

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



Band 26

2017



FH·W·S

Hochschule
für angewandte Wissenschaften
Würzburg-Schweinfurt

Vorträge

**37. Wissenschaftlich-Technische
Jahrestagung der DGPF**

8. – 10. März 2017
in Würzburg

*Kulturelles Erbe erfassen und bewahren -
Von der Dokumentation zum virtuellen Rundgang*

ISSN 0942-2870

Thomas P. Kersten, Hrsg.

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



Band 26

2017

Beiträge

37. Wissenschaftlich-Technische
Jahrestagung der DGPF

8. – 10. März 2017
in Würzburg

*Kulturelles Erbe erfassen und bewahren -
Von der Dokumentation zum virtuellen Rundgang*

ISSN 0942-2870

Thomas P. Kersten, Hrsg.

ISSN 0942-2870

Publikationen der
Deutschen Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
Band 26, 536 S., Hamburg 2017
Hrsg.: Thomas P. Kersten

© Deutsche Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
München 2017

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

Gemeinsam laden die Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation sowie die Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt zur Jahrestagung vom 7. bis zum 10. März 2017 nach Würzburg ein.

Wir haben für die gemeinsame Tagung als Motto gewählt:

KULTURELLES ERBE ERFASSEN UND BEWAHREN – VON DER DOKUMENTATION ZUM VIRTUELLEN RUNDGANG

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

Würzburg blickt auf eine lange Standorttradition zurück. Wenn auch die erste urkundliche Erwähnung als „Castellum Virteburch“ erst von 704 stammt, so lassen doch Funde einer Wehranlage von um 1000 v. Chr. auf eine längere Geschichte schließen. Die erste Ansiedlung am Schalksberg geht sogar auf die Zeit des Pleistozän (mind. 10.000 v.Chr.) zurück. Damit bildet der Standort Würzburg geradezu eine Herausforderung für das Motto unserer Jahrestagung. Wir freuen uns sehr auf den Eröffnungsvortrag von Herrn Univ.-Prof. Mag. Dr. Michael Doneus, Institut für Urgeschichte und Historische Archäologie, TU Wien, der uns einen Einblick in die Nutzung der Methoden der Photogrammetrie, Fernerkundung und Geoinformation im Rahmen der Archäologie geben wird.

Mit dem Sieg der Franken über die Alemannen 496 findet in Würzburg die fränkische Landnahme statt und die Burg über dem linken Mainufer wird Herzogshof der Merowinger. Im Jahr 779 wird erstmals der Würzburger Markt beschrieben und der Weinbau in der Stadt erwähnt. Das Bild um die Stadt wird noch heute durch den Weinbau geprägt, der besonders durch das Würzburger Bürgerspital, erstmals 1317 urkundlich erwähnt, vorangetrieben wurde. Im Mittelalter war die Stadt ein bedeutendes wirtschaftliches, geistliches und hoheitliches Zentrum. Im Jahr 1576 wurde durch Fürstbischofs Julius Echter von Mespelbrunn das Juliusspital gestiftet, das heute u.a. ein Krankenhaus, ein Seniorenstift und ein Weingut umfasst. Das sei an dieser Stelle erwähnt, da in der Zehntscheune des Weinguts Juliusspital unsere festliche Abendveranstaltung der DGPF-Jahrestagung stattfinden wird.

Das akademische Leben in Würzburg zog nach einer ersten Universitätsgründung 1402 durch eine Wiedergründung im Jahre 1582 durch Julius Echter von Mespelbrunn erneut in die Stadt ein. Der Fürstbischof wurde aber auch durch die Landesvertreibung von Protestanten und die intensivierete Hexenverfolgung bekannt. Neben der Universität wurde 1797 das Collegium musicum academicum und 1971 die Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt (FHWS) gegründet, deren Wurzeln bis 1807 zurückreichen. Der Vorstand der DGPF hat sich sehr über die Zusage vom DGPF-Arbeitskreisleiter Ansgar Brunn und der Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, gefreut, die diesjährige DGPF-Jahrestagung zu beherbergen.

Die Hochschule Würzburg-Schweinfurt ist auf die beiden Standorte Würzburg und Schweinfurt verteilt, an denen über 9000 Studierende eingeschrieben sind. Die Wahl der Tagungsstätte am Standort Würzburg erschien aufgrund der Verkehrsanbindung ideal, da sie nur 450 m vom Hauptbahnhof Würzburg entfernt liegt. Der Studienbereich Geo der FHWS ist Mitnutzer des Gebäudes, in dem Conrad-Röntgen 1895 die nach ihm benannten Röntgenstrahlen entdeckte. Ein Besuch in der Gedenkstätte ist während der Tagung sehr zu empfehlen!

Dem Aufruf zur Einreichung von Beiträgen folgten Autoren von 77 wissenschaftlichen Beiträgen, von denen 52 auf der Tagungs-CD als Publikationen enthalten sind. Insgesamt sind 127 Autoren in den Beiträgen aufgeführt, wobei einige Standorte aus ihren Arbeits- und Forschungsgruppen mehrere Beiträge eingereicht haben. Ein besonderer Dank sei hier auch den Reviewern für die Begutachtung der eingereichten Beiträge und dem Editor für die Zusammenstellung des Tagungsbands ausgesprochen.

Da der Vorstand der DGPF die wissenschaftlich-technische Kooperation mit den Firmen intensivieren will, begrüßen wir besonders den diesjährigen Zuspruch der 22 Aussteller bei der Fachfirmenausstellung. Im Anwenderforum geben Vorträge von den ausstellenden Fachfirmen einen kurzen Überblick über das Firmenprofil und die Hauptproduktlinien und stellen dann im Detail die Exponate und verschiedene Anwendungsmöglichkeiten der Produkte vor.

Ein weiterer Arbeitsschwerpunkt des Vorstands liegt in der Förderung der Aus- und Weiterbildung. Dazu wird am Dienstag vor der Eröffnung der Jahrestagung erstmals ein Studentenforum stattfinden, bei dem wir den Studierenden durch Vorträge von Firmenvertretern einen Einblick in die berufliche Praxis vermitteln wollen. Dies steht neben dem bereits gut eingeführten Wettbewerb um den Karl-Kraus-Nachwuchspreis. Bei dem stellen bis zu zehn ausgesuchte Absolventen aus Deutschland, Österreich und Schweiz, die von den Betreuern als sehr gute empfohlen wurden, in Konkurrenz ihre Arbeiten an Postern persönlich vor. Diese Posterveranstaltung bildet auch immer eine interessante Plattform für das Kennenlernen von Absolventen und Firmenvertretern. Das Weiterbildungsforum findet mit zwei Vortragsblöcken am Mittwochvormittag statt.

Neben diesen Aktivitäten flankieren Treffen zum Doktorandenkolloquium der Deutschen Geodätischen Kommission (DGK)/DGPF und der Abteilung Geoinformatik der DGK die Veranstaltung in Würzburg.

Der Vorstand der Gesellschaft bedankt sich bei allen Sponsoren und Ausstellern für die Unterstützung der Tagung. Im Namen des Vorstands der DGPF sowie des örtlichen Mitveranstalters laden wir herzlich nach Würzburg ein und wünschen Ihnen spannende Vorträge und Fachdiskussionen.

Prof. Dr.-Ing. Uwe Stilla
Präsident



Prof. Dr.-Ing. Ansgar Brunn

FH·W·S

Hochschule
für angewandte Wissenschaften
Würzburg-Schweinfurt

INHALTSVERZEICHNIS

FERNERKUNDUNG - SCHWERPUNKT VEGETATION

- G. GRENZDÖRFFER & U. SCHRÖDER:
UAS unterstützte Vegetationserfassung tidebeeinflusster
Makrophytenbestände in der Unterelbe. 1

SENSOREN UND PLATTFORMEN - WELTRAUM

- K. JACOBSEN:
Verbesserung der Geometrie von Satellitenbildern durch Höhenmodelle. 13
- R. HOHN, C. WILLIGES, M. HARLANDER, J. HINGER, H. ROSSMANN & R. REULKE:
Verifizierung der Fokalebene (FPA) von Sentinel-4. 22

OPTISCHE 3D MESSTECHNIK I

- T. KERSTEN & M. LINDSTAEDT:
Mit Kamera, Maßstab und Laptop – Best Practice zur Erstellung
von 3D-Modellen für eine Museumsausstellung
am Beispiel vietnamesischer Kulturobjekte. 35
- J. BRAUNECK & H. FELDHAUS:
Hyperspektrale und photogrammetrische Datenaufnahme
am ehemaligen Zisterzienserkloster Wörschweiler. 45
- F. MAIWALD & D. SCHNEIDER:
Erfassung und Analyse dreidimensionaler Deformation an Gemälden
zur Unterstützung des Restaurierungsprozesses. 55

