bei der Visualisierung des DGM eine Bedeutung, da es hier auf eine möglichst exakte Wiedergabe ankommt, welche dann aber zu einer erheblich größeren Ergebnisdatei führen kann.
Da aber in der Praxis für die betrachtete Projektregion ein hochgenaues und hochauflösendes DGM bereits existiert, ist das durch den 3Dfier erstellte DGM nur als Zwischenergebnis zu betrachten und dieses sollte in der Folge in Form eines Postprozesses durch das erwähnte existierende DGM ersetzt werden. Daher kann der Parameter „simplification" zur Prozessbeschleunigung empfohlen werden.

Gebietsrand
Beim Vernähen von Kacheln muss darauf geachtet werden, dass am Gebietsrand angrenzende Flächen in 2-facher Nachbarschaft zur Randberechnung mit einbezogen werden, da es sonst zu Kachelrandfehlern in Form virtueller Wände kommt.

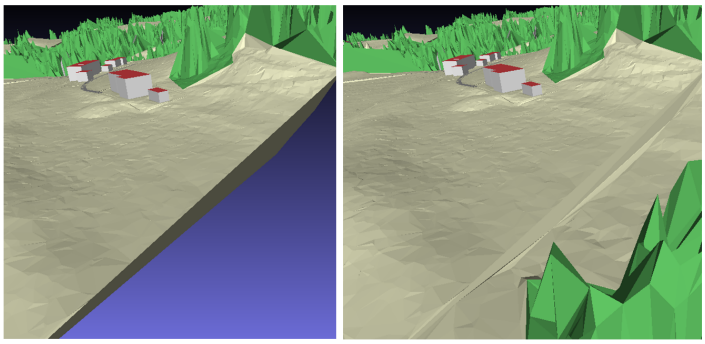


Abb. 5: Kachelrandfehler

4 Abbildung auf CityGML-Modell

4.1 Semantisches Mapping

Da vom 3Dfier nur die Klassen Building, Bridge, Road, Water, Forest, Terrain und Separation erzeugt werden, ist im Nachgang eine Semantische Transformation bzw. eine Anpassung entsprechend der Mapping-Tabelle aus Projektphase 1 (MARX et al. 2017) notwendig, um ein möglichst feingranulares CityGML-Modell zu erhalten.
So werden beispielsweise die Gleise, die nach Anwendung des 3Dfiers auf die thematische Klasse Road abgebildet wurden, im Zielmodell wieder in Klasse Railway abgebildet. Sportanlagen, die zunächst als Building abgebildet wurden (da die Oberfläche der sportlichen Anlagen als ebene Fläche dargestellt werden sollten), werden im anschließenden Mapping auf die Klasse LandUse abgebildet.

4.1.1 Mapping der Attribute

Alle Attribute der ATKIS-Daten werden entsprechend des CityGML-Standards überführt.
Für den Fall, dass es für ein Attribut der ATKIS-Daten kein passendes Pendant im CityGML-Standard gab, wurden diese Informationen als generische CityGML-Attribute überführt.

4.2 Geometrie

Alle Klassen wurden als MultiSurfaces im LoD1 modelliert. Lediglich die im Nachgang eingefügten Gebäudemodelle des LDBV und des LGL liegen im LoD2 vor.

4.3 Zusätzliche Transformationen

4.3.1 Setzen des korrekten Koordinatensystems

Die CityGML-Daten werden vom 3Dfier aktuell im EPSG:7415 (Niederländisches Bezugssystem) ausgegeben. Dies erfordert es, dass im Nachgang das korrekte Koordinatensystem (EPSG:31468) gesetzt wird.

4.3.2 Einfügen zusätzlicher Attribute

Neben den in den ATKIS-Daten vorhandenen Attributen werden zusätzliche Attribute erzeugt. Dies sind zum einen die beiden generischen Attribute *min_Hoehe* und *max_Hoehe*. Das

Speichern der minimalen und maximalen absoluten Höhe als Attribut ist deshalb sinnvoll, weil mittels dieser Attribute in einfachen Viewern räumliche Selektionen ausgeführt werden können. Zum anderen wird für jedes Objekt die Fläche berechnet und im Attribut *Flaecheninhalt* abgespeichert. Dies kann dann sinnvoll sein, wenn man beispielsweise für Bauarbeiten an Straßen wissen muss, wie groß die Oberfläche eines Straßenabschnitts ist.

5 Austausch von Objektarten

5.1 Gebäude

Für das dreidimensionale Landschaftsmodell sollen die original LoD2-Gebäudemodelle, die vom LDBV und LGL bereitgestellt wurden genutzt werden. Da für die Anwendung des 3Dfiers dennoch zunächst eine lückenlose Flächenabdeckung vorliegen muss, müssen die mittels des 3Dfiers erstellten Klötzchenmodelle im Nachgang mit den LoD2-Gebäudemodellen ausgetauscht werden.

5.2 Wald

Nach Absprache mit den Auftraggebern sollte statt der durch den 3Dfier erzeugten Vegetationsobjekte die Waldfläche besser lediglich über das Gelände gebogen werden und diese Flächen mit der entsprechenden Punktwolke verlinkt werden.

5.3 DGM

Durch den 3Dfier wird ausgehend von den 2D-Polygonen „Terrain“ (dies entspricht den vorhandenen Lücken in den aufbereiteten 2D-Daten) ein neues DGM erzeugt. Da sowohl für das LDBV, als auch für das LGL amtliche hochaufgelöste DGM-Daten vorliegen, sollte statt des erzeugten Terrains dieses DGM genutzt werden.

6 Qualitätsbewertung

6.1 Umgang mit Inkonsistenzen in Eingangsdaten

Generell ist der 3Dfier tolerant gegenüber inkonsistenten Eingangsdaten, das Ergebnis ist dann gegebenenfalls lokal unbefriedigend. Das Programm ist daher für ein iteratives Vorgehen auf Basis von nur teilweise korrekten Eingangsdaten gut geeignet. Im Folgenden wird auf den Umgang des 3Dfiers mit üblichen Datenfehlern eingegangen.

6.1.1 Nicht freigestellte Objekte

Nicht freigestellte Objekte in 2D (z.B. Straße in Wald) führt zu Überlappungen, im schlimmsten Fall zu Triangulierungsfehlern. Die korrekte Aufbereitung der 2D-Ausgangsdaten im Sinne der vom 3Dfier vorgegebenen Randbedingungen ist daher essentiell für die Ergebnisqualität.

6.1.2 Lücken in der Punktwolke

Lücken in der Punktwolke (zum Beispiel unter Brücken) führen zu Fehlern im 3D-DLM.
Der 3Dfier erscheint sehr sensibel gegenüber Datenlücken, wie sie in der Praxis naturgemäß in 3D-Punktwolken vorkommen. Bei Anwendung der aktuellen Version des 3Dfiers ist daher ein

Präprozess sinnvoll, der solche Datenlücken über einfache Interpolationsprozesse oder durch Nutzung des vorhandenen DGMs auffüllt. Dabei können aber naturgemäß nur konkrete Lücken, wie z.B. unter Brücken behandelt werden.

6.1.3 Inkonsistenzen in 2D-Ausgangsdaten

Inkonsistenzen in 2D-Ausgangsdaten (Polygone) werden vom 3Dfier soweit möglich korrigiert, z.B. Multipolygone, Selbstüberschneidungen.

6.1.4 Lücken im 2D-Datenbestand

Der 3Dfier toleriert bei der 3D-DLM-Erzeugung Lücken im 2D-Datenbestand, es kommt nicht zum Abbruch. Das Ergebnis beinhaltet aber naturgemäß entsprechende Lücken im 3D-DLM.

6.2 Verwendung von unklassifizierten Punktwolken

Die Verwendung eines (unklassifizierten) bDOM ist zwar technisch möglich, aber faktisch nicht sinnvoll. Der 3Dfier setzt bei den verwendeten LAS-Daten klassifizierte Bodenpunkte voraus. Eine klassifizierte Vegetation wird empfohlen, um Gebäudehöhen exakter abbilden zu können und das Terrain sauberer abzubilden, ist aber nicht zwingend notwendig.
So zieht zum Beispiel die Klasse Road die Höheninformation nur aus den Bodenpunkten (Class 2: Ground). Wurden keine Bodenpunkte klassifiziert, werden die Objekte dieser Klasse mit der Höhe 0 versehen.

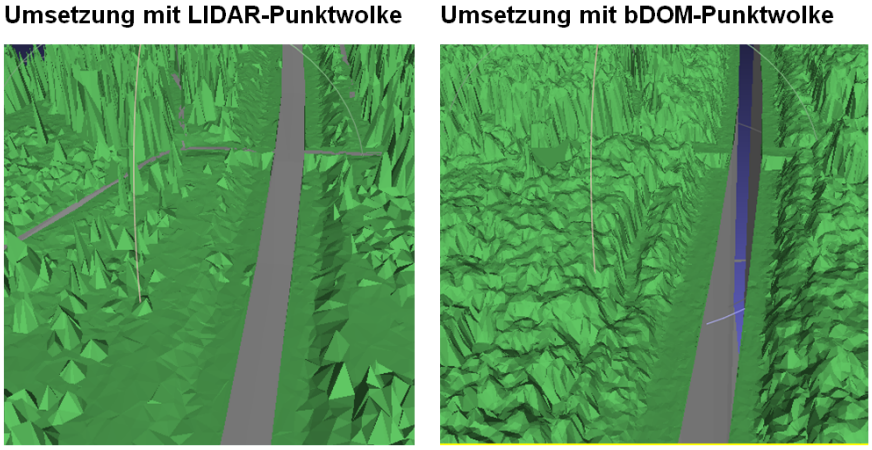


Abb. 6: Vergleich Umsetzung LIDAR / bDOM

6.3 Betrachtung der einzelnen Objektarten

Im Folgenden werden die einzelnen Objektarten hinsichtlich ihrer Qualität betrachtet und auf mögliche Ansätze zur Vermeidung von Fehlern eingegangen.

6.3.1 Building

Grundlage für die Umsetzung ist die für die amtliche LoD Erzeugung verwendeten Grundrisse (ALKIS). Die Erzeugung im Rahmen des 3Dfiers ist wegen Nachbarschaftsbeziehungen

notwendig und erfolgt immer in LoD1. Im Postprozess werden dann die erzeugten LoD1-Modelle durch amtliche LoD2-Modelle ersetzt.
Um Lücken im 3D-DLM-Modell zu vermeiden, muss die Lage der verwendeten amtlichen 3D-Gebäuden exakt den für den 3Dfier-Lauf verwendeten Grundrissen entsprechen. Kritisch können die unterschiedlichen Zeitpunkte zwischen Modellerzeugung und 3Dfier-Lauf sein.
Für die nahtlose Integration des amtlichen 3D-Modells in das umgebende Gelände ist die Verwendung des amtlichen DGMs Voraussetzung

6.3.2 Bridge

Vor allem und gerade im Bereich von Brücken besteht eine hohe Sensibilität in Bezug auf die Ausgangsdaten. In vielen Fällen ist das vollautomatische Ergebnis aber bereits zufriedenstellend.
Neben den geschlossenen 3D-Informationen aus Punktwolke und DGM sind ebenso die lückenlosen und überlappungsfreien 2D-Informationen entscheidend für eine saubere Umsetzung von Brücken. Andernfalls entstehen auch hier Triangulationsfehler, wie z.B. Lücken im 3D-DLM. Generell können Triangulationsprobleme durch Verwendung des DGMs statt der Bodenpunkte aus der Punktwolke minimiert werden.