GEOINFORMATIK I

- M. CHRISTEN & M. FEHR:
Web-Visualisierung komplexer 3D-Stadtmodelle mit pyRT. 67
- A. VOLKMANN:
Geoinformatik und Archäologie - GIS basierte Kulturraumstudie antiker
und frühmittelalterlicher Besiedlungsmuster am mittleren Main. 76

FERNERKUNDUNG - SCHWERPUNKT SENSOREN UND FUSIONEN

- S. BALTRUSCH & J. PETERS:
mySentinel – Fernerkundungsdaten
für das Bundesland Mecklenburg-Vorpommern. 90
- C. VILLAMIL LOPEZ & U. STILLA:
Fusion of High Resolution SAR Images with Optical and GIS Data
and its Application to Change Detection in Urban Areas. 97
- S. KELLER, A. BRAUN, S. HINZ & M. WEINMANN:
Investigation of the Potential of Hyperspectral EnMAP Data
for Land Cover and Land Use Classification. 110

BILDANALYSE UND BILDVERSTEHEN I

- H. MAYER, H. HUANG, A. KUHN, M. MICHELINI,
W. NGUATEM, L. ROTH, M. SCHMITZ & M. DRAUSCHKE:
Bilder vom Boden und vom UAV: Orientierung, dichte 3D Rekonstruktion
und funktionale Modellierung. 122
- D. IWASZCZUK, T. KOCH & U. STILLA:
Innenraumrekonstruktion aus semantisch angereicherten 3D Punkten
und Linien. 129

GEOINFORMATIK II

- J. OEHRLEIN, T. C. VAN DIJK & J.-H. HAUNERT:
Gleichwertige Ziele in dynamischen Navigationskarten. 138
- L. RUNCEANU, S. BECKER, N. HAALA & D. FRITSCH:
Indoor Point Cloud Segmentation for Automatic Object Interpretation. 147

SENSOREN UND PLATTFORMEN - LUFTBILD UND TERRESTRISCH

- A. HANEL & U. STILLA:
Calibration of a Vehicle Camera System
with Divergent Fields-of-view in an Urban Environment. 160
- G. KEMPER:
Hochauflösende Multisensor-Systeme. 170

OPTISCHE 3D MESSTECHNIK II

- M. WEINMANN, C. MALLET & B. JUTZI:
Involving Different Neighborhood Types for the Analysis
of Low-level Geometric 2D and 3D Features and
their Relevance for Point Cloud Classification. 179
- D. RUEß & R. REULKE:
Kalibrierung von Kameras mit allgemeinen, stetigen
und nicht-symmetrischen Kameramodellen. 192

GEOINFORMATIK III

- P. BURGGRAFF, E. KNEIPS, J. SCHULTHEIß & M. TEMPEL:
KuLaDig – eine Kooperationsplattform zur Erfassung
und Bewahrung der Kulturlandschaft und ihrer Entwicklung. 203

RADARFERNERKUNDUNG UND LASERSCANNING

- M. F. MEYER, I. PFEFFER & C. JÜRGENS:
Die automatisierte Auswertung von LiDAR-Daten in der Archäologie. 216
- T. BAHR, N. HOLZER & P. SMITH:
Geiger-Mode - LiDAR-Technologien der nächsten Generation. 223
- C.-H. YANG & U. SÖRGEL:
Urban Monitoring by 4D Change Detection
Using Multi-temporal SAR Images. 233

BILDANALYSE UND BILDVERSTEHEN II

- C. KRUSE, H. NEUBERGER, F. ROTTENSTEINER,
T. HOBERG, M. ZIEMS, J. HUTH & C. HEIPKE:
Automatische Detektion von Bombenkratern in Kriegsluftbildern
mittels markierter Punktprozesse. 245
- B. BORGMANN, M. HEBEL, M. ARENS & U. STILLA:
Konzept zur Gefährdungserkennung im städtischen Verkehrsraum
durch Personendetektion in MLS-Punktwolken. 262

STANDARDISIERUNG UND QUALITÄTSSICHERUNG

- A. BRUNN, H. WEICHELT & R. REULKE:
Bildqualität von optischen Fernerkundungsdaten. 276
- S. BALTRUSCH & R. REULKE:
DIN-Normungsarbeit in der Photogrammetrie und Fernerkundung –
Stand und Perspektiven. 280

3D STADTMODELLE

- I. HIJAZI, T. KUTZNER & T. H. KOLBE:
Use Cases and their Requirements on the Semantic Modeling
of 3D Supply and Disposal Networks. 288
- J. GEHRUNG, M. HEBEL, M. ARENS & U. STILLA:
Filtern bewegter Objekte aus mobilen LiDAR-Daten
auf Grundlage volumetrischer Modelle. 302
- Z. YAO & T. H. KOLBE:
Dynamically Extending Spatial Databases to support
CityGML Application Domain Extensions using Graph Transformations. 316

AUS- UND WEITERBILDUNG

- A. ORTWEIN, B. KREBS, J. MEHRING, J. SCHULTZ & A. RIENOW:
Neue Dimensionen der Erdbeobachtung im Schulunterricht:
Raumfahrt, 3-D und Augmented Reality. 332
- W. REINHARDT & I. M. HOSSAIN:
Nutzung eines Body of Knowledge für das Curriculum Design. 342
- T. LUHMANN:
Softwareunterstützte Lehre in Photogrammetrie mit PhoX. 350

FERNERKUNDUNG - SCHWERPUNKT GEOLOGIE UND ARCHÄOLOGIE

- T. ULLMANN, E. LANGE & R. BAUMHAUER:
Transformationen der antiken Stätte Bubastis/Tell Basta (Ägypten)
im letzten Jahrhundert: Analyse moderner und
historischer kartographischer und fernerkundlicher Quellen. 363

POSTER

- R. BOERNER, L. HOEGNER & U. STILLA:
Concept of a Method for Change Detection
in Topobathymetric Point Clouds..... 371
- A. CONJO & R. REULKE:
Untersuchungen des Einflusses von Satellitenjitter auf die Bildqualität
von hochauflösenden optischen Sensoren. 383
- C. SPIESS & R. REULKE:
Vergleich der optischen Eigenschaften von hyperspektralen Kameras. 391
- W. DORNER, A. WEINFURTNER, L. DOSTÁLEK,
K. PAČLÍKOVÁ, K. HORNÍČKOVÁ & M. PREUSZ:
Crowd basierte Verortung historischer Fotografien und Konzepte zur ver-
besserten Aufarbeitung historischer Archive mit Hilfe geoinformatischer
Konzepte – Methodischer Ansatz des EU Projektes PhotoStruk. 403
- M. PÖCHTRAGER:
Segmentierung großer Punktwolken mittels Region Growing. 414
- Q. SCHREYER, E. GÜLCH & J. TAUTZ:
Konzept zur 3D-Erfassung, Analyse und Visualisierung
eines dynamischen Bienennests. 422
- A. DINKEL, L. HOEGNER & M. STEPHANI:
Integration und Bewertung photogrammetrischer Methoden
und GIS zur Dokumentation von archäologischen Projekten
am Beispiel von Gheriat el-Garbia (Libyen). 428
- M. CHIZHOVA, A. GURIANOV, D. KOROVIN, A. BRUNN & U. STILLA:
Automatically Extraction and Reconstruction of Cupola Geometries
of Orthodox Churches from Precision Point Clouds. 434
- T. VIETZE, D. SCHNEIDER & F. MAIWALD:
Untersuchung der Eignung photogrammetrischer Methoden zur
Erzeugung von 3D-Punktwolken aus historischen Bilddatenbeständen. 444

BEITRÄGE KARL-KRAUS-NACHWUCHSPREIS

S. DEGGIM, F. TSCHIRSCHWITZ & T. KERSTEN: Entwicklung eines virtuellen Museums für ein historisches Gebäude am Beispiel des Alt-Segeberger Bürgerhauses.	454
L. KNOPP, K. EDER, L. HOEGNER & U. STILLA: 3D Rekonstruktion von Biberdämmen zur Simulation des Wasserrückhaltes an Fließgewässern.	464
A. KÖHN: BRDF Analysis of Different Land Cover Types Using Aerial and Satellite Imagery.	472
N. METTIG, R. REULKE, J. FISCHER, R. PREUSKER & E. LORENZ: Sensitivitätsstudie zur Bestimmung der Meeresoberflächentemperatur aus Messungen des Satelliten TET-1.	482
K. MEYER: Aufbau einer Geodateninfrastruktur am Beispiel Sokodé, Togo.	492
I. PFEIL, C. REIMER & W. WAGNER: Determination of the Freeze/Thaw Surface State from ERS-2 Backscatter Data.	498
M. PILZ & K. RAUTZ: Entwicklung einer Tourenplanungssoftware (SNOWMAPS) auf Basis von abgeleiteten Karten aus Höhendaten.	508
R. ROFALLSKI & T. LUHMANN: Untersuchung und Modellierung des Rolling-Shutter-Effekts für photogrammetrische Einzel- und Mehrbildauswertungen.	514
M. STEINSIEK, P. POLEWSKI, W. YAO & P. KRZYTEK: Semantische Analyse von ALS- und MLS-Daten in urbanen Gebieten mittels Conditional Random Fields.	521
AUTORENVERZEICHNIS	532
PROGRAMMKOMITEE	536