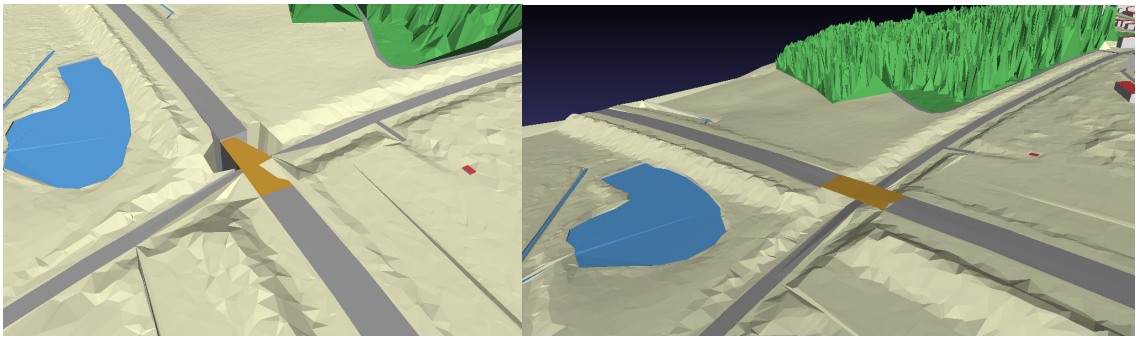


Abb. 7: Behebung von Triangulationsproblemen durch Austausch der Bodenpunkte (links: Umsetzung mit Bodenpunkten aus LAS, rechts: Bodenpunkte aus DGM)

Die Lage der Brücken passt im Testgebiet zu anschließenden Verkehrsobjekten, so dass keine Lücken in X und Y zu den anschließenden Straßen entstehen. Die Höhe der Brücken passt dagegen häufig nicht exakt zu den angrenzenden Straßenobjekten.

6.3.3 Road

Liegen Lücken im Punktdatenbestand vor (z.B. unter Brücken) erfolgt eine unsaubere Interpolation der Höhe, welche zu in der Realität nicht vorhandenen Tälern und Kuppen im Straßenverlauf führen.

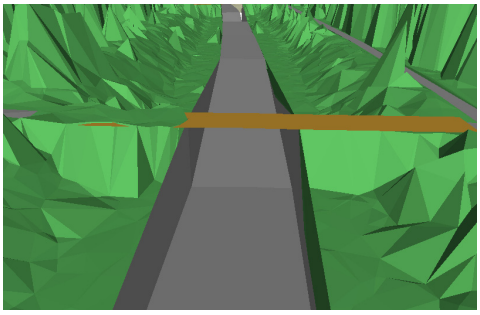


Abb. 8: Unsaubere Straßenverlauf durch Lücken in der Punktwolke

Korrekt wäre eine Interpolation ausschließlich auf Basis vorhandener Punkte. Dabei müsste der 3Dfier die Semantik des Straßenobjektes berücksichtigen und so starke Gefällesprünge oder Unstetigkeitsstellen vermeiden. Aktuell kann das Problem nur durch automatische Lückenfüllung z.B. durch Austausch der Bodenpunkte durch Punkte aus dem amtlichen DGM vermieden werden.
Ist die Punktdichte des verwendeten DOM zu niedrig, kommt es zu einer unruhigen Straßenoberfläche. Das semantische Wissen um eine ebene Straßenoberfläche kommt hier offensichtlich nicht oder nicht ausreichend gewichtet zum Tragen. Auch bei uneinheitlichem Gelände, z.B. sich änderndem Gefälle oder durch Lageversatz verursachten uneinheitlichen Höhen der auf der Straße liegenden Punkte entstehen Unstetigkeiten im Straßenverlauf (Querneigung, Knicke etc.). Somit ist für einen ebenen Straßenverlauf eine Pufferung der linienhaften 2D-Straßenelemente mit einer der Realität entsprechenden Breite elementar.
Der 3Dfier bietet außerdem in der aktuellen Version keine Möglichkeit zur Unterscheidung der unterschiedlichen Verkehrswege (z.B. Straßen und Gleise). Aus Sicht der 3D-Modellierung im Anwendungsfall dieses Projektes erscheint dies auf Grund des Generalisierungsgrades unkritisch. Für die Unterscheidung in der Visualisierung erfolgt im Rahmen eines Postprozesses eine Trennung der Verkehrstypen.

6.3.4 Water

Bei der Qualitätsprüfung wird zwischen stehenden und fließenden Gewässern unterschieden.
Für stehende Gewässer gilt die Annahme, dass die Neigung 0 sein muss. Die entsprechende Überprüfung der Daten des LDBV ergab, dass 19,8% der stehenden Gewässer keine horizontale Fläche repräsentieren und damit als fehlerhaft klassifiziert wurden.
Für die Qualitätsbewertung der Fließgewässer gilt die Annahme, dass alle Flächen eines Objekts in etwa (auf 5 Grad genau) in die gleiche Richtung schauen müssen. Ist dies nicht der Fall, werden die Objekte als fehlerhaft bewertet. Mit der oben aufgestellten Annahme ergibt sich somit für die Daten des LDBV ein Prozentanteil von 58% der Fließgewässer, die fehlerhaft sind.
Der Grund hierfür ist den auftretenden Fehlern bei Straßen sehr ähnlich. Demnach ist auch hier eine Pufferung der Fließgewässer mit einer der Realität entsprechenden Breite notwendig.
Außerdem sind einige der ursprünglichen Gewässerachsen durch sehr viele kleine Abschnitte repräsentiert. Dadurch kann es zwischen den einzelnen Objekten zu starken Höhendifferenzen kommen. Dies kann bereits in der Vorverarbeitung der Daten berücksichtigt und verbessert werden.

6.3.5 Terrain

Da das vom 3Dfier erzeugte Terrain im Nachgang durch das ursprüngliche hochaufgelöste und genaue DGM ersetzt wurde, erübrigt sich hier eine Qualitätsbewertung.

6.3.6 Forest

Das gleiche gilt auch für den Wald. Dieser soll im Ergebnis lediglich flächig über das Gelände gelegt werden.

6.3.7 Zusammenfassung und Fazit

Zusammengefasst ergeben sich folgende Kriterien für eine Qualitätsverbesserung des resultierenden 3D-DLMs:

- Bodenpunkte der Punktwolke durch Informationen aus amtlichem DGM (z.B. DGM1) ersetzen
- Möglichst hohe Übereinstimmung in der Lage zwischen DLM und DGM
- Realistische Breiteninformation pro Abschnitt eines Verkehrsweges

Zufriedenstellende Ergebnisse werden aktuell erreicht bei:

- Ebenen oder leicht welligem Gelände
- Gerader oder nur leicht kurviger Achsenverlauf
- Daten entsprechen möglichst weitgehend oben genannten Kriterien zur Qualitätsverbesserung
- Realistische Darstellung bei Übersichtsbetrachtung

Grobe Fehler beschränken sich im Wesentlichen auf Kreuzungsbereiche und Über-/Unterführungen.

7 Bereitstellung der Ergebnisdaten

Abgesehen vom DGM wurden alle Daten als CityGML-Instanzdokumente erzeugt und mittels des CityGML-Importers in eine 3D City Database[4] auf einem PostgreSQL DBMS (mit PostGIS) importiert. Nach Festlegen der BoundingBox und Anpassung diverser Einstellungen (z.B. Füllfarbe der Objekte) werden die Daten als COLLADA/glTF exportiert. Dieses Datenformat ist notwendig, um die Inhalte der Datenbank für eine 3D-Web-Visualisierung mittels Cesium aufzubereiten.
Eine Anreicherung der Objekte um thematische Attribute erfolgt mittels des in die 3D City Database integrierten Spreadsheet-Exporters.
Die Attribute eines Objektes werden aus einem exportierten Spreadsheet ausgelesen, welches über das Internet zugreifbar sein muss, z.B. via Google Spreadsheet.
Die Visualisierung des 3D-Landschaftsmodells wird über einen Web-Browser realisiert. Dies geschieht ohne Installieren oder Aktivieren zusätzlicher Erweiterungen (Plug-Ins) für den Browser.

[4] https://www.3dcitydb.org

8 Fazit & Ausblick

In der Demonstrationsphase konnte die grundsätzliche Machbarkeit einer hochgradig automatisierten Ableitung eines flächendeckenden 3D-DLM aus den vorliegenden ALKIS- und ATKIS-Daten nachgewiesen werden. Das verwendete Werkzeug 3Dfier lieferte mit den beschriebenen Ausgangsdaten im vollautomatischen Prozess ein fachlich sinnvolles Ergebnis, das aber im Detail noch inhaltliche Fehler enthält. Der Hardware-Ressourcenbedarf sowie die Laufzeiten der Prozessierung des Testgebiets erscheinen für eine Anwendung auch landesweiter Datenbestände geeignet.
Für ein sinnvolles Ergebnis ist eine Aufbereitung der Ausgangsdaten notwendig. Das betrifft sowohl die Vektordaten, als auch die verwendenden Oberflächenpunkte (Austausch der Bodenpunkte durch hochauflösendes DGM). Auch hat die Verwendung von realen Breiten anstelle von Klassen bei Straßen- und Gewässerobjekten elementare Bedeutung. Die notwendige Vorverarbeitung der Daten, genauso wie das abschließende Mapping auf ein CityGML-Zielmodell, lässt sich gut in einen automatisierten Gesamtprozessablauf integrieren.
Somit bietet der untersuchte Ansatz eine praktikable Möglichkeit um 2D-Daten automatisiert und konsistent in die dritte Dimension anzuheben und dabei alle gewünschten fachlichen Informationen zu übernehmen. Weitere vertiefte Untersuchungen könnten z.B. die Einbeziehung weiterer Softwarewerkzeuge sowie die Verkettung der Teilschritte zu einem Gesamtprozess beinhalten.

9 Literaturverzeichnis

COORS, V., HUBER, U., JOOS, G., REINELT, A., TEEGE, G. & VOERKELIUS, U., 2016: Stellungnahme des Beitrags zum Projektvorhaben Digitales Landschaftsmodell (3D-DLM). unveröffentlicht.

GRÖGER, G., KOLBE, T.H., NAGEL, C. & HÄFELE, K.-H., 2012: OpenGlS City Geography Markup Language (CityGML) Encoding Standard, Version 2.0.0., Open Geospatial Consortium.

KOLBE, T.H., 2009: Representing and Exchanging 3D City Models with CityGML. 3D Geo-Information Sciences, Lee, J. & Zlatanova, S. (Hrsg.), Springer.

MARX, C., FIUTAK, G., WILLKOMM P. & DONAUBAUER, A., 2017: Projekt 3D Digitales Landschaftsmodell (3D-DLM) am Runden Tisch GIS e.V. Abschlussbericht (Konzeptionsphase): Anwendungsszenarien, Datenevaluierung, Daten-modellierung & Methoden zur 2D -> 3D-Transformation. https://rundertischgis.de/images/2_publikationen/stud_berichte/3D-DLM%20-%20Abschlussbericht.pdf

STOTER, J., VALLET, B., LITHEN, T., PLA, M., WOZNIAK, P., KELLENBERGER, T., STREILEIN, A., ILVES, R. & LEDOUX, H., 2016: State-of-the-art of 3D national mapping in 2016. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **41**(B4), 653-660.

Geoinformationen in der Immobilienbewertung

RICHARD OTT[1]

Zusammenfassung: Geoinformationen gewinnen in der Immobilienwirtschaft auch bei der Bewertung und Risikoanalyse zunehmend an Bedeutung. Voraussetzung ist aber, dass neben einer hohen Qualität auch die flächendeckende Verfügbarkeit gewährleistet werden muss. Die Bundesregierung hat hier mit ihrer Initiative zu einem verbindlichen Open-Data-Gesetz einen guten Beitrag geleistet. Bisherige Anwendungen zeigen, dass vorhandene Geoinformation im Zusammenspiel mit dem nötigen Knowhow zur Verarbeitung der Daten und dem Fachwissen der Anwender kombiniert werden und Nutzen entfalten können. Der Beitrag zeigt beispielhaft, wie Lösungen realisiert werden können und welche Potentiale und Erwartungen hinsichtlich der Nutzung von Geoinformation seitens der Immobilienwirtschaft noch bestehen.