UAS unterstützte Vegetationserfassung tidebeeinflusster Makrophytenbestände in der Unterelbe

GÖRRES GRENZDÖRFFER¹ & UWE SCHRÖDER²

Zusammenfassung: Die langfristige und objektive Überwachung und Vermessung schwer zugänglicher Gebiete, wie die Elbauen mit hohen Röhrichten und tidenbeeinflussten offenen Wattflächen sind ein ideales Einsatzszenario für UAS. Gleichwohl stellen gerade solche Areale in praktischen Projekten besondere Herausforderungen im Hinblick auf die Bildflugplanung und Realisierung dar. Der Beitrag geht auf die größten Herausforderungen bei der Bildflugplanung und Realisierung ein. Anschließend wird die Leistung verschiedener Softwareprodukte im Hinblick auf Genauigkeit und Vollständigkeit bei der Ableitung von photogrammetrischen Standardprodukten (Punktwolke und „True Orthophoto“) verglichen. Das hohe geometrische Genauigkeitspotential ermöglicht es, selbst geringe Veränderungen der Wattflächen von wenigen Zentimetern über einen Zeitraum von ca. zwei Jahren zu beschreiben. Die in diesem Zeitraum zu beobachteten saisonalen Veränderungen auf den Wattflächen betragen bis zu ± 20 cm, gleichen sich aber oftmals wieder aus.

1 Einleitung

Im Zuge der geplanten Fahrrinnenanpassung der Tideelbe und in Ergänzung der WRRL-Monitoringsaktivitäten sollen an mehreren Standorten der Unterelbe Makrophyten jährlich bis 12 Jahre nach Ausbau untersucht werden, WSD-NORD 2012. In diesem Zusammenhang sollen Luftbildaufnahmen von unbemannten Luftfahrtsystemen (UAS) für eine lagegenaue flächenhafte Erfassung der Bestände verwendet werden. Die UAS-Befliegungen sollen zudem im Zusammenhang mit der Vermessung der Topographie zur Ableitung von Erosions- und Sedimentationstendenzen flächenhafte Informationen dienen. Eine zentrale Fragestellung im Zusammenhang einer etwaigen Fahrrinnenvertiefung und sonstigen Veränderungen am Flussverlauf sowie dem Wasser- und Strömungsregime der Elbe ist die Veränderung der Wattflächen und der Küstenlinie, bzw. der Grenze zwischen Vegetation und Watt. Vor diesem Hintergrund hat die Bundesanstalt für Gewässerkunde (BfG) ein Forschungsprojekt initiiert, das diesen Fragen nachgehen soll. Darüber hinaus sollen die Möglichkeiten zur Analyse von 3D-Punktwolken in Kombination mit digitalen Farbluftbildern und daraus abgeleiteten Indizes zur Differenzierung von Biomasse, Bestandsdichte und Vegetationstypen untersucht werden. Aus der Lage der zu untersuchenden Gebiete (Abb. 1) wird ersichtlich, dass drei Standorte auf der schleswig-holsteinischen Seite der Elbe liegen und einer auf der niedersächsischen.

¹ Universität Rostock, Professur für Geodäsie und Geoinformatik, J.-v.-Liebig Weg 6, 18059 Rostock, E-Mail: goerres.grenzdoerffer@uni-rostock.de

² Bundesanstalt für Gewässerkunde, Referat U3 – Vegetationskunde, Landschaftspflege, Am Mainzer Tor 1, 56068 Koblenz, E-Mail: uwe.schroeder@bafg.de

werden kann, bzw. kurze sonnige Abschnitte ausgenutzt werden können, stellt der Wind in Küstennähe in Kombination mit der Tide ein nicht zu unterschätzendes Problem dar. Aus diesen Gründen mussten die Befliegungen teilweise um mehrere Wochen verschoben werden, bis sich ein entsprechendes Zeitfenster aufgetan hat, in dem sowohl Tide als auch Wind und Wetter eine Befliegung sicher zulassen.

Verhinderung von Überstrahlungen. Bei Befliegungen der nassen Wattflächen um die Mittagszeit, bzw. höher stehender Sonne, kann es zu spiegelnden Reflexionen und einem so genannten Hot Spot kommen. Dieser Effekt tritt dann ein, wenn der Sensor die Sonnenspur kreuzt und der Einfallswinkel gleich dem Ausfallswinkel ist. Eine wirksame Strategie ist demzufolge nur, in den Morgen- oder Abendstunden zu fliegen. Das wurde aber verworfen, da dadurch die möglichen Zeitfenster noch mehr eingeschränkt würden.

Messung von Passpunkten. Die Genauigkeit des Bildverbands steht und fällt mit der Entwicklung von transportablen Passpunkten, bzw. kodierten Messmarken, die auch bei Tide und hoher Vegetation verwendet werden können. Die ursprüngliche Idee permanenter Passpunkte ließ sich unter praktischen Gesichtspunkten nicht realisieren, da im Tidenbereich kontinuierlich große Kräfte wirken, und eine entsprechende konstruktive Lösung sehr aufwändig, bzw. u.U. auch genehmigungspflichtig wäre. Aufgrund der Tatsache, dass schwarze Messmarken auf einem weißen Hintergrund unter ungünstiger, d.h. heller Belichtung immer kleiner werden (GRENZDÖRFFER 2011), haben die als signalisierte Passpunkte verwendeten Aluminiumscheiben mit einem Durchmesser von 20 cm ein helles Zentrum und einen schwarzen Rand. Im Einzelnen wurden drei verschiedene Arten von Passpunkten verwendet. Im Röhricht wurden Passpunkte auf 1 bzw. 2 m hohe Vermessungsstäbe montiert. Auf Wegen, Sand oder sonstigen Flächen mit kurzer Vegetation wurden die Messmarken direkt auf den Boden gelegt. Auf den teilweise nassen und wellenbeeinflussten Wattflächen wiederum wurden umgedrehte Eimer verwendet, die nicht einsinken oder durch Sand überspült werden können.

Folgende Qualitätsfragen der photogrammetrischen Prozesskette und der abgeleiteten Produkte sind untersucht worden, auf die nachfolgend detailliert eingegangen wird:

- Wie genau sind die 3D-Punktwolken und wie zuverlässig kann die Vegetationshöhe und die Topographie zur Berechnung von Erosion und Sedimentation abgebildet werden?
- Welchen Einfluss hat die Auswertesoftware (Vergleich von Agisoft Photoscan, Pix4D und Trimble UAS-Master) auf die Ergebnisse?
- Lassen sich mit Hilfe von Vegetationsindices aus RGB-Aufnahmen Aussagen über die Vegetationstypen machen, bzw. die visuelle Interpretation unterstützen bzw. erleichtern?
- Ist die Qualität der erzeugten „True“ Orthophotos für eine Vegetationskartierung ausreichend, bzw. welche Stellschrauben bei der Erzeugung von True Orthophotos sind von Bedeutung?

Durch den multitemporalen Ansatz können verschiedene Prozesse erfasst und begleitet werden. Beispielsweise die Entwicklung der Vegetation, die sich zu verschiedenen Zeitpunkten unterschiedlich gut differenzieren lässt. Aufbauend auf der Frühjahrsaufnahme lassen sich darüber hinaus auch Vegetationshöhen ungefähr ableiten. Problematisch in diesem Zusammenhang geschlossene Schilfbestände, die auch im Winter einen dicht geschlossenen Röhrichtgürtel bilden können.

2.1 Photogrammetrische Prozessierung und Datenqualität

Die photogrammetrische Datenauswertung der UAS-Bildflüge erfolgte mit der Software Photoscan Professional der Firma Agisoft. Inhaltliche Auswertungen, Volumenberechnungen etc. erfolgten mit der Software ArcGIS 10.2.

Die Software Photoscan Professional liefert in einer sehr benutzerfreundlichen Art die automatische Bildorientierung für hochgradig überlappende Bildsequenzen einschließlich der Kamerakalibrierung und 3D-Punktwolken sowie abgeleitete Produkte z.B. ein digitales Oberflächenmodell, Orthophotos und Vegetationsindices. Verschiedene Vergleichsstudien zur Genauigkeit unterschiedlicher Softwareprodukte, sind in den letzten Jahren durchgeführt worden (MOUTINHO 2015). Dabei wurde festgestellt, dass kommerzielle Produkte durchwegs sehr gute, d.h. genaue und zuverlässige Ergebnisse geliefert haben. Allerdings beziehen sich diese Aussagen in erster Linie auf Genauigkeiten (Residuen) an Pass- oder Kontrollpunkten.