1 Allgemein

Immobilien bewegen sich nicht. Und das ist auch gut so. Sie befinden sich immer an der gleichen Stelle. Und das gilt natürlich auch für alle Attribute, die man ihr anheftet.
Man muss also nur noch die wertrelevanten Informationen finden und der Immobilie eindeutig zuordnen. Dann sind diese im Bedarfsfall sehr schnell und im Idealfall - automatisiert abrufbar.
Stellt sich die Frage: Welche Informationen werden in der Immobilienbewertung benötigt? Es gilt noch immer: Bewerten heißt vergleichen. Also benötigen die Gutachter*innen Parameter, die es ihnen ermöglichen, den Bewertungsgegenstand, oder einzelne Merkmale mit anderen Objekten zu vergleichen, um daraus den Wert der Immobilie abzuleiten.
Dabei wird im Wesentlichen unterschieden zwischen eigennutzungsfähigen Immobilen (Eigentumswohnungen und Einfamilienhäuser), bei denen der Besitz im Vordergrund steht und Renditeimmobilien. Hier wird der Fokus auf den Ertrag und die Rendite gelegt.
Bei der Bewertung von Einfamilienhäusern gilt: Sind die Immobilien aufgrund ihrer Objektmerkmale direkt vergleichbar.
Bei Renditeobjekten steht in der Regel keine ausreichende Anzahl von Immobilien zur Verfügung, die direkt miteinander vergleichbar sind. Deshalb erfolgt die Bewertung indirekt, insbesondere über die erzielbaren Erträge und die marktüblichen Renditen.
Geoinformationen können diese Bewertungsprozesse erheblich unterstützen, wenn sie bewertungsrelevante Parameter liefern, die im Rahmen des Bewertungsprozesses - am besten teil-/automatisiert - den Gutachter*innen zur Verfügung gestellt werden können.
Dabei unterscheidet man grundsätzlich zwischen zwei Arten von Vergleichsparametern:
Parameter,

- die sich auf das Umfeld der Immobilie beziehen (Vergleichspreise, Mieten, Renditen, Lageinformationen, Wettbewerb etc.)
- die direkt an der Immobilie „festgemacht“ werden können (Objekttyp, Wohn- und Nutzfläche, Ausstattung, Zustand etc.)

[1] Kurze Strasse 17, D-12167 Berlin

Ein ganz wesentlicher Ausgangsparameter für die Bewertung von Immobilien ist – und das gilt nicht erst seit der Schneider-Affäre – die vorhandene Wohn- bzw. Nutzfläche. Ist diese Eingangsgröße fehlerhaft, führen auch alle folgenden Berechnungen, unabhängig davon, wie genau sie durchgeführt werden, zu einem falschen Ergebnis.
Der Berechnung bzw. der Plausibilisierung der vorgelegten Flächenangaben kommt somit eine ganz zentrale Bedeutung zu.
HypZert, die Zertifizierungsgesellschaft für Gutachter in der Finanzwirtschaft hat deshalb zusammen mit der Firma M.O.S.S. Computer Grafik-Systeme eine teil-/automatisierte Lösung entwickelt, die es ermöglicht, sehr schnell und mit einer hohen Genauigkeit auf der Basis von Geoinformationen die zur Verfügung gestellten Angaben zur Mietfläche zu prüfen.
Dabei handelt es sich um einen internetbasierten Web-Dienst, der grundsätzlich auch in jede Bewertungssoftware über eine Schnittstelle integriert werden kann. Die Anwendung ist zusammen mit Gutachter*innen entwickelt und speziell auf deren Bedürfnisse abgestellt worden.
Im Vorfeld der Entwicklung fanden regelmäßig „Round-Table-Gespräche“ zwischen der Firma M.O.S.S. und den HypZert-Gutachter*innen statt. Diese Form der Kommunikation war Voraussetzung, um ein gemeinsames Verständnis für die Anforderungen an den geplanten Dienst zu entwickeln. Dabei konnte die Machbarkeit schon früh nachgewiesen werden. Es war jedoch noch eine lange Wegstrecke, um die Einbindung der erforderlichen Daten mit den Datenbereitstellern zu gewährleisten.

2 Eine konkrete Anwendung

Die von der Firma M.O.S.S. realisierte Flächenplausibilisierung basiert unter anderem auf dem Produkt „3D-Gebäudemodelle LoD1“ der Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV). Die Daten werden als Basis der Grundflächenberechnung sowie der Ermittlung der Anzahl der Stockwerke genutzt. Zum „3D-Gebäudemodell LoD1“ heißt es in der Produktspezifikation der AdV:
Ein 3D-Gebäudemodell ist ein digitales, numerisches Oberflächenmodell der im AAA-Modell der AdV definierten Objektbereiche Gebäude und Bauwerke. Der Gebäudegrundriss wird grundsätzlich der amtlichen digitalen Liegenschaftskarte entnommen.
Gerade die Übereinstimmung der Gebäudegrundrisse mit der amtlichen digitalen Liegenschaftskarte sichert die Eignung der Lösung für die Verwendung zur Plausibilisierung von Flächenangaben bei der Immobilienbewertung.
In einem ersten Schritt wird über die Eingabe der Adresse das Grundstück auf einer Flurkarte identifiziert und farblich markiert. Der Nutzer kann die Lage des Gebäudes auf dem Grundstück überprüfen, korrigieren und ggf. noch weitere Gebäude-Teile hinzufügen.
Wird die Auswahl bestätigt, errechnet das System auf Basis der amtlich eingemessenen Grundfläche des Gebäudes und dem 3D-Stadt-Modell die Anzahl der Geschossebenen. Daraus errechnet sich automatisiert die oberirdische Geschossfläche.
Die Informationen zur Unterkellerung und zum Dachausbau werden automatisch aus der Bezeichnung des „Objekttyps“ übernommen.
Als Ergebnis werden die Bruttogrundfläche (als Basis für die Ermittlung der Herstellungskosten) und die Wohn-/Nutzfläche (als Basis für die Berechnung des Ertragswertes) angezeigt. Dabei

erfolgt die Umrechnung der Grundfläche in Wohn- bzw. Nutzfläche über einen - objektarttypischen - Koeffizienten, der voreingestellt oder auch manuell eingegeben werden kann.
Alle Wertparameter werden in einer Tabelle dargestellt und können individuell verändert oder ergänzt werden.
Neben der Plausibilisierung der Angaben über die Oberfläche des Dienstes ist die Dokumentation wesentlicher Bestandteil der Anwendung. Als Ergebnis wird ein PDF Dokument mit interaktiver 3D Gebäudedarstellung, Lagebeschreibung des Objektes (Karte) und den Bewertungsparametern erstellt und dem Gutachten beigefügt.
Diese teil-/automatisierte Prüfung der Flächen erspart viel Zeit und bietet gleichzeitig die Möglichkeit, die im Gutachten in Ansatz gebrachten Flächen nachzuvollziehen und zu überprüfen.

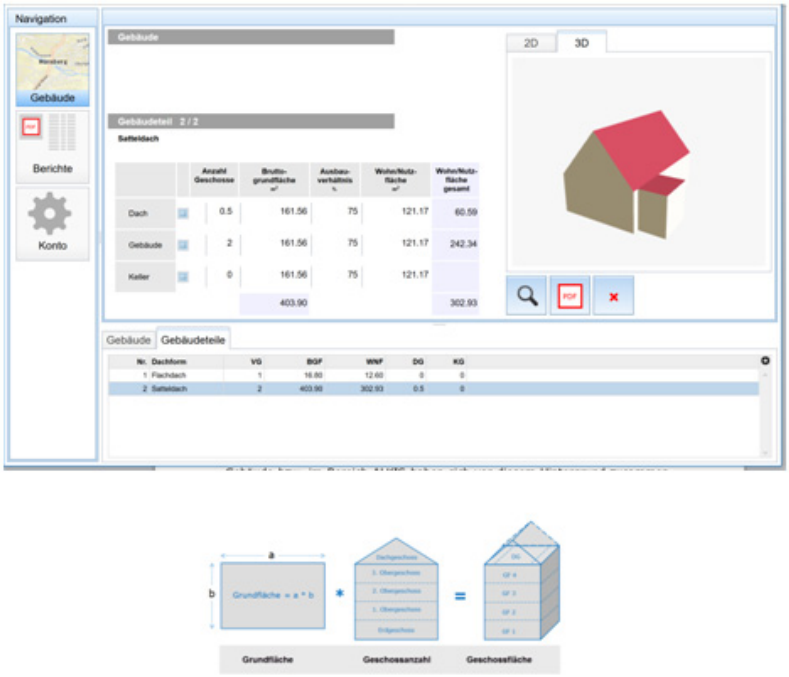


Abb. 1: Muster einer Flächenplausibilisierung

Weitere Anwendungen sind geplant: Analysen zum energetischen Zustand, Hochwasserrisiko, Solarstrompotential etc. ermöglichen eine noch differenziertere Bewertung.

3 Die Hürden

Man kann also feststellen: Der künftige Anwendungsbereich von Geoinformationen ist sehr groß. Das gilt aber auch für die Hürden, die noch bestehen.
Alle Geoinformationen müssen wesentliche Anforderungen erfüllen:

1. Elektronische Verfügbarkeit
2. Hohe Qualität
3. Ausreichende Anzahl
4. Aktualität
5. Quellennachweis
6. Datenschutz

Bereits der erste Punkt führt dazu, dass viele Projekte scheitern, obwohl die nachfolgenden Punkte grundsätzlich erfüllt werden können.
Ein Beispiel:
Die Gutachterausschüsse in Deutschland erhalten alle Kaufpreise und eine Vielzahl von Informationen zum Kaufgegenstand. Diese Informationen werden in einer Kaufpreissammlung abgelegt. Zu dieser Datensammlung haben Gutachter*innen mit berechtigtem Interesse und entsprechendem Nachweis Zutritt.
Die Suche ist in vielen Fällen mühsam und kann nur auf Einzelobjektebene mit ausgewählten Kriterien erfolgen. Ein direkter Zugriff auf die Datenbank wird verwehrt. Die Suchergebnisse werden auf eine - vom Gutachterausschuss festgelegte - Anzahl eingeschränkt. Die Handhabung ist von Bundesland zu Bundesland unterschiedlich und das Kostenmodell ist nicht einheitlich. Bekommt der Gutachter in Berlin die Informationen kostenlos, muss er in Hamburg dafür erheblich in die Tasche greifen.

4 Zusammenfassung und Ausblick

„Die Bundesregierung hat dieses Jahr zum “ Jahr der Geowissenschaften“ erklärt. Die Politik hat die Geoinformation inzwischen in ihrer Bedeutung erkannt.“ Das Zitat stammt vom 16. Januar 2002. Am 18. Mai 2017 – fünfzehn Jahre später - hat der Deutsche Bundestag den Entwurf eines Ersten Gesetzes zur Änderung des E-Government-Gesetzes verabschiedet. Dieser setzt die Forderungen nach einem verbindlichen Open-Data-Gesetz um. Er gibt den Behörden der unmittelbaren Bundesverwaltung auf, die bei ihnen vorhandenen elektronischen Daten zukünftig zu veröffentlichen.
Das Prinzip der offenen Daten "Open Data" bekommt lt. Bundesministerium des Innern weltweit eine immer größere Bedeutung. Die Verfügbarkeit von Daten wird zunehmend auch zu einem wichtigen Wirtschaftsfaktor. Die Daten werden somit zu einem Teil einer modernen Infrastruktur.
Der Entwurf der Bundesregierung wird von der Finanz- und Immobilienwirtschaft unterstützt. Er muss jedoch auch zur Anwendung kommen. Wenn Geoinformationen nicht rechtzeitig zur Verfügung gestellt werden, führt das dazu, dass künftig Lösungsmodelle auf der Basis von verfügbaren Daten (z.B. aus Internetportalen) entwickelt werden.
Bei der Entfaltung des Nutzens von Geoinformation ist zu beachten, dass die Immobilienbewertung, wie viele andere Branchen zur Zeit unter dem Schlagwort „Digitalisierung“ einen Wandel vollzieht, für den heute niemand voraussagen kann, wie genau die Prozesse der Zukunft aussehen werden. Welche Rolle Geoinformationen in digitalisierten Immobilienbewertungsprozessen spielen, hängt nicht zuletzt auch davon ab, dass intensive Kommunikation und Zusammenarbeit zwischen den Branchen stattfindet und die Beteiligten schnell und flexibel aufeinander reagieren.

Kollaborative Analyse und Entscheidungsfindung an Tangible User Interfaces – Herausforderungen für die Modellierung und Visualisierung von Geodaten

JOCHEN SCHIEWE[1]

Zusammenfassung: Im Zuge der Digitalisierung von Prozessen bis hin zu ganzen Städten kommen Beteiligungsverfahren mit unterschiedlichen Nutzergruppen eine immer größere Bedeutung zu. Bei solchen kollaborativen Workshops spielen häufig auch Geoinformationen bei der Analyse und Entscheidungsfindung eine wichtige Rolle. Dieser Beitrag stellt ein größeres Bürgerbeteiligungsprojekt vor, das mit Hilfe von Tangible User Interfaces realisiert worden ist. Aus den Erfahrungswerten dieses Vorhabens werden Besonderheiten und weiter zu lösende Herausforderungen im Hinblick auf die Modellierung von Geoinformationen sowie die Gestaltung von Kartengraphik und Interaktionen mit dieser Benutzerschnittstelle herausgestellt.

1 Einleitung

Geoinformationen spielen bei der Analyse und Entscheidungsfindung in unterschiedlichen fachlichen bzw. politischen Fragen eine zentrale Rolle. Vor dem Hintergrund komplexer Problemstellungen bzw. des Willens zur Einbindung von Akteuren und Bürgern gewinnt die kollaborative Bearbeitung immer mehr an Bedeutung.

Unter **kollaborativer Bearbeitung** wird hier verstanden, dass mehrere Personen gemeinsam an allen Arbeitsschritten einer Aufgabe tätig sind – im Gegensatz zur kooperativen Herangehensweise, bei den Personen jeweils eigene Teilaufgaben bewältigen. Die Kollaboration kann räumlich getrennt oder konzentriert sowie zeitlich synchron (d.h., parallel) oder asynchron (d.h., zeitversetzt) erfolgen. Typische Beispiele für kollaborative Prozesse unter Berücksichtigung von Geoinformationen sind Bürgerbeteiligungsverfahren, die als Präsenz-Workshops (räumlich konzentriert sowie synchron) oder über ein webbasiertes GIS (räumlich getrennt sowie asynchron) durchgeführt werden können. Im Zuge der Digitalisierung von Prozessen bis hin zu ganzen Städten spielen solche Beteiligungsverfahren – mit unterschiedlichen Nutzergruppen (Bürger, Experten oder eine Kombination hiervon) – eine immer größere Rolle.

In Abhängigkeit vom Kollaborationsformat sind angepasste graphische und interaktive Gestaltungsformen zu Analyse- und Visualisierungszwecken bereitzustellen. In diesem Beitrag soll ein bisher selten genutztes, aufgrund fallender Kosten aber immer attraktiveres Mittel zur Darstellung und Interaktion vorgestellt werden – das sogenannte **Tangible User Interface** (TUI). TUIs sind anfassbare Benutzerschnittstellen, die eine Interaktion zwischen Mensch und Maschine durch physische Objekte (und i.d.R. zusätzlich auch durch Berührungen) erlauben.

Dieser Beitrag stellt ein größeres Bürgerbeteiligungsprojekt vor, das mit Hilfe von Tangible User Interfaces realisiert worden ist – es handelt sich dabei um die Suche nach Flüchtlingsunterkünf-

[1] HafenCity Universität Hamburg, Lab for Geoinformatics and Geovisualization (g2lab), Überseeallee 16, D-20457 Hamburg, E-Mail: jochen.schiewe@hcu-hamburg.de

ten in Hamburg (Kapitel 2). Aus den Erfahrungswerten dieses Vorhabens werden Besonderheiten und weiter zu lösende Herausforderungen im Hinblick auf die Modellierung von Geoinformationen sowie die Gestaltung von Kartengraphik und Interaktionen herausgearbeitet (Kapitel 3).

2 Projekt *FindingPlaces*

2.1 Hintergrund

Der große Strom von Flüchtlingen sorgte 2015/2016 für großen politischen Handlungsdruck. Allein in der Freien und Hansestadt Hamburg wurde im März 2016 ein Bedarf an Folgeunterkünften (d. h. für ca. drei bis fünf Jahre Wohndauer nach den ersten sechs Monaten in einer Erstaufnahmeeinrichtung) für rund 40.000 Geflüchtete geschätzt. Die Notwendigkeit, eine sozialverträgliche Lösung für Flüchtlinge und Bewohner gleichermaßen zu finden, wurde zum vorherrschenden und öffentlich sehr emotional und kontrovers diskutierten Thema.

Vor diesem Hintergrund entstand das Bürgerbeteiligungsprojekt *FindingPlaces* zwischen der Stadt Hamburg und der HafenCity Universität Hamburg (URBACT 2018). Die grundlegende Idee bestand darin, die Expertise und Ortskenntnis der Bürgerinnen und Bürger zu nutzen, um weitere, bisher unbeachtete Flächen für Unterkünfte zu identifizieren. Darüber hinaus konnte auch ein stadtweiter Diskurs im Kontext von verschiedenen Interessen und rechtlichen Planungsanforderungen angeregt werden. Insgesamt besuchten rund 400 Teilnehmer die 35 Workshops, dabei wurden 160 Flächen identifiziert, von den 44 die weitere Begutachtung durch die Behörden passierten.

Detaillierte Informationen zum Projekt geben SCHIEWE et al. (2017) und MENSING (2017).

2.2 Workshops

Zwischen Mai und Juli 2016 wurden für jeden der sieben Hamburger Bezirke bis zu sieben Workshops angeboten, zu denen sich jeweils maximal 30 Personen anmelden konnten. Die Teilnehmer diskutierten unter professioneller Moderation jeweils zwei Stunden lang über potenzielle Areale.

Nach einer Einführung und einem Gesamtüberblick über Zielsetzung und Ablauf des Workshops erfolgte die kollaborative Arbeit an zwei Tischen mit Tangible User Interfaces. Am **„Bezirks-Tisch“** wurde ein Luftbild sowie Zusatzdaten des für den Workshop relevanten Bezirks auf einen 2 m × 2 m großen Tisch projiziert. Da aus Zeitgründen immer nur Teile eines Bezirkes untersucht werden konnten, erfolgte hier in gemeinsamer Diskussion die Auswahl von drei bis vier Fokusräumen.

Am **„Grundstücks-Tisch“** derselben Größe konnten alle Attribute eines dargestellten Flurstücks und der zugehörigen Umgebung aus einer Datenbank abgefragt und an einem zusätzlichen Monitor dargestellt werden. Auf dieser Informationsbasis wurde diskutiert, ob dieses Flurstück aus Sicht der Teilnehmer für den Bau einer Flüchtlingsunterkunft geeignet erschien.

Die dabei ausgewählten Flächen wurden an die Verwaltung weitergeleitet, der innerhalb von zwei Wochen eine Einschätzung zu jedem Vorschlag erstellte und im Internet veröffentlichte.

2.3 Tangible User Interfaces

Zentrales Element der Workshops war ein interaktives Stadtmodell der Stadt Hamburg, welches am *CityScienceLab* in Kooperation mit dem Lab for Geoinformatics und Geovisualization (*g2lab*) der HafenCity Universität Hamburg innerhalb von nur zwei Monaten entwickelt wurde. Das Modell basiert auf einem *CityScope*, einem am Media Lab des Massachusetts Institute of Technology (MIT) entwickelten, kollaborativen Tangible User Interface (MIT, 2018). Das ursprüngliche *CityScope* wurde für das Projekt *FindingPlaces* grundlegend überarbeitet. Es wurde vergrößert, um Funktionen erweitert und insbesondere an ein webbasiertes Geographisches Informations-System (GIS) angebunden.

Die Projektion der Karteninformationen auf die beiden Tischoberflächen erfolgte durch je zwei links bzw. rechts darüber angebrachte Beamer (Abb. 1). Die Auswahl zu bearbeitender Regionen am „Bezirks-Tisch“ erfolgte über die Registrierung aufgelegter „Suchrahmen“ (Abb. 2).

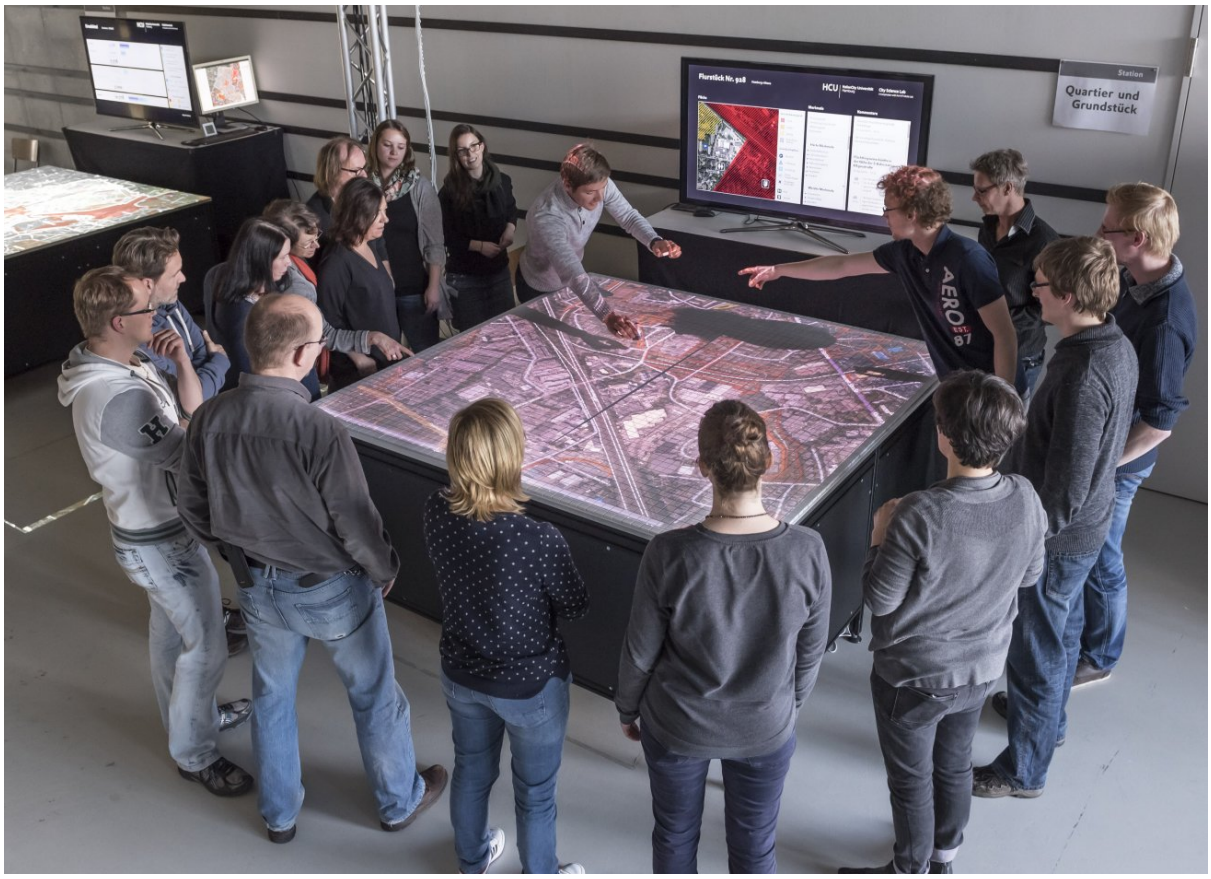


Abb. 1: Workshop-Gruppe am „Grundstücks-Tisch“

Abb. 2: Auswahl einer zu bearbeitenden Region mit „Suchrahmen“

Die Oberfläche des „Grundstück-Tisches“ wurde mit 55×55 quadratischen Kunststoffplättchen belegt, die an einer gewünschten Rasterzelle entfernt und durch einen an der Unterseite kodierten LEGO^{TM}-Baustein ersetzt werden konnte (Abb. 3). Vier unter der Plexiglasscheibe angebrachte Kameras filmten die Bewegungen der Steine bzw. den Status jeder Rasterzelle. Die Steine konnten zwei einfache Funktionalitäten auslösen. Ein „Suchstein“ sorgte für die Abfrage von Attributen des Grundstücks, während unterschiedliche „Datensteine“ die Anzahl von Personen repräsentierten, die nach Ansicht der Teilnehmer in Unterkünften an der betreffenden Stelle untergebracht werden konnten. Diese Zahlen wurden nach Bestätigung in die Datenbank übernommen.