Grundsätzlich können die an den Passpunkten erzielten Genauigkeiten nicht auf alle 3D-Punkte übertragen werden. Das hat mehrere Ursachen. Z.B. ist auf Wasseroberflächen und anderen bewegten Oberflächen, z.B. durch Wind bewegte Bäume kein Matching möglich. Ähnliche Probleme ergeben sich auf sehr homogenen Flächen, bzw. Flächen mit gleichförmigen Strukturen, bei denen identische Punkte anhand fehlender Merkmale nicht identifiziert werden können bzw. fehlerhaft verknüpft werden. Das betrifft im vorliegenden Fall einige Wattbereiche, interessanterweise auch nahezu unabhängig von der verwendeten Software. Abb. 2 zeigt beispielhaft die ungleiche Verteilung der Verknüpfungspunkte (Keypoints) auf einem Bild. Abb. 3 zeigt die dichte 3D-Punktwolke der Wattfläche mit den fehlerhaften Bildzuordnungen (Rauschen), welche durch die strukturlosen Bildinhalte verursacht wurden.



Abb. 2: Ungleiche Verteilung der Keypoints in einem Bild (Bildflug Bielenberg, 9.6.2015)



Abb. 3: 3D-Punktwolke mit fehlerhaften Bildzuordnungen im Watt

Ebenfalls problematisch sind durch „Hot Spots“, Glitter oder Glare überstrahlte Flächen, bei denen ein Matching dieses Bilds zu seinen Nachbarn ebenfalls versagt. Das ist auf den Wattflächen allerdings nicht so problematisch, da durch die hohe Redundanz im Allgemeinen genügend Bilder zur Ermittlung einer akkuraten 3D Position zur Verfügung stehen. Zwischen belaubten Bäumen ist teilweise der Waldboden zu erkennen. Die ermittelten 3D-Koordinaten dieser Bereiche sind

aufgrund ungünstiger Schnittbedingungen gleichwohl mit besonderer Vorsicht zu genießen. Auf allen anderen Oberflächen lassen sich 3D-Koordinaten mit hoher Genauigkeit (< 5 cm) erfassen. Umfangreiche Tests in der Vergangenheit, z.B. HAALA (2014), haben das große Potential von Dense Matching Algorithmen anhand von digitalen flugzeuggestützten Luftbilddaten eingehend untersucht. Dabei hat sich gezeigt, dass gerade bei der Vegetation verschiedene Softwareprodukte zu unterschiedlichen Ergebnissen gekommen sind. Für höher aufgelöste UAS-Aufnahmen ist ähnliches zu beobachten (GRENZDÖRFFER 2014).

Aus diesem Grund wurde der 3. UAS-Flug im Testgebiet Hollerwettern vom 9.6.2015 mit drei verschiedenen Softwareprodukten ausgewertet, um vor allem die dabei abgeleiteten Punktwolken besser vergleichen zu können. In dem Softwarevergleich wurden Agisoft PhotoScan Professional (Version 1.2), Pix4Dmapper Pro (Version 2.3) und Trimble UASMaster einbezogen. Der Bildflug Hollerwettern am 9.6.2015 mit insgesamt 115 Aufnahmen bietet in dreierlei Hinsicht Herausforderungen für das Matching und Mosaicing:

1. Volle Vegetationsbedeckung und etwas Wind
2. Auf den Wattflächen Hot Spot Effekte am Bildrand
3. Teilweise sehr homogene Strukturen auf den Wattflächen

Die Bilder konnten mit allen drei Programmen problemlos ausgewertet werden. Die Residuen an den Passpunkten sind minimal und vergleichbar, wie Tab. 1 zeigt.

Tab. 1: Vergleich der Residuen an 12 Passpunkten für das Projekt Hollerwettern vom 9.6.2015

RMS Fehler [m]	X [m]	Y [m]	Z [m]	Total [m]
Agisoft Photoscan	0.0072	0.0053	0.0033	0.0095
Pix4D	0.0066	0.0059	0.0073	0.0115
Trimble UAS-Master	0.0035	0.0024	0.0132	0.0139

Anzumerken ist, dass die Residuen im Allgemeinen zu optimistische Daten liefern, und somit nur eine eingeschränkte Aussagekraft haben. Deshalb wurden die 4 Passpunkte im inneren des Bildverbands als Kontrollpunkte definiert. Im Ergebnis liegen die RMS Werte bei allen Produkten im ähnlichen Bereich (x) 0.017 m, (y) 0.002 m und (z) 0.024 m (Pix4D), (x) 0.011 m, (y) 0.001 m und (z) 0.005 m (Photoscan), sowie (x) 0.014 m, (y) 0.015 m und (z) 0.037 m (UAS Master). Dies entspricht den Erwartungen, da die tatsächlich erzielbaren Genauigkeiten üblicherweise bei dem 0,5 - 2 fachen der Bodenauflösung (GSD) liegen.

Die Punktwolken unterscheiden sich im Hinblick auf ihre Dichte und Vollständigkeit jedoch deutlich, wie der nachfolgende graphische Vergleich (Abb. 5 - 5) zeigt. Die größten Unterschiede sind in der hohen Vegetation (Bäume) und im Watt zu beobachten, d.h. genau an den Stellen an denen die Matching Algorithmen oftmals Schwierigkeiten haben. Dabei ist die Frage, welche Software nun besser ist, so einfach nicht zu beantworten. Das hängt auch damit zusammen, dass die (Standard) Parameter zur Erstellung der Punktwolken bei den verschiedenen Softwareprogrammen durchaus unterschiedlich sind. So wurden beispielsweise bei Pix4DMapper alle 2-fach verknüpften Punkte gewählt, um auch bei den Bäumen 3D-Punkte zu bekommen. Gleichzeitig sind deshalb

auf den Wattflächen Fehler aufgetreten. Die Oberfläche der Büsche und Bäume sowie der Priel sind bei Photoscan Professional am vollständigsten und detailliertesten beschrieben.

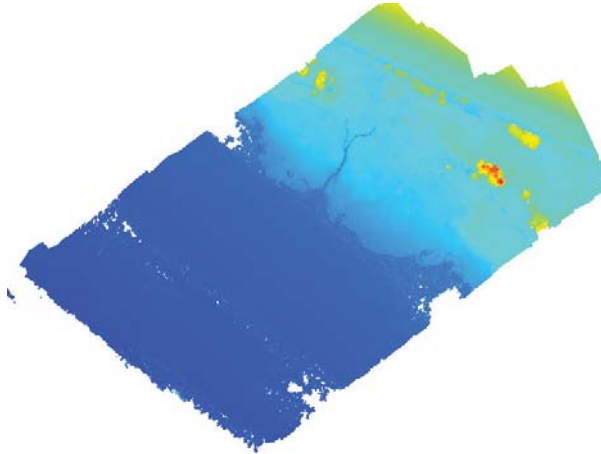


Abb. 4: 3D-Punktwolke mit der Software Trimble UAS-Master 19,6 Mio. Punkte

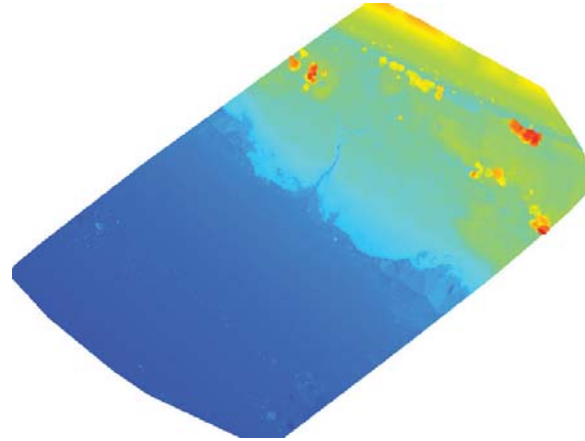


Abb. 5: 3D-Modell der Punktwolke mit der Software Pix4D Mapper (143,8 Mio. Punkte)

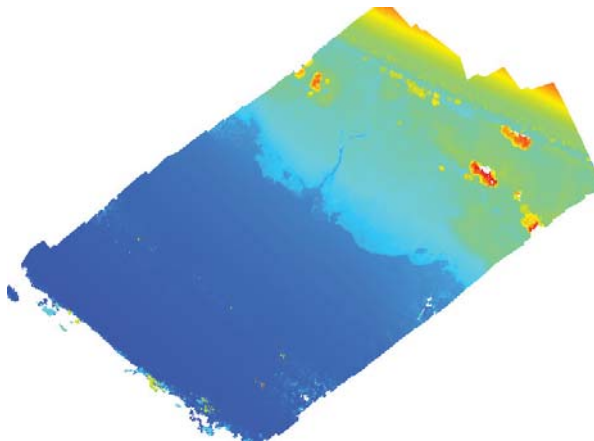


Abb. 6: 3D-Punktwolke mit der Software Photoscan Professional 47,3 Mio. Punkte

Der Ableitung eines digitalen Geländemodells (DGM) aus dem photogrammetrischen Oberflächenmodell durch eine Klassifikation der 3D-Punktwolke sind leider natürliche Grenzen gesetzt, da die Vegetation und hier insbesondere das Schilf auch im Winter einen dicht geschlossenen Röhrichtbestand bilden kann. Für andere Bereiche, z.B. Laubbäume können laubfreie Aufnahmen genutzt werden, um Informationen über die Geländetopographie zu erhalten.