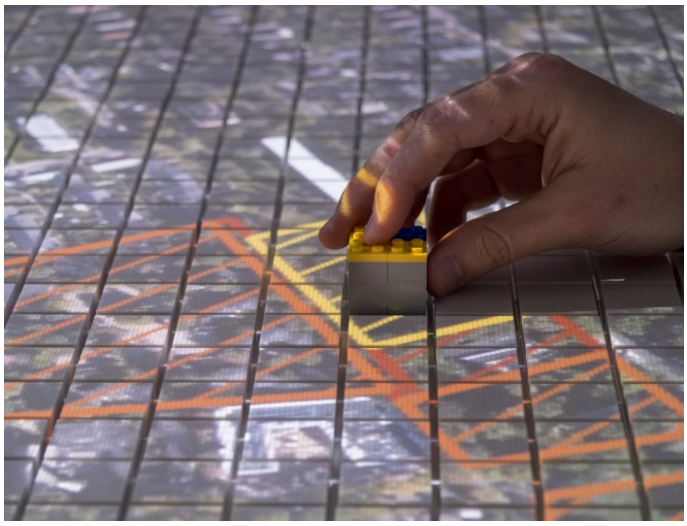


Abb. 3: Platzierung eines „Such- bzw. Datensteins“

Die Datenbasis für das Projekt *FindingPlaces* stammte vom Landesbetrieb Geoinformation und Vermessung Hamburg (LGV). Die grundsätzliche Eignung der Flächen, die sich aus rund 30 betrachteten Kriterien ergab, wurde aus Komplexitätsgründen vorab nur grob qualitativ über die Einteilung in drei, farblich markierte Eignungsklassen visualisiert.

3 Herausforderungen

Der Einsatz von Tangible User Interfaces zur kollaborativen Analyse und Entscheidungsfindung weist einige Besonderheiten sowie noch zu lösende Probleme im Zusammenhang mit der Modellierung von Geoinformationen sowie der Gestaltung von Kartengraphik und Interaktionen auf.
Der **Multi-User-Ansatz** in einer prinzipiell räumlich konzentrierten und zeitlich synchronisierten Umgebung zeichnet sich je i.d.R. durch eine Gruppe mit heterogenem Vorwissen hinsichtlich der eigentlichen Thematik sowie der Nutzung von kartographischen Darstellungen aus. Zum anderen gibt es immer Ablenkungspotenzial, das die prinzipielle Synchronität der Bearbeitung gefährden kann.
Konsequenter Weise ist eine sehr klare und möglichst einfache Gestaltung der – auf ein Minimum beschränkten – Kartengraphik und Interaktionselemente notwendig. Hierzu gehören z. B. einfache Datenklassifikationen (im Idealfall nach dem Ampelprinzip), ständig lesbare Legenden, aber auch Doppelkodierungen (z. B. über Farbe und Text). Im Projekt *FindingPlaces* wurde diese Notwendigkeit rechtzeitig erkannt und unter Mithilfe eines Kommunikationsdesigners und mit der Durchführung mehrerer Testläufe berücksichtigt.
Die hier vorgestellte kollaborative Bearbeitung stellt ferner einen Sonderfall einer **multimedialen Darbietung** dar, bei der kartographische und sprachliche Elemente gleichberechtigt zusammengeführt werden müssen. Die Schwierigkeit besteht darin, dass nur ein gewisser Anteil an Erläuterungen planbar ist. Zum großen Teil erfolgt eine freie Diskussion, bei der die kartographische Darstellung „nur“ eine unterstützende oder reaktive Funktion hat. Es hat sich im o. g. Projekt gezeigt, dass zumindest für den planbaren Teil der Einführung ein detailliertes und mehrfach zu testendes „Drehbuch“ mit Unterstützung professioneller Moderatoren oder Dramaturgen einen großen Mehrwert darstellt.
Schließlich folgen die kartographischen Elemente in der kollaborativen Umgebung einem **Multipurpose-Prinzip**. Zum einen dienen sie dem Zweck der Motivation – sowohl die großflächige, variable und detaillierte Darstellung der eigenen Umgebung, als auch die Möglichkeit der Interaktion über die Such- und Datensteine hatte einen offensichtlich positiven Einfluss auf die Mitwirkung. Zum anderen dient die Darstellung sowohl reinen Präsentations-, als auch einfachen Explorations-Funktionen im Sinne des *Map Use Cube* (MACEACHREN & KRAAK 1997): Im Rahmen des oben beschriebenen Bürgerbeteiligungsverfahren stand die reine Präsentation der Grundstücke mit ihrer topographischen Umgebung sowie der zugehörigen Eignungsklassen im Fokus. Es gibt inzwischen aber auch weitergehende Entwicklungen, z. B. für interaktive Erreichbarkeitsanalysen mit Hilfe des TUI (SCHABER, 2017), die auf eine explorative Analyse abzielen. Diese Aufgaben sind eher im Experten-Bereich angesiedelt, nichtsdestotrotz sollte auch hierfür ein Satz ein vorgefertigten, aufgabenorientierten Gestaltungen bereitgestellt werden, um einen gezielten und schnellen Einsatz zu ermöglichen.

4 Zusammenfassung & Schlussfolgerungen

Tangible User Interfaces in verschiedenen Ausprägungen haben inzwischen eine größere Marktreife und günstigere Preise erzielt und sind somit für kollaborative Anwendungen mit Unterstützung von Geodaten und Karten interessant geworden. Es ist zu erwarten, dass sich die Breite an

Nutzern und Anwendungen vergrößern wird, und somit auch notwendige Weiterentwicklungen im Kontext der Modellierung von Geoinformationen sowie der Gestaltung von Kartengraphik und Interaktionen eine wichtige Rolle im Bereich der Forschung und Entwicklung einnehmen werden.
Diese Weiterentwicklungen bedingen einen stark interdisziplinär ausgerichteten Ansatz, der in Zukunft auch verstärkte Untersuchungen der Usability sowie User Experience bedingt, um den propagierten Mehrwert nachzuweisen bzw. den initialen Kostenaufwand zu rechtfertigen.

5 Literaturverzeichnis

MACEACHREN, A.M. & KRAAK, M.-J., 1997: Exploratory cartographic visualization: advancing the agenda. In: Computers and Geosciences, 23(4), 335-343.

MIT, 2018: http://cp.media.mit.edu/city-simulation. letzter Zugriff 15.01.2018.

MENSING, T., 2017: Entscheidungsunterstützung in der partizipatorischen Stadtplanung durch CityScopes. Kartographische Nachrichten, **66**(4), 219-220.

SCHABER, L., 2017: Entwicklung und Visualisierung dynamischer und interaktiver Impact-Modelle auf CityScopes. Unveröffentlichte Master-Thesis im Studiengang Geomatik an der HafenCity Universität Hamburg.

SCHIEWE, J., KRÖGER, J., ZOBEL, K. & MENSING, T., 2017: Aufbau, Nutzung und Evaluation eines interaktiven Stadtmodells für die partizipative Flächensuche für Flüchtlingsunterkünfte. GIS Science, **30**(2), 31-39.

URBACT, 2018: http://urbact.eu/finding-places. letzter Zugriff 15.01.2018.

Community mapping in the commercial environment using HERE base maps as an example

GEORG HERRIG[1]

Abstract: The HERE map, used by millions worldwide through mobile apps, social networks and car navigation systems, combines information of more than 80.000 sources. One of those sources is our community of map enthusiasts that voluntarily provide edits in their local areas. Learn more on how HERE builds its maps and how we engage with communities worldwide to keep our map fresh and build the reality index.

1 About HERE

HERE empowers the world with location data. The HERE Map is used by millions worldwide through mobile apps, social networks and in 4 out of 5 cars equipped with a navigation system.
We are committed to delivering the highest quality location data as the go-to destination for location services for any need. From healthcare, transportation and logistics to infrastructure, manufacturing and financial services.
The core of our solutions is a rich, always evolving, map that captures the complexity of the world in the greatest possible detail. Offering nearly one thousand attributes, we create flexible solutions that enrich customer products with location intelligence and help their applications reach the scalability.

2 HERE map updating process

HERE combines information from more than 80.000 sources including a global fleet of mapping vehicles, GPS probe data, vehicle sensors, local authorities and a global community of map editors.

3 HERE Community program

Our global community of map editors improve their cities and neighborhoods by contributing to the HERE Map through HERE Map Creator or a Map feedback form. Contributors around the world can add, delete or update roads, places (e.g. business details, categories) and house numbers to improve navigation, attract new customers and much more.
Members of our global community can access HERE Map Creator through iOS and Google Play apps and mapcreator.here.com.

[1] HERE Technologies, Friedrich-Ebert-Straße 55, D-45127 Essen, E-Mail: georg.herrig@here.com

4 Community inputs are an essential element to keep the HERE map fresh

4.1 Benefits for the HERE map

The map feedback provided by the HERE map community is essential to reflect the real world in our map. Our community allows us to:

- Detect and implement immediate changes by bringing in local knowledge
- Confirm changes detected by other sources
- Integrate features where no other or source is available
- Detect changes before they actually happen
- Alert on temporary changes like road works, speed traps, fuel prices.[2]

4.2 Types of map contributors

The global HERE map community is built on different types of map contributors, depending on their level of mapping knowledge, their engagement and their contribution goal. Through different activities, our community managers engage with them on a personal level, going from teaching digital mapping tools to students to working with advanced users on enhancing our map feedback tools.

Why do these groups of people contribute to the HERE map?

- Map enthusiasts: individuals that like mapping and want to make sure that the map reflects the real world. They take pride in the fact that the map is used by so many people worldwide
- Students: to get a practical knowledge of digital mapping or on development topics (location related APIs and SDKs) in the scope of their academic curriculum.
- Locals and tourism: to promote local highlights, make sure that emergency services arrive on time
- Small businesses: to keep their info like opening hours and contact information updated
- Car drivers or other people that use our maps: they want to make sure that an error they have noticed gets corrected

5 HERE Community Management

Around the world, more than 40 HERE community managers take care of the users of more than 100 countries.

[2] These attributes cannot be changed via Map Creator. However, users do alert us on their changes via other communication means and our community managers do the integration for them.

They are in contact daily with the Map Creator users to discuss features, answer their questions and provide personal support. We organize regular trainings and events, mapathons, contests, and smaller meetups.

5.1 Ambassador Programs

Our best editors (called Map Masters or HERE Ambassadors) get specific moderation rights allowing them to moderate their own and other user edits. They also benefit from accelerated validation and integration.
They represent the HERE Map Creator community. Our ambassadors spread the word, by actively recruiting new users and working on community engagement. They provide valuable feedback on new features and implementations and work with the developer team on future enhancements. They help us out by providing user support via our forums, when responding to other users' questions, but also by co-organizing our local events. Their presence is essential to make our meetings a success.

5.2 Academic Programs

HERE has a long-lasting relationship with universities and high schools all over the world. Together with the teachers we setup a real academic training program around digital mapping (combining theory and practical sessions). The students mostly work around updating their local area, but can also focus on specific tasks like working on specific map attributes like speed limits or geometry. Thanks to our partner Mapillary (www.mapillary.com), a tool that allows imagery collection, we can provide the students a complete theoretical and practical understanding of the complete mapping cycle, from collection over editing into validated changes.
We are also expanding our academic approach with a focus on future developers: students gain practical development experience by working on location related applications using the HERE APIs and SDKs.
3 audiences:

- High schools
- Universities
- Young entrepreneurs

Since 2017, we benefit from back-up from the European Union through the Erasmus+ program. We also partner with inter(national) organizations like iGEO where digital mapping challenges via Map Creator become part of the official program.

5.3 Associations

A variety of associations use Map Creator to update the map in remote or rural areas or to focus on specific attributes. For instance, sports associations add bike pathways or NGO's focus on local wheelchair access.
We have more long term agreements with Auto-Moto clubs for continuous imagery collection (using our partner tool Mapillary) and then integration of the detected changes into Map Creator, often also working as ambassadors

5.4 Local Government

We work with local government to enhance the map for citizens and tourists. Locals update the map with shops, emergency access roads etc. In some cases, they can also use our APIs to create their own application and provide it to the citizens.
If they dispose of large volume datasets, they can also use our supplier portal – upload.here.com – to upload these to the HERE map for further integration.

5.5 Partner Programs

Here has been partnering with 3rd parties to ensure a feedback loop for our map users (e.g. Ford, BMW), to update the map on specific attributes or to provide edits in restricted areas that are inaccessible for non-authorized persons. Other parties, like insurance companies, use our map data to verify road attributes and correct them if needed through Map Creator, creating a real win-win situation.

6 Community statistics 2017

Number of Monthly active users globally: more than 5.000

Number of edits globally: 24.7M edits, 1.3M added kilometers

Number of edits in Germany: 700.000

Number of active users in Germany: 5.000

Number of communities: 316 active communities globally, 45 in Germany

Events in Germany: User Meetings in north, east, west and south Germany, Map Master meeting, Map Job campaigns, participation in conferences

Map Masters: 11 Map Masters in Germany

Academic Programs: 5 Universities with more than 200 students in Germany

7 Conclusion

HERE Technologies combines decades of map building expertise with high quality sources, valuable customer feedback and advanced engineering capabilities to orchestrate the publishing of the freshest map on the market. We detect new road geometry and verify changes in real world conditions quickly, building a digital representation of the world – a “Reality Index”.
This Reality Index is only possible thanks to the valuable inputs of a highly motivated community of map enthusiasts that reflect real time and hyperlocal changes in our maps. To foster this community, HERE has set up a global program, where community managers work with students and mentors to build the map of the future.

Wie die Elektromobilität unsere Infrastrukturplanung verändert

FRANK STEINBACHER[1] & MARTIN VEH[1]

Zusammenfassung: Die Elektrifizierung unserer Mobilität wird uns in ähnlicher Weise in unseren gesellschaftlichen aber auch infrastrukturtechnischen Strukturen verändern wie das Smartphone oder der Internetanschluss. Diese neue Technologie ermöglicht erstmal eine Sektorenkopplung zwischen Infrastruktur die uns die Mobilität ermöglich und der Kraft die für unsere Mobilität benötigt wird. Die elektrifizierte Mobilität wird ihre Wertschöpfung nicht in der Mobilität alleine haben, sondern in den Dienstleistungen und Produkten die mit Mobilität verknüpft werden.

1 Infrastruktur für die elektrifizierte Mobilität

Der Markt Zusmarshausen (Bayern, Schwaben, A8), die Sortimo International GmbH, die Steinbacher-Consult Ingenieurgesellschaft und die eLoaded GmbH haben im Rahmen der Gewährung von Zuwendungen für die Errichtung öffentlich zugänglicher Ladeinfrastruktur für Elektrofahrzeuge mit der Planung und baulichen Umsetzung der derzeit weltweit größten Ladeinfrastruktur in 2017 begonnen. Am 22.12.2017 wurde die rechtliche Baureife für dieses regionsübergreifende Infrastrukturprojekt erlangt. Die Strahlwirkung des Standorts liegt nach derzeitiger Bewertung auf Grund seiner Größe bei über 100km Reichweite und auf Grund seiner technologischen Ansätze auf höchstem internationalem Niveau (NewYorkTimes, CleanTechnica, Artikel 13.09.2017). Im Januar 2018 beginnt die bauliche Umsetzung und Errichtung von ca. 150 DC-Ladepunkten mit Ladeleistungen von mind. 50kW bis 350kW pro Ladepunkt durch ein neuartiges DC-Bussystem.

1.1 Ziel des Projektes

Der Standort ist als zentraler „Hub“ mit regionaler Infrastrukturvernetzung konzeptioniert und soll die Synergie und die damit verbundenen Potentiale unterschiedlicher Nutzergruppen der elektrifizierten Mobilität technologisch, wirtschaftlich und sozio-psychologisch eröffnen. Das Zusammenspiel kommunaler und privatwirtschaftlicher Belange im Bereich der Elektromobilität wird den öffentlichen Verkehrssektor und eine Wirtschaftsregion neu definieren – das Vorhandensein von multifunktionaler Ladeinfrastruktur wird zu einem der zentralen Standortfaktoren der Zukunft. Mobilität wird ein wichtiger Wirtschaftsfaktor und eine Region wird energieeffizient, klima- und umweltverträglich sowie wirtschaftlich neu geprägt.
Die Elektrifizierung der Mobilität führt nur dann zu einer signifikanten CO2-Reduzierung, wenn der für den Antrieb benötigte Strom regenerativ erzeugt wird. Vor diesem Hintergrund ist die Verknüpfung der Ladesteuerung mit dem regionalen regenerativen Energieangebot Bestandteil

[1] Steinbacher-Consult Ingenieurgesellschaft mbH&Co. KG, Richard-Wagner-Str.6, D-86356 Neusäß
und eLoaded GmbH, Feldstr. 1b, A-6020 Innsbruck,
E-Mail: f.steinbacher@steinbacher-consult.com, info@elaoded.eu

der Untersuchungen. Durch die Einbeziehung eines regionalen Energieversorgungsunternehmens wird einerseits die netztechnische und energiewirtschaftlich-regulatorische Begleitung des Projektes sichergestellt. Gleichzeitig werden über die vertrieblichen Ansätze kommerzielle Anreize und Geschäftsmodelle zur nachhaltigen Nutzung und Optimierung der Anwendung entwickelt.

Die Forschungsdissemination soll über die Beteiligten die Elektrifizierung der Mobilität und der damit verbundenen Technologien unter dem Gesichtspunkt der Skalierbarkeit positiv entwickeln und effizienter gestalten. Gerade international soll die Vorreiterrolle einer der Kernkompetenzen der Deutschen Wirtschaft, nämlich der komplexen Infrastrukturplanung sowie dessen Betrieb, herausgearbeitet werden und sich somit als Leitmarkt positionieren. Insgesamt sind 12 deutsche und zwei österreichische Unternehmen maßgeblich an der Projektgestaltung beteiligt.

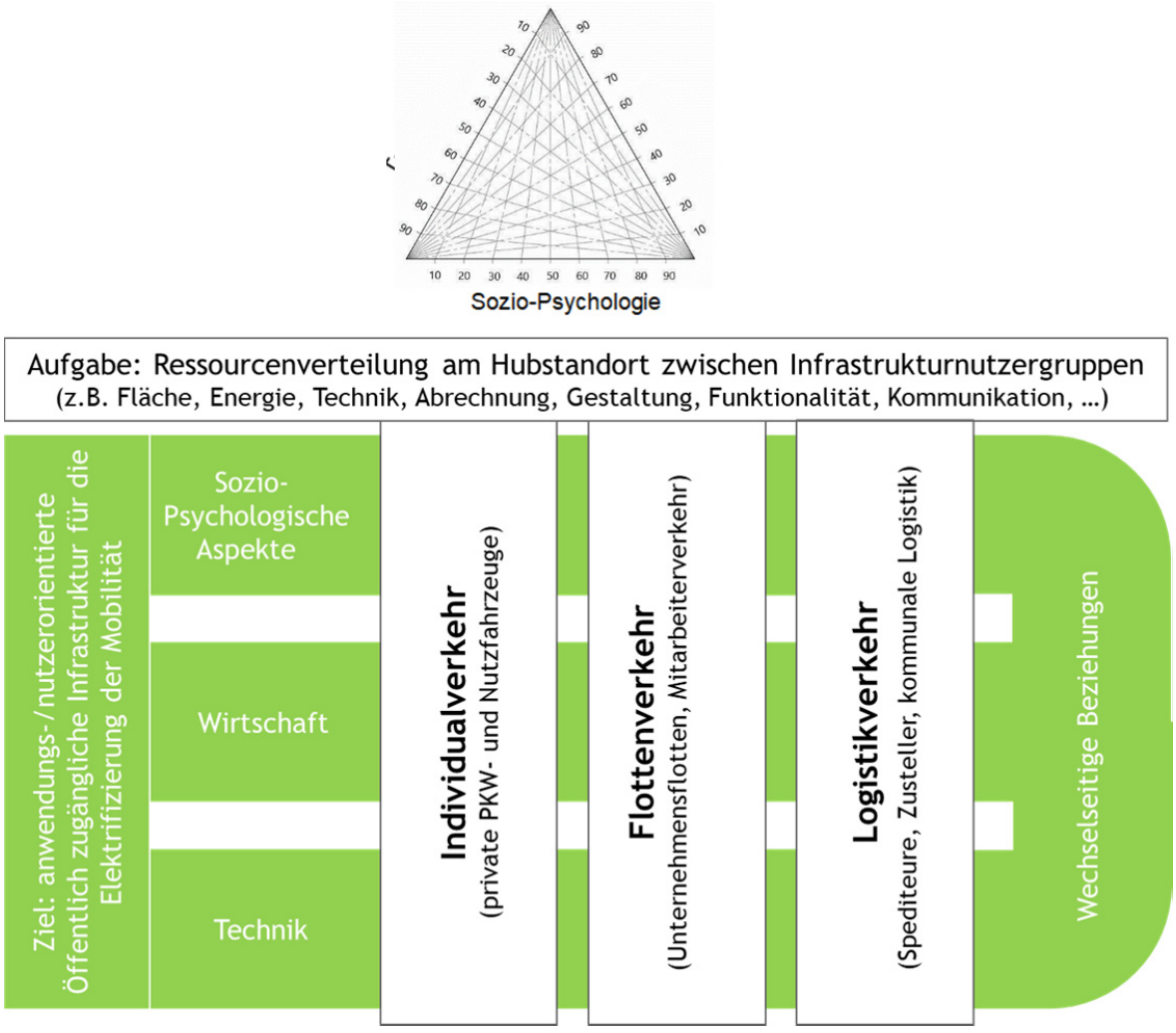


Abb. 1: Methodischer Aufbau einer Standortbewertung

1.2 Methodik der Infrastrukturplanung und -ausrichtung

Damit öffentliche Ladeinfrastruktur nutzer- und anwendungsorientiert betrieben werden kann, besteht ein Zielkonflikt zwischen technischen Aspekten, wirtschaftlichen Aspekten und sozio-psychologischen Aspekten für die unterschiedlichen Nutzergruppen. Am Hub-Standort in Zusmarshausen werden die Nutzergruppen Individualverkehr, Flottenverkehr und Logistikverkehr im Zuge der Auslegung zugelassen (kein ÖPNV, kein LKW-Verkehr). Diese Nutzergruppen unterliegen jedoch unterschiedlichen Anforderungen an die Infrastruktur. Es wird geklärt, in wie fern diese Nutzergruppen durch ihren unter-schiedlichen Anspruch an die Ressource „Ladeinfrastruktur“ im Konflikt stehen oder durch eine gezielte Betriebsgestaltung und ein Hubmanagement sogar eine Ressourcennutzung bei Fragen wie z.B. Flächenbedarf, Energiebedarf, Technikbedarf und -auslegung, Abrechnungssysteme, Gestaltung, Funktionalität, Kommunikation, usw. ermöglichen. Der Hub-Standort in Zusmarshausen ist derzeit als das weltweit größte Living-Lab öffentlicher Ladeinfrastruktur zu verstehen.

1.3 Standortaufbau

Alle Bausteine des Projektes zielen mit ihrer Modularität, Skalierbarkeit, Erweiterbarkeit, Adaptionsfähigkeit und der Verknüpfung und Einbindung in eine öffentliche Verkehrsinfra-struktur darauf ab, den neuen Markt und die neue Technik für eine elektrifizierte Mobilität als eine nachhaltige, langfristige und wirtschaftliche Infrastruktur zu entwickeln und etablieren.

2 Fazit & Ausblick

Der USP dieses Projektes ist, auch international in Alleinstellung, in seiner Hubfunktion, Größe, energetischer Vernetzung, der Nutzerstrukturen, seiner technologischen Funktionen und den ortsgebundenen und mobilen Kundendienstleistungen zu finden. Dies ist weltweit einzig-artig. Die multifunktionale Nutzungsform ermöglicht z.B. neue **Ansätze in der Wirtschaftlich**keitsbetrachtung, der Energieverteilung oder technologischen Skalierbarkeit. Die zentrale Hubstruktur ermöglicht eine Bündelung von Ressourcen im öffentlichen Raum. Unterschiedliches Verkehrsverhalten und die daraus resultierenden Bedürfnisse an Lade**ninf**rastruktur werden zusammengeführt und die geschaffene Infrastruktur effizienter genutzt. Gleiches gilt für die zum Laden benötigte Energie und deren Bereitstellung: Primärenergie-einsparung durch Reduktion von Transformations- und Verteilverlusten, Abwärmenutzung von Leistungselektronik, Transformation und Verteilung zur Gebäudeklimatisierung oder die Bündelung erneuerbarer Energieressourcen.