Es zeigt sich zudem, dass die Ermittlung eines Zuwachses durch Vergleiche zweier digitaler Oberflächenmodelle auch nicht immer so einfach ist, wie gedacht. Vor allem dann nicht, wenn starker Wind das Schilf oder andere Vegetation herunterdrückt.

2.2 True Orthophotos - Vor und Nachteile

Im Rahmen der Datenprozessierung von UAS-Aufnahmen werden neben 3D-Punktwolken und digitalen Oberflächenmodellen üblicher Weise so genannte "True Orthophotos" erzeugt.

Per Definition (DIN 18740-3) ist ein „*True Orthophoto ... ein Orthophoto, in dem sichttote Räume ... weitgehend beseitigt sind.*“

Zur Erzeugung eines True Orthophotos, in dem alle Objekte lagerichtig abgebildet sind, müssen alle im Bild sichtbaren Objekte und deren Einzelteile im zugrundeliegenden Digitalen Oberflächenmodell repräsentiert sein. Während dies bei Gebäuden grundsätzlich machbar ist, von Problemen an den Gebäudekanten einmal abgesehen, gibt es bei Vegetation generell Probleme, vgl. Abb. 7. Die Probleme treten vor allem bei höheren Objekten, insbesondere Bäumen auf, deren Form nur unzureichend im DOM abgebildet ist. Gleichförmig hohe und dichte Vegetationsbestände, wie z.B. Schilf, werden im DOM als homogene Oberfläche erfasst und demzufolge im Orthophoto korrekt dargestellt.

Da bei entsprechendem Kontrast selbst Objekte kleiner als die Bildauflösung im Bild sichtbar sein können (z.B. Äste oder Freileitungen), scheitert deren lagerichtige Abbildung auch dann, wenn man ein Digitales Oberflächenmodell mit der Auflösung des zu erstellenden Digitalen Orthophotos zu Grunde legt. Das ist aus mehreren Gründen kaum machbar. Zum ersten können für ein Objekt, z.B. ein Baum, Haus, Brücke auch mehrere Höhen an einer Fassade oder im Geäst existieren, die zu einem digitalen Oberflächenmodell intelligent vermascht werden müssen. Zum zweiten ist das Matching jedes Bildpunktes in zwei und mehr Bildern, z.B. auf Grund von Bewegungen und Abschattungen, ungünstigen Schnittbedingungen etc. nicht möglich, siehe Abb. 9. Nicht zuletzt können im DOM auch Fehler und Artefakte vorhanden sein. Deshalb ist die Auflösung des DOM im Allgemeinen um den Faktor 4 - 9 geringer als das Luftbild, Abb. 10.

Neben einer möglichst detaillierten Beschreibung der tatsächlichen Oberfläche ist die Mosaikbildung der zweite problematische Schritt bei der Erzeugung eines True Orthophotos. Die in einem Bild auftretenden sichttoten Bereiche müssen aus anderen, benachbarten Bildern gewonnen werden. Grundvoraussetzung ist eine hohe Überdeckung, die bei UAS Befliegungen für Vermessungsaufgaben im Allgemeinen gegeben ist (Längs- und Querüberdeckung 80/60%) um jedes Objekt aus jeder Perspektive betrachten zu können. Dennoch können z.B. in einer Straßenschlucht oder zwischen Bäumen sichttote Räume verbleiben. Zusätzlich muss berücksichtigt werden, dass die einzelnen Bildpixel in der Regel mit ihren Nachbarpixeln korreliert sind, sei es durch das Bayer-Pattern auf der Sensor Ebene oder auch durch die Optik etc. Dies kann an der Grenze zwischen den zu verwendenden Bildbereichen zu Farbsäumen führen. Radiometrische (BRDF, Hotspots) und geometrische Restfehler können die Ergebnisse weiter beeinträchtigen, wie auf Abb. 12 deutlich zu sehen ist. Eine alternative Berechnungsmethode ist die Verwendung aller verfügbaren Farbwerte aus den zugrundeliegenden Bildern für die Erstellung des True Orthophotos. Dies bedeutet eine Mittelwertbildung und führt damit nahezu unweigerlich zu einem Verlust an Schärfe, vgl. Abb. 11. Gefordert ist also eine intelligente Mosaikbildung, die die vorhandenen Bilder und deren Blickrichtung für die Zusammenstellung des Mosaiks berücksichtigt, dabei aber weitgehend auf Mittelwertbildung verzichtet (WIEDEMANN & WICKI 2010).



Abb. 7: True Orthophoto auf der Grundlage eines UAS-DOM mit nahezu lagerichtiger Darstellung der Bäume und Artefakten



Abb. 8: Herkömmliches Orthophoto auf der Grundlage eines UAS-DGM, welches um Bäume und Gebäude bereinigt ist



Abb. 9: Aus UAS-Befliegung abgeleitete 3D-Punktwolke für True Orthophoto

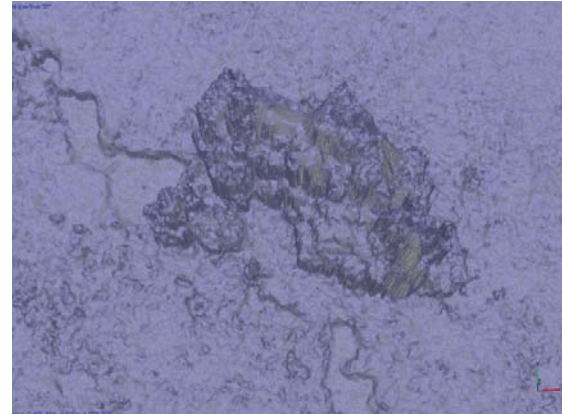


Abb. 10: Aus 3D-Punktwolke abgeleitetes Oberflächenmodell (Dreiecksvermaschung)



Abb. 11: Orthophoto mit Mittelwertbildung und Glättung im Übergangsbereich zwischen benachbarten Bildern



Abb. 12: True Orthophoto mit Artefakten im Gewässerbereich, die von geometrischen Restfehlern und radiometrischen Unterschieden benachbarter Bilder herrühren

3 Temporale Veränderungen im Watt

Eine zentrale Fragestellung im Zusammenhang einer etwaigen Fahrrinnenvertiefung und sonstigen Veränderungen am Flussverlauf und dem Wasser- und Strömungsregime der Elbe ist die Veränderung der Wattflächen und der Küstenlinie, bzw. der Grenze zwischen Vegetation und Watt. In Abhängigkeit des Standorts ist die Trennlinie zwischen der Vegetation und der Wattfläche in einer unterschiedlichen Höhe über dem Meeresspiegel und auch das Gefälle des Watts ist auf den Standorten sehr unterschiedlich, wie die nachfolgende Abb. 13 von vier Profilen aus den digitalen Oberflächenmodellen der UAS-Befliegungen zeigt. Die Profile wurden quer zur Küstenlinie gewonnen und beginnen jeweils kurz vor der Grenze zwischen der Vegetation und dem Watt. Das Gefälle im Watt ist recht konstant. In der Abb. 6 ist allerdings auch das Rauschen zu sehen, dass durch die fehlerhaften Punktzuordnungen auftreten kann. Die genauen Ursachen für die Unterschiede zwischen den Standorten hängen mit den lokal vorherrschenden Gelände- und Strömungsverhältnissen zusammen.

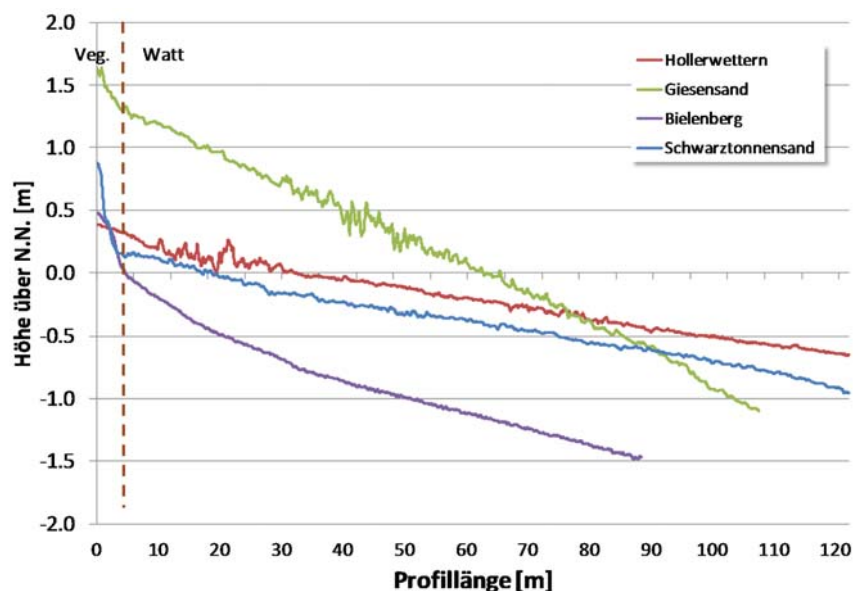


Abb. 13: Aus dem DOM abgegriffene Profillinien durch das Watt, quer zum Küstenverlauf, für die vier beflogenen Standorte, gestrichelte senkrechte Linie: geschlossene Vegetationsgrenze. Das zu beobachtende Rauschen ist auf Unzulänglichkeiten bei der Ableitung des DOM zurückzuführen.