Standortfarbmarkierung von oben nach unten im Lageplan:

- *Ladeinfrastruktur für Logistikprozesse (~20 Ladepunkte)*
- *Kurzzeitladen für Dienstleistungsinfrastruktur und Behindertenfahrzeuge (~20 Ladepunkte)*
- *Supraladen bis 350kW für Schnellstversorgung (~24 Ladepunkte)*
- *Mittelfristladen für Kundendienstleistungen (~40 Ladepunkte)*
- *Langzeitparken für Mitarbeiter (~30 Ladepunkte)*
- *Langzeitparken für Park&Charge&Ride (~10 Ladepunkte)*

Abb. 2: Standortaufbau Ladehub Zusmarshausen, A8

Softwaregestützte Fortführung der Straßenbestandsdaten im GIS

RALF BEHRENS[1]

1 Kurzbeschreibung

Mit der Einführung eines Straßenmanagements und der damit einhergehenden Bestands- und Zustandserfassung rückt sowohl mit der ersten Wiederholungsaufnahme als auch der täglichen Weiterentwicklung der Straße die Frage nach den Mechanismen zur korrekten Datenfortführung in den Fokus. Eine hundertprozentige Automatisierung des Fortführungsfalles in puncto Bestand und Zustand ist sicherlich kaum zu erreichen, doch können geeignete Werkzeuge den Anwender soweit unterstützen, dass diese Aufgabe überschaubar zu bewältigen ist. Der Vortrag zeigt hierzu erste Lösungsansätze auf.

2 Softwaregestützte Fortführung der Straßenbestandsdaten – grundlegend für ein umfängliches Erhaltungsmanagement

Steht der Zustand der Straßen im Fokus, kommt sicherlich die Frage auf, wie die Haushaltsmittel zur Straßenerhaltung bestmöglich einzusetzen und welche Maßnahmen zu welchem Zeitpunkt idealerweise anzuwenden sind. Somit wird ein Erhaltungsmanagement für die Straßen angestrebt. Damit soll grundsätzlich die Entwicklung der Straßenflächen bezüglich Bestand und Zustand aufgezeigt werden, um so mit geeigneten Hilfsmitteln und fachlichem „Know How“ die richtigen Schlüsse ziehen zu können.

2.1 Umfassende Bestandsaufnahme

Als Ausgangbasis ist dafür eine umfassende Bestandsaufnahme notwendig, welche neben einem sogenannten Knoten-Kanten-Modell (kurz KKM) auch den Bestand (im kommunal-städtischen Bereich als Realflächen erhoben) und die dazu zugehörigen Zustandsdaten umfasst. So kommt es zu einer ersten, aktuellen Momentaufnahme, die eine gute Übersicht über den Status Quo gibt. Doch der Straßenbestand entwickelt sich stetig weiter. So ändert sich der Bestand, wenn Maßnahmen z. B. zur Verbreiterung eines Radweges durchgeführt werden oder eine Ausfallstraße um eine Busspur erweitert wird. Zu diesem Zeitpunkt stellt sich zum ersten Mal die Frage nach der Fortführung des Straßendatenbestandes: Wird der alte Radweg einfach gelöscht und der neue an dessen Stelle gesetzt? Oder besteht die Möglichkeit, den alten Radweg zu archivieren, so dass dieser über einen Zeitstempel im alten Zustand über eine zeitliche Abfrage wieder aufzeigbar wäre. Fragen wie: „Welche Ausprägung hatte die Kreuzung vor der Umbau- und Sanierungsmaßnahme“ oder „Welchen Verlauf nahm der Radweg vor der Neuplanung“ können durch Abfragen zum Zeitraum über den Zeitstempel einfach beantwortet werden.

[1] IP SYSCON GmbH, Tiestestr. 16-18, D-30171 Hannover, E-Mail: ralf.behrens@ipsyscon.de

2.2 Datenfortführung

Die Fortführung spielt auch beim Aspekt der Zustandsbewertung eine wichtige Rolle. Eine Prognose bezüglich der Zustandsentwicklung eines Straßenabschnittes lässt sich aufstellen, wenn Messwerte aus mehreren Kampagnen über konstante Zeitintervalle vorliegen. Dabei ist entscheidend, dass die Historie dieser Messwerte komplett und durchgehend geführt und den richtigen Flächen exakt zugewiesen ist, damit diese dann den Algorithmen hinsichtlich einer Zustandsprognose zugeführt werden.
In einer ersten Umsetzung für die manuelle Fortführung durch den Fachanwender am GIS- Arbeitsplatz können einige Werkzeuge wie folgt angewendet werden:

- wird eine neue über eine alte Fläche gezeichnet, so erfolgt das Ausstanzen der älteren Fläche
- letztere wird in einem Archiv fortgeführt und erhält ein „Lebenszeitende“
- wohingegen die neue Fläche ein „Lebenszeitbeginn“ eingetragen bekommt
- ferner wird an der archivierten Fläche der Bezug zur „Nachfolgefläche“ hinterlegt
- Zustandsdaten verbleiben an der archivierten Fläche
- an der neuen Fläche sind in der Regel neue Zustandserhebungen durchzuführen

Ein vergleichbares Vorgehen trifft u.a. sowohl auf Straßenabschnitte (KKM), als auch auf linien- oder punkthafte Straßenobjekte zu. So ist es möglich, eine komplette Historie der Bestands- und Zustandsentwicklung zu dokumentieren und als Grundlage für ein Erhaltungsmanagement heranzuziehen.
Eine sich direkt anschließende Fragestellung ergibt sich beim Blick auf Massenfortführungsdaten, die z. B. von einem Dienstleister bereitgestellt werden. Hierbei ist zu meistern, dass die Änderungen am Bestand und Zustand nicht nach und nach vom GIS-Anwender eingebracht und individuell überwacht werden, sondern Massendaten in einem Prozessschritt inhaltlich und fachlich korrekt übernommen werden sollen. Um auch diesen Vorgang abbilden zu können, wurde gemeinsamen mit einem führenden Datendienstleister ein Konzept erarbeitet, welches in seiner Umsetzung die Massendaten über die richtige Verschlüsselung fachlich korrekt in die Daten einsortiert und die GIS-Objekte aktualisiert. Zugrunde liegt eine Matrix, welche die verschiedenen Sachverhalte, die aufeinander treffen können, beschreibt und den daraus resultierenden Importvorgang anstößt. Dabei sind Fälle aufgeführt wie: (a) nur Sachdatenänderungen, (b) nur Zustandsänderungen, (c) neue Geometrie mit Sachdaten und (d) neuer Verlauf des Abschnitts (mit zugehörigen Flächen). Ausgehend von diesen Merkmalen liefert das System zukünftig automatisierte Abläufe, um ebenso die Fortführung von Massendaten am Bestand und am Zustand zu dokumentieren.

3 Fazit

Um ein aussagekräftiges Erhaltungsmanagement durchzuführen, ist das Vorhalten einer historisierten Bestands- und Zustandsentwicklung maßgeblich zielführend. Dabei spielt es grundsätzlich keine Rolle, ob Fortführungen in Einzelschritten, manuell und nach Bedarf erfolgen, oder ob Fortführungen zu festen Intervallen massenhaft dem Datenbestand zugeführt werden. Wichtig ist allein die Auswertbarkeit im Rahmen der anschließenden Prozesse – entweder durch das Fachpersonal oder durch Softwareroutinen zur Entwicklungsprognose der Straßen und einhergehender Maßnahmen und Kosten.

Einsatz von Real-Time Daten zur Optimierung des Winterdienstes in Bayern

MARC KLEEMANN[1] & MARKUS STECKER[2]

Zusammenfassung: Zur Optimierung der Steuerung von Winterdiensteinsätzen in Bayern werden verschiedene Real-Time Daten eingesetzt. Das Spektrum reicht von Daten des Deutschen Wetterdienstes (Niederschlagsradar Informationen, Mess- und Prognosewerte der Straßenwetterstationen, Unwetterwarnungen), Webcambilder bis hin zu den Positions- und Aktivitätsdaten der Winterdienstfahrzeuge. Der Vortrag stellt dar, welche Herausforderungen an die Architektur einer auf Real-Time Daten basierten Anwendung gestellt werden, und wie diese umgesetzt worden sind, um die Sicherheit auf Bayerns Straßen zu erhöhen.

1 Einleitung und Problemstellung

Für den Straßenbetriebsdienst in Bayern stellen die Wintermonate aufgrund der extremen Witterungsverhältnisse eine besondere Herausforderung dar. Entscheidend für die Aufrechterhaltung der Funktionsfähigkeit der Straßen ist die Organisation eines vorrausschauenden, effektiven Winterdienstes. Eine der Hauptaufgaben besteht darin, die 600 öffentlichen und 700 privaten Einsatzfahrzeuge so zu koordinieren, dass ein möglichst reibungsloser Verkehrsfluss auf den 22.000 Straßenkilometern sichergestellt wird.

Die Bayerische Straßenbauverwaltung unternimmt alle Anstrengungen, den Winterdienst laufend zu optimieren, um die Verkehrssicherheit und den Verkehrsfluss bestmöglich zu gewährleisten. Entsprechend hoch waren die Anforderungen an ein neues digitales Winterdienstmanagementsystem. Unter anderem galt es, eine Vielzahl relevanter Daten aus unterschiedlichen Quellen zusammenzuführen, zu analysieren und darzustellen. Ein Prognose-Modul sollte – möglichst präzise und weit im Vorfeld – Auskunft darüber geben, auf welchen Streckenabschnitten mit Glätte zu rechnen ist. Zudem musste ein Rechte- und Rollenkonzept etabliert und leicht bedienbare Komponenten, sowohl für die Einsatzzentrale als auch für die Bordcomputer der Fahrzeuge, entwickelt werden.

Die Autobahndirektion Nordbayern ist eine zentrale Landesbehörde des Freistaates Bayern. Ihr ist die Landesbaudirektion angegliedert. Die Zentralstelle für den Straßenbetriebsdienst an der Autobahndirektion unterstützt die Autobahn- und Straßenmeistereien in Bayern bei der Durchführung des Winterdienstes.

[1] con terra GmbH, SDI & E-Government, Karl-Liebknecht-Str. 5, 10178 Berlin, E-mail: m.kleemann@conterra.de

[2] con terra GmbH, SDI & E-Government, Martin-Luther-King-Weg 24, 48155 Münster, E-mail: m.stecker@conterra.de

2 Lösung

Zur Winterdienst-Saison 2015/2016 ist das neue Bayerische Winterdienstmanagementsystem (WDMS-BY) in die Pilotphase gestartet. Genutzt wird es zunächst von insgesamt 16 Straßen- und 5 Autobahnmeistereien sowie der Metropolregion Nürnberg. Herzstück des WDMS-BY ist ein webbasiertes Geoinformationssystem, welches unter Federführung der Autobahndirektion Nordbayern in Zusammenarbeit mit der con terra entwickelt wurde.