Neben beliebigen Profilen sind die geometrisch hochaufgelösten UAS-Oberflächenmodelle besonders geeignet, um flächenhafte Veränderungen im Höhenprofil durch Vergleiche zwischen den Befliegungen zu erfassen. Das trifft im besonderen Maße auf die vegetationslosen Wattflächen zu somit das Oberflächenmodell dem Geländemodell entspricht. Die nachfolgenden drei Abbildungen zeigen die Differenzen zwischen den durchgeführten Befliegungen (Abb. 14, Abb. 15, Abb. 16). Aufgrund unterschiedlicher Tidenverhältnisse ist das zu beobachtende Watt unterschiedlich breit. Das vegetationslose Watt beginnt bei etwa 0.3 m ü. N.N. und reicht je nach Flugtermin zwischen -0,8 m für die Winteraufnahme und bis -1,5 m bei der Frühjahrs- und Sommeraufnahme. Die an obiger Stelle bereits erläuterten Matchingfehler führen teilweise zu einem Rauschen, was sich negativ auf das Ergebnis auswirkt.

Wie vielleicht zu erwarten war, kam es über die Wintermonate an den südlichen Teilen des Standorts Bielenberg zu einem leichten Abtrag von teilweise mehr als 10 cm, vgl. Abb. 14. Im Norden des Standorts ist hingegen keine signifikanten Veränderungen zu beobachten. Diese Entwicklung hat sich im Zeitraum zwischen der Frühjahrs- und der Sommeraufnahme wieder umgekehrt, wie die Abb. 15 anschaulich verdeutlicht. Im Saldo eines Jahres (Abb. 16) gleichen sich die Differenzen weitgehend aus. Allerdings ist an diesem Standort im Zeitraum Juni 2015 – August 2016 wieder eine Versteilung des Wattprofils zu beobachten. Ebenfalls festzuhalten bleibt, dass sich die Grenze zwischen dem Watt und der Vegetation kaum verändert hat.

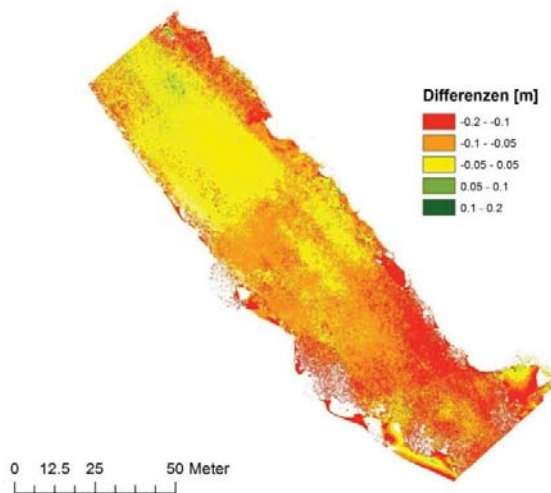


Abb. 14: Vertikale Veränderungen der Wattoberfläche am Standort Bielenberg von Herbst (30.10.2014) bis Frühjahr (09.03.2015)

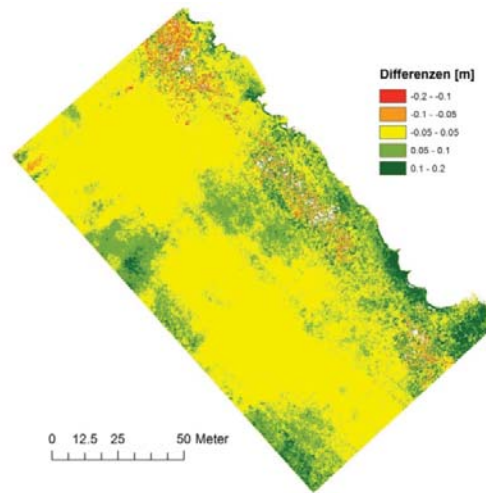


Abb. 15: Vertikale Veränderungen der Wattoberfläche am Standort Bielenberg von Frühjahr (09.03.2015) bis Sommer (09.06.2015)

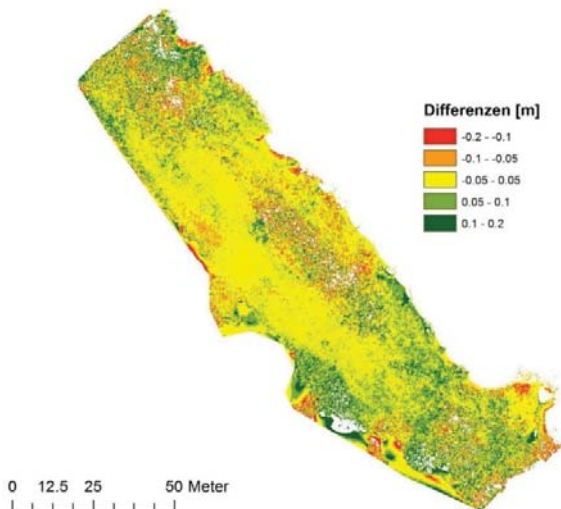


Abb. 16: Vertikale Veränderungen der Wattoberfläche am Standort Bielenberg von Herbst (30.10.2015) bis Sommer (09.06.2015)

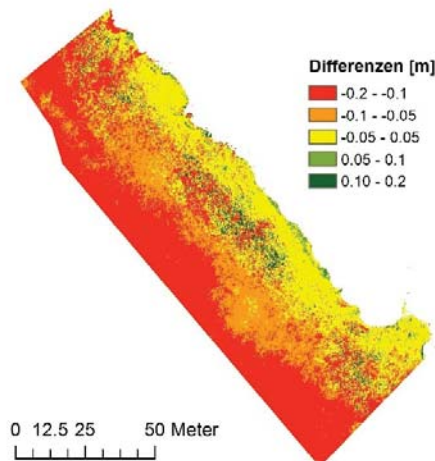


Abb. 17: Vertikale Veränderungen der Wattoberfläche am Standort Bielenberg von Sommer (09.06.2015) bis Sommer (04.08.2016)

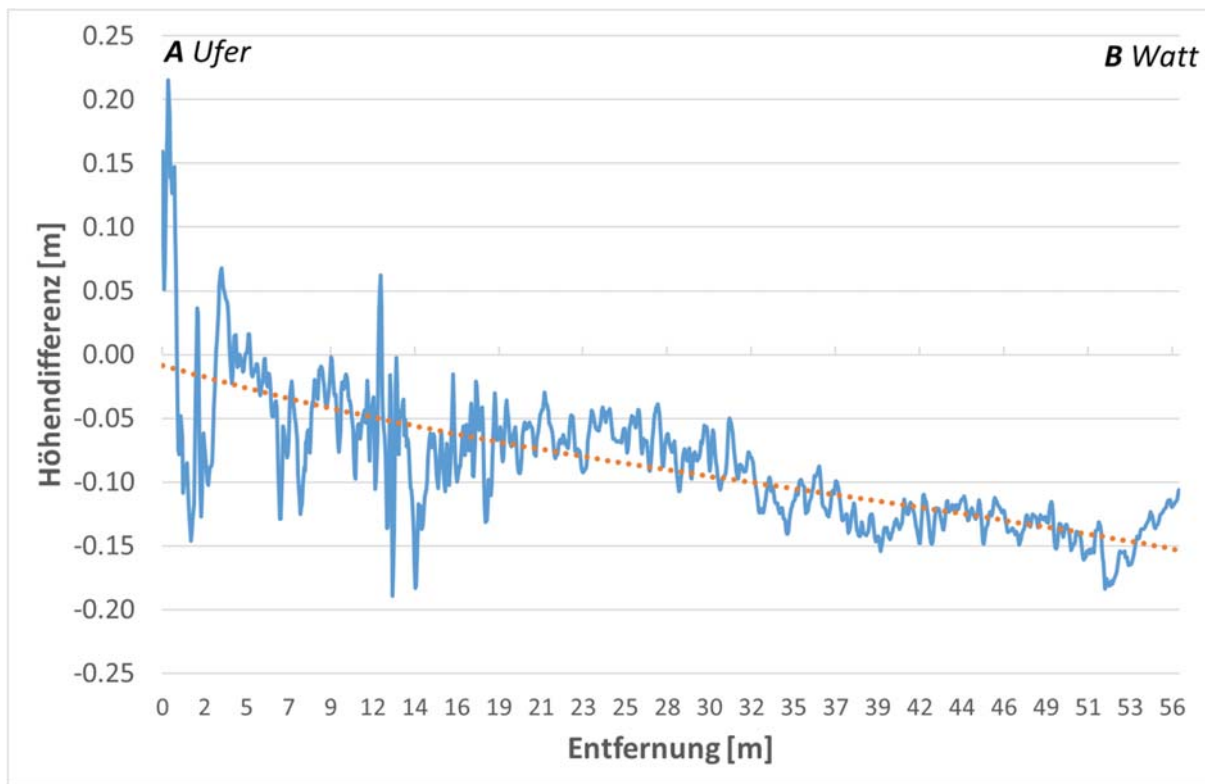


Abb. 18: Veränderung des Profils der Wattflächen am Standort Bielenberg vom 4.8.2016 im Vergleich zur Befliegung vom 9.6.2015 (etwa 50 m Breite) vom Ufer (links) bis zur maximal auswertbaren Tiefe von -1 m ü. N.N. (rechts)

Tab. 2: Vergleich der Höhenwerte [m ü. N.N.] bisheriger UAS Befliegungen entlang des Profils in Bielenberg

Befliegungsdatum	A - Ufer				B - Watt
30.10.2014	-0.03	-0.37	-0.65		
09.03.2015	0.14	-0.38	-0.66	-0.87	-1.00
09.06.2015	-0.08	-0.33	-0.66	-0.84	-0.98
04.08.2016	-0.05	-0.44	-0.73	-0.96	-1.09

Insgesamt zeigten sich also auf den untersuchten Standorten im Zeitraum der letzten zwei Jahre durchaus signifikante temporale Veränderungen an dem Profil der Wattflächen, jedoch nicht an dem Verlauf der Watt / Vegetationsgrenze bzw. der Küstenlinie.

4 Zusammenfassung

Der Ansatz eines kontinuierlichen (multitemporalen) landschaftsökologischen Monitorings mit Hilfe von UAS ist ein längerfristiger Ansatz, mit dessen Hilfe verschiedene Prozesse erfasst und begleitet werden können. Beispielsweise die Entwicklung und Verbreitung verschiedener Vegetationstypen und -gesellschaften, die sich zu verschiedenen Zeitpunkten unterschiedlich gut differenzieren lassen. Die Interpretation und Erfassung der Veränderungen sollte manuell erfolgen, denn für eine weitgehend automatische Auswertung / Klassifikation des Bildmaterials bedarf es aus heutiger Perspektive noch weitere Untersuchungen, da die sehr hohe Auflösung und die variable Radiometrie, die bei wechselnden Belichtungsverhältnissen während einer Befliegung auftreten kann, zu viele Sonderfälle erzeugt.

Fehler im DOM bzw. ein DOM mit einer Auflösung unterhalb der Bildauflösung führen zu Artefakten im "True" Orthophoto. Das betrifft insbesondere Vegetations- und Wasserbereiche, in denen die zugrundeliegenden Matchingverfahren aus verschiedenen Gründen Probleme haben. Diese sind für eine visuelle Interpretation in der Regel unproblematisch, bei automatisierten Verfahren jedoch zu berücksichtigen.

Insgesamt zeigten sich auf den untersuchten Standorten signifikante Veränderungen an den Wattflächen bzw. der Küstenlinie. Veränderungen über den Winter haben sich über das Frühjahr bis zum Sommer meist wieder ausgeglichen. Um die Relevanz und die Ursachen der Veränderungen beobachten zu können, ist ein zeitlich engeres Befliegungsschema notwendig.

5 Literaturverzeichnis

- GRENZDÖRFFER, G., 2011: Geodatenerfassung mit Unmanned Airborne Systems (UAS). Terrestrisches Laserscanning - TLS 2011 mit TLS Challenge, DVW-Schriftenreihe Band **66**, Wißner-Verlag, 203-214.
- GRENZDÖRFFER, G., 2014: Crop Height Determination with UAS Point Clouds. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **40**(1), 135-140.
- HAALA, N., 2014: Dense Image Matching Final Report. EuroSDR Official Publication **64**, 115-146.
- MOUTINHO, O., 2015: Evaluation of Photogrammetric Solutions for RPAS: Commercial vs. Open Source. Master thesis, Universidade do Porto, <https://repositorio-aberto.up.pt/handle/10216/82382>, last access January 22, 2017.
- WIEDEMANN, A. & WICKI, P., 2010: Mythos True Orthophotos - Vom Sinn und Unsinn eines Produkts. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 19, Kohlhofer, G. & Franzen, M. (Hrsg.), 271-279.
- WSD-NORD, (2012): Planfeststellungsbeschluss für die Fahrrinnenanpassung der Unter- und Außenelbe für 14,5 m tiefgehende Containerschiffe. Wasser- und Schifffahrtsdirektion Nord, Kiel, AZ.: P-143.3/46.

Verbesserung der Geometrie von Satellitenbildern durch Höhenmodelle

KARSTEN JACOBSEN¹

Zusammenfassung: Nicht alle hochauflösenden optischen Satellitenbildkameras haben eine ausreichende und stabile Kalibrierung. In Flugrichtung hängt die Bildgeometrie von der relativen Genauigkeit und Frequenz der Winkelbestimmung ab. Nach einer schnellen Orientierungsänderung kann es zu Schwingungen kommen, die nicht immer ausreichend erfasst werden und in den rationalen Polynomkoeffizienten, die die Sensororientierung ausdrücken, berücksichtigt sind. Die original bestimmten Sensorrotationen werden meistens nicht mit den Bilddaten ausgeliefert und können dementsprechend nicht für eine geometrische Rekonstruktion der Bildorientierung genutzt werden. Auf Satellitenstereomodellen basierende digitale Höhenmodelle können durch nicht ausreichend bekannte Bildorientierungen mit üblicher Anzahl von Passpunkten zur Biaskorrektur deformiert und nicht ausreichend horizontiert sein. Systematische, teilweise szintillierende, Differenzen der γ -Parallaxen der Vorwärtsschnitte zur Objektkoordinatenbestimmung der Höhenmodelle deuten auf systematische Höhenfehler hin.

Auf verschiedenen hochauflösenden optischen Satelliten basierende Höhenmodelle wurden mittels nahezu weltweiter Höhenmodelle auf systematische Höhenfehler untersucht und korrigiert, bevor sie mit ausreichend genauen Referenzhöhenmodellen verglichen wurden. Als Höhenmodelle zur Korrektur wurden die kostenlos verfügbaren digitalen Oberflächenmodelle (DOM) von SRTM, ASTER GDEM2 und das ALOS World 3D (AW3D30) mit einer Bogensekunde Rasterweite verwendet. AW3D30 hat zwar die größte Höhengenaugigkeit dieser nahezu weltweiten Höhenmodelle, zur Bestimmung der systematischen Höhenfehler reichen jedoch alle aus. Die ermittelten systematischen Höhenfehler sind durch Tiefpassfilterung zu verbessern um einen Einfluss lokal veränderter Vegetationshöhe zu reduzieren.

1 Einleitung

Die innere Orientierung optischer Satellitenkameras wird vor dem Satellitenstart kalibriert. Durch die Startbeschleunigung, Temperatureinflüsse und Austrocknungseffekte kann sie sich später verändern und erfordert eine periodische Überprüfung im Weltraum. Die innere Orientierung ist durch die Sensorkonfiguration komplex, so werden z.B. die panchromatischen Zeilen von WorldView-2 und WorldView-3 aus 60 CCD-Elementen zusammengesetzt (Abb. 1).

Da die CCD-Elemente nicht lückenlos aneinanderstoßen können, sind sie in 2 Gruppen aufgeteilt, die in Aufnahme­richtung gegeneinander versetzt sind. Um eine einheitliche Zeile zusammensetzen zu können, müssen die Teilbilder leicht unterschiedlicher Aufnahmezeitpunkte miteinander kombiniert werden. Die innere Orientierung ist somit auch von der äußeren Orientierung abhängig, die sich durch ein leicht verändertes Aufnahmezentrum und veränderte Aufnahme­richtungen bemerkbar machen kann.