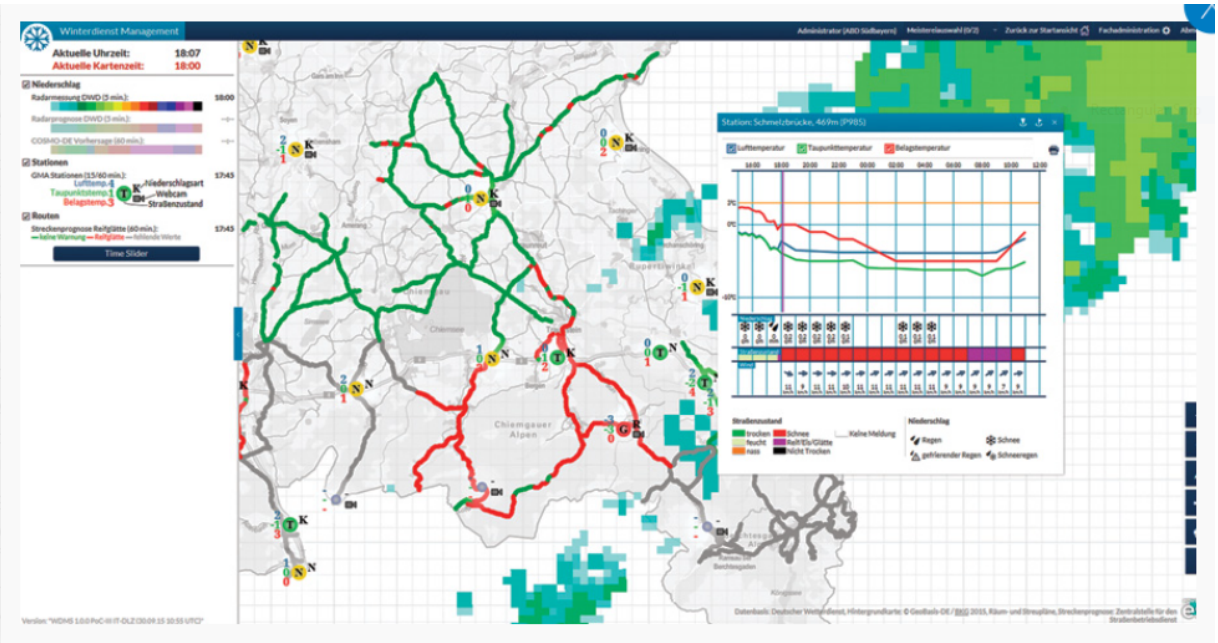


Abb. 1: Prognose zur streckengenauen Auskunft zur Glättebildung in den nächsten 18 Stunden

Das System bündelt die permanent einlaufenden Daten, analysiert sie und fügt sie zu einem aktuellen Gesamtbild zusammen. Einbezogen werden hierfür unter anderem Niederschlags- und Vorhersagedaten des Deutschen Wetterdienstes sowie die Daten von zurzeit 250 Glättemeldeanlagen in Bayern. Hinzu kommen Bilder der WebCams, die alle 10 Minuten von den Glättemeldeanlagen übermittelt werden. Das Besondere des Systems ist jedoch das Alarmierungs- und Prognose-Modul, welches streckengenau Auskunft darüber gibt, wo in den nächsten 18 Stunden mit Glätte auf der Fahrbahn zu rechnen ist. Die Einsatzleitung kann so effektiver als bisher die Einsätze der Räum- und Streufahrzeuge koordinieren und für sichere Straßenverhältnisse sorgen.
Eingesetzte Technologie:

- ArcGIS for Server: Bereitstellung von Kartendiensten und GIS-Funktionen
- FME Server, FME Desktop: ETL-Prozesse zur lfd. Verarbeitung komplexer Echtzeitdaten
- map.apps: Geo-Apps erstellen, organisieren und betreiben
- security.manager: Absicherung von Diensten und Rechte- und Rollenmanagement
- GDI/INSPIRE-Dienste.

3 Fazit & Ausblick

Durch die Verwendung eines modernen GI-Systems auf Basis der Technologie von Esri und con terra stehen den Winterdiensteinsatzleitern nun erstmalig alle entscheidungsrelevanten Daten übersichtlich auf einer Oberfläche zur Verwendung. In Pilotbetrieben hat sich das System erfolgreich bewährt und leistet nun einen wertvollen Beitrag zu Erhöhung der Verkehrssicherheit auf den Straßen Bayerns.
Weiterführende Informationen: siehe www.abdnb.bayern.de und https://www.conterra.de/wdms.

AUTORENVERZEICHNIS

A

ALLROGGEN, NIKLAS 34
ARENS, MICHAEL 341, 363
ATKINSON, COLIN 829

B

BAHR, THOMAS 71
BARETH, GEORG 285
BEARDSLEY, PAUL 743
BEHNISCH, MARTIN 658
BEHRENS, RALF 921
BEIL, CHRISTOF 717
BERGER, CHRISTIAN 806
BEYER, FLORIAN 521
BIRKENBEUL, HENNING 20
BLEISCH, SUSANNE 141
BLOMLEY, ROSMARIE 408
BOERNER, RICHARD 76
BORGMANN, BJÖRN 363
BRAUSEWETTER, PATRICK 187
BREITWIESER, FLORIAN 778
BRENNER, CLAUS 376
BREUER, MICHAEL 710
BRINKHOFF, THOMAS 164
BRUNN, ANSGAR 450
BURKART, ANDREAS 274
BUSCH, STEFFEN 86, 376

C

CAKIR, SERCAN 177
CARSTENSEN, DORIS 877
CAVEGN, STEFAN 431
CHANG, JUIWEN 213
CHRISTEN, MARTIN 328, 624
CLEMENT, URS 328
COORS, VOLKER 559, 646, 670

D

DIEDRICH, ERHARD 547
DONAUBAUER, ANDREAS 199, 886, 888
DUTTA, AVISHEK 633

E

EISFELDER, CHRISTINA 679
ELFOULY, MOSTAFA 199
ELTNER, ANETTE 389, 573
ENGELS, JOHANNES 633
EUGSTER, HANNES 431
EWALD, WOLFGANG 857

F
FILIPPOVSKA, YEVGENIYA 317
FIUTAK, GEORG 888
FRECKMANN, PETER 219

G
GEHRUNG, JOACHIM 341
GERHARDS, ROLAND 633
GERKE, MARKUS 242
GITAHI, JOSEPH M. 633
GHIMIRE, PRASHANT 633
GLÄßER, CORNELIA 465
GLATZ, WOLFGANG 71
GLOCK, CLEMENS 20
GRAF, MARCO 624
GRENZDÖRFFER, GÖRRES 521, 851
GÜLCH, EBERHARD 151
GÜMGÜMCÜ, HASRET 743
GÜNTHER-DIRINGER, DETLEF 534

H
HAALA, NORBERT 701
HAHMANN, STEFAN 187
HAHMANN, THOMAS 547
HAHN, MICHAEL 633
HANEL, ALEXANDER 355
HEBEL, MARCUS 341, 363
HECHT, ROBERT 658
HEIPKE, CHRISTIAN 584, 611
HENNIG, SABINE 294
HENZE, FRANK 103
HERRIG, GEORG 913
HINZ, STEFAN 34, 49, 767
HODAM, HENRYK 117, 129
HOEGNER, LUDWIG 76
HOLDENER, DAVID 755
HU, YUNCAI 232
HUGENTOBLER, URS 729
HURTIG, THOMAS 306

I
ISYA, NOORLAILA H. 483
IWASZCZUK, DOROTA 729

J
JACKISCH, CONRAD 34
JACOBSEN, KARSTEN 422
JÜRGENS, CARSTEN 508
JUTZI, BORIS 408, 767

K
KADA, MARTIN 317
KARUT, IZABELA 317
KASS, STEVE 794

KAZIMI, BASHIR 597
KELER, ANDREAS 397
KELLER, SINA 34, 49
KELM, THORSTEN 787
KERSTEN, THOMAS 471
KIM, HYUN OK 679
KLEEMANN, MARC 923
KLEIN, ULRIKE 787
KLINGBEIL, LASSE 274
KOLBE, THOMAS H. 199, 717, 831, 835, 886, 888
KOPF, STEPHAN 187
KORKUT, SAFAK 141
KOUKOFIKIS, ATHANASIOS 670
KRALISCH, LISA-MARIE 710
KRÖHNERT, MELANIE 573
KRÜGER, MARC 787
KUENZER, CLAUDIA 679
KUNZ, KAROLIN 232
KUNZ, PETER 1
KUPER, PAUL V. 823
KUTZNER, TATJANA 827, 835

L

LATIFI, HOOMAN 151
LAUMER, DANIEL 743
LAUPHEIMER, DOMINIK 701
LIEBEL, LUKAS 790
LILIENTHAL, HOLGER 710
LINDNER, CLAUDIA 117, 129
LINDSTAEDT, MAREN 471
LOESCH, BENJAMIN 431
LUDWIG, CHRISTINA 794
LUSSEM, ULRIKE 285

M

MAAS, HANS-GERD 389, 573
MACHL, THOMAS 857
MAIWALD, FERDINAND 103
MAIER, PHILIPP M. 49
MANDL, PETER 778
MARX, CAROLINE 888
MATHEUS, ANDREAS 864
MEHLTRETTER, MAX 584
MEINERT, AXEL 187
MEYER, ADRIAN 328
MEYER, M. FABIAN 508
MINK, ROBIN 633
MOSLEHIRAD, OMIDREZA 729
MUDERSBACH, CHRISTOPH 787
MWANGI, PATRICIA W. 691

N
NEBIKER, STEPHAN 431
NETZEL, FABIAN 787
NGUYEN, UYEN D. X. 611
NIEMEIER, WOLFGANG 242, 483
NIESSNER, ROBIN 767
NOACK, PATRICK O. 840
NOVAK, DAVID 431

O
ORTWEIN, ANNETTE 117, 129
OTT, RICHARD 903

P
PETEINATOS, GERASSIMOS 633
PFEFFER, INGO 508
PFISTER, ANGELA 787
PIPELIDIS, GEORGIOS 729
PITER, ANDREAS MAXIMILIAN 749
PIVETEAU, NADINE 798
POLITZ, FLORIAN 177, 597
PREHOFER, CHRISTIAN 729
PREUß, NADINE 536
PREY, LUKAS 258

Q
QUEHL, JANNIK 376

R
RASCHER, UWE 274
RAUBAL, MARTIN 798
RIEDEL, ANIKA 242, 483
RIEDEL, BJÖRN 242, 483
RIENOW, ANDREAS 117, 129
RIESE, FELIX M. 34
RIFFLER, MICHAEL 794
RÖDER, MARIUS 151
ROSSEM VAN, RUUD 862
ROTTENSTEINER, FRANZ 611
RUHDORFER, ROLAND 802

S
SARDEMANN, HANNES 389, 573
SCHÄR, MARKUS 761
SCHÄR, PASCAL 431
SCHARLACH, HOLGER 306
SCHIEWE, JOCHEN 213, 646, 907
SCHILLING, HENDRIK 767
SCHINDLER, KONRAD 743
SCHITO, JORAM 798
SCHMIDHALTER, URS 232, 258
SCHNEIDER, DANILO 103
SCHNEIDER, SVEN 559
SCHOLZ, JOHANNES 778

SCHOTTEN, MAIKE 646
SCHOTTMÜLLER, CHRISTINE 242
SCHREINER, VERA 465
SCHULTZ, JOHANNES 117, 129
SEFERCIK, UMUT 422
SELG, FABIAN 117, 129
SESTER, MONIKA 597
SIEGMANN, BASTIAN 274
SIHOMBING, ROSANNY 646
STECKER, MARKUS 923
STEINBACHER, FRANK 917
STILLA, UWE 76, 341, 355, 363

T

TEUCHER, MIKE 465
THIEMANN, FRANK 177
TRAUN, CHRISTOPH 461
TSCHIRSCHWITZ, FELIX 471

U

V

VEH, MARTIN 917

W

WEGNER, JAN DIRK 743
WEIBEL, ROBERT 798
WEIDNER, UWE 58
WEIGEL, MICHAEL 258
WEINMANN, MARTIN 58, 408
WEINMANN, MICHAEL 408
WENDT, TIM 658
WICHMANN, ANDREAS 317
WILKE, NORMAN 274
WILLKOMM, PHILIPP 888

X

XU, YUSHENG 76

Y

YAO, ZHIHANG 831

Z

ZANDA, ADRIANA 141

PROGRAMMKOMITEE

Abmayr, Thomas
Aumann, Gabriele
Baltrusch, Sven
Bareth, Georg
Brunn, Ansgar
Buchroithner, Manfred
Burghardt, Dirk
Crom, Wolfgang
Dickmann, Frank
Edler, Dennis
Grenzdörffer, Görres
Gülch, Eberhard
Günther-Diringer, Detlef
Haala, Norbert
Haunert, Jan-Henrik
Hochschild, Volker
Höfle, Bernhard
Illert, Andreas
Jacobsen, Karsten
Kaufmann, Viktor
Kersten, Thomas
Kleinschmit, Birgit
Kolbe, Thomas H.
Kröger, Johannes
Krzystek, Peter
Matthias, Ekkehard
Meng, Liqiu
Moser, Jana
Nebiker, Stephan
Petzold, Bettina
Przybilla, Heinz-Jürgen
Schiewe, Jochen
Schneider, Danilo
Sester, Monika
Sörgel, Uwe
Stilla, Uwe
Theile, Erik
Weidner, Uwe
Weinmann, Martin
Wetzel, Hans-Ulrich
Wolff, Kirsten