¹ Leibniz Universität Hannover, Nienburger Str. 1, 30167 Hannover,
E-Mail: jacobsen@ipi.uni-hannover.de

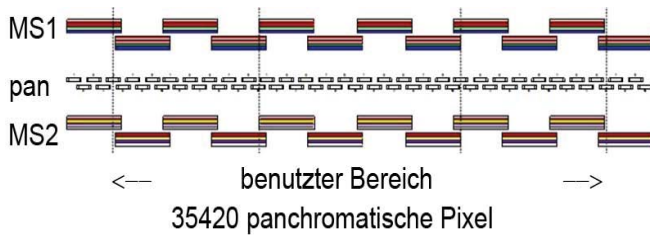


Abb. 1: CCD-Konfiguration von WorldView-2 und WorldView-3
pan = panchromatisch
MS1, MS2 = multispektral

Die erste Information über eine mögliche Beeinflussung eines Höhenmodells durch Probleme der inneren und der äußeren Orientierung liefern die y -Parallaxen der Vorwärtsschnitte für die Punkte der Höhenmodelle nach der automatischen Bildzuordnung. Abbildung 2 zeigt die systematischen y -Parallaxen für die Berechnung eines WorldView-2 Höhenmodells nach Mittelbildung in X - und Y -Richtung. Die gruppenweise Mittelbildung der y -Parallaxen in der X -Richtung (quer zum Orbit) zeigt Widersprüche mit der Frequenz der Länge der einzelnen CCD-Elemente was sich durch eine hochfrequente Vibration des Satelliten erklären lässt, während die Mittelwerte in Y -Richtung auf Fehler der absoluten Orientierung hindeuten. Das östliche Nachbarmodell des WorldView-2-Stereopaars, das während der gleichen 1,5 Minuten wie das Modell 2 (Abb. 2) aufgenommen wurde, zeigt als Funktion von X die gleiche Frequenz und Amplitude. In der Y -Richtung sind ähnliche Werte wie im Modell 2 vorhanden.

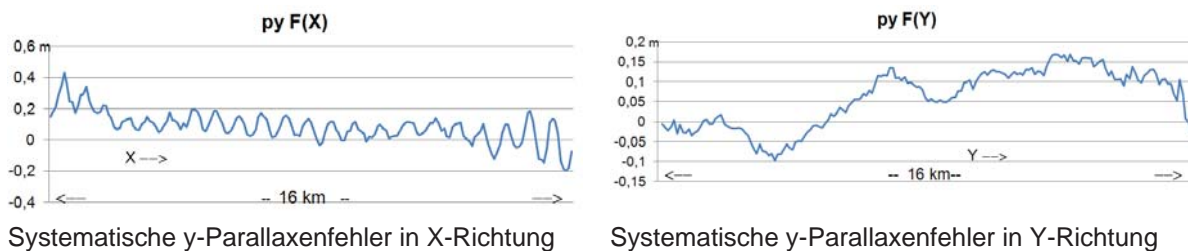


Abb. 2: Systematische Parallaxenfehler WorldView-2, Karaburun, Modell 2

Systematische Sensororientierungsfehler lassen auch systematische Höhenfehler erwarten. Vom chinesischen optischen Stereosatelliten Ziyuan-3 wird über eine begrenzte Genauigkeit der Kalibrierung und der direkten Sensororientierung berichtet (PAN & ZOU 2016; JIANG et al. 2015a; JIANG et al. 2015b; RADHADEVI et al. 2016; JACOBSEN 2016a). Satellitenvibrationen nach schneller Bewegung von einem Aufnahmegebiet zum nächsten können die Satellitenbildgeometrie beeinflussen. Wenn hochfrequente Änderungen zwar durch die Kreisel registriert wurden, aber nicht direkt in der Bildgeometrie berücksichtigt sind, können sie nicht durch die rationalen Polynomkoeffizienten (RPC) dritter Ordnung ausreichend berücksichtigt werden und verursachen systematische Höhenfehler. Diese lassen sich allerdings durch nahezu weltweite Referenzhöhenmodelle, wie SRTM, ASTER GDEM2 und AW3D30, nachweisen und a posteriori berücksichtigen.

Eine Reihe von Höhenmodellen, erstellt aus optischen Satellitenbildpaaren, wurde durch Vergleich mit Referenzhöhenmodellen auf systematische Fehler untersucht und korrigiert. Hier stellt sich das Problem, dass SRTM, ASTER GDEM2 und AW3D30 digitale Oberflächenmodelle sind, die die Höhen der sichtbaren Oberflächen repräsentieren. Das SRTM DOM basiert auf interferometrischen synthetischen Apertur-Radar (InSAR) im C-Band. Das C-Band Radar dringt nur geringfügig in die Vegetation ein. Die Vegetationshöhe kann sich seit der Aufnahme im Februar 2000 verändert haben. ASTER GDEM2 nutzt alle brauchbaren optischen

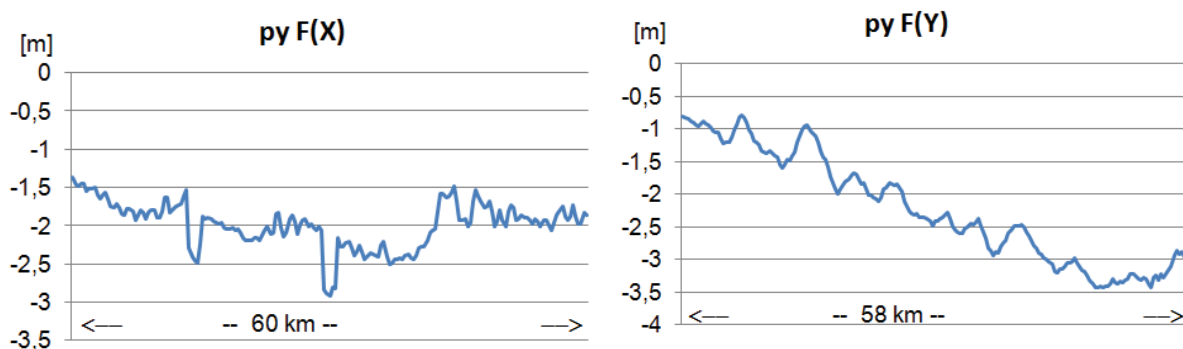
Stereoaufnahmen von Terra ASTER, die seit Dezember 1999 aufgenommen wurden und präsentiert damit eine mittlere Vegetationshöhe des gesamten Aufnahmezeitraums. Für das seit Mai 2016 verfügbare AW3D30 (JACOBSSEN 2016b) wurden alle brauchbaren ALOS PRISM Tri-Stereoaufnahmen verwendet, aufgenommen vom Januar 2006 bis zum May 2011. Die Oberflächendefinition ist stark von der Bodenauflösung abhängig, die bei Terra ASTER 15m und bei ALOS PRISM 2,5m beträgt. Alle drei Höhenmodelle haben einen Punktabstand von 1 arcsec, entsprechend etwa 31m am Äquator. ASTER GDEM2 ist nicht so genau wie das SRTM DOM und AW3D30 (Jacobsen 2016b) weswegen es im Folgenden nicht weiter erwähnt wird.

Bei allen durchgeführten Untersuchungen wurden dreidimensionale Lageabweichungen der Höhenmodelle durch Ausgleichung a priori berücksichtigt.

2 Systematische Fehler digitaler Höhenmodelle aus Satellitenbildern

2.1 Ziyuan-3

Mittels der beiden äußeren Aufnahmen des chinesischen Tri-Stereo-Satelliten Ziyuan-3, die 3,4m Bodenauflösung haben, wurde mit den ISPRS-Testdaten Sainte-Maxime ein Höhenmodell durch Kleinste-Quadrate-Bildzuordnung erstellt. Die Bildorientierung erfolgte, wie üblich, durch bias-korrigierte rationale Polynomkoeffizienten (RPC), andere Informationen über die direkte Sensororientierung lagen nicht vor.



Systematische y-Parallaxenfehler in X-Richtung

Systematische y-Parallaxenfehler in Y-Richtung

Abb. 3: Systematische y-Parallaxenfehler des Ziyuan-3 Stereomodells Sainte-Maxime

Die systematischen y-Parallaxenfehler (Abb. 3) unterscheiden sich von denen der von WorldView-2 (Abb. 2). Es treten zwar auch in X-Richtung Deformationen auf, jedoch zeigen sich periodischen Einflüsse hauptsächlich in Y-Richtung (in Orbit-Richtung). Der Größenunterschied der Parallaxenabweichungen ist durch die Bodenauflösung von 0,5m bei WorldView-2 und 3,4m bei Ziyuan-3 zu erklären. Der leichte Trend der y-Parallaxen in X-Richtung bei WorldView-2 und der starke Trend in Y-Richtung bei Ziyuan-3, sind durch Orientierungsunsicherheiten hervorgerufen. Das Ziyuan-3 DOM wurde mit einem Referenz-DOM des französischen IGN verglichen, das eine Standardabweichung von etwa 20cm hat. Es basiert auf digitalen Luftbildaufnahmen die fast zeitgleich mit den Ziyuan-3 Bildern aufgenommen wurden und 10cm Bodenauflösung haben. Außerdem sind in Abbildung 4 die systematischen Höhenfehler gegenüber dem SRTM-DOM und gegenüber AW3D30 dargestellt. Die systematischen Höhenfehler in X- und Y-Richtung wurden

durch ein gleitendes Polynom dritter Ordnung geglättet, das die 9 nächst benachbarten Werte berücksichtigt, um den zufälligen Einfluss der Vegetationsänderungen zu reduzieren.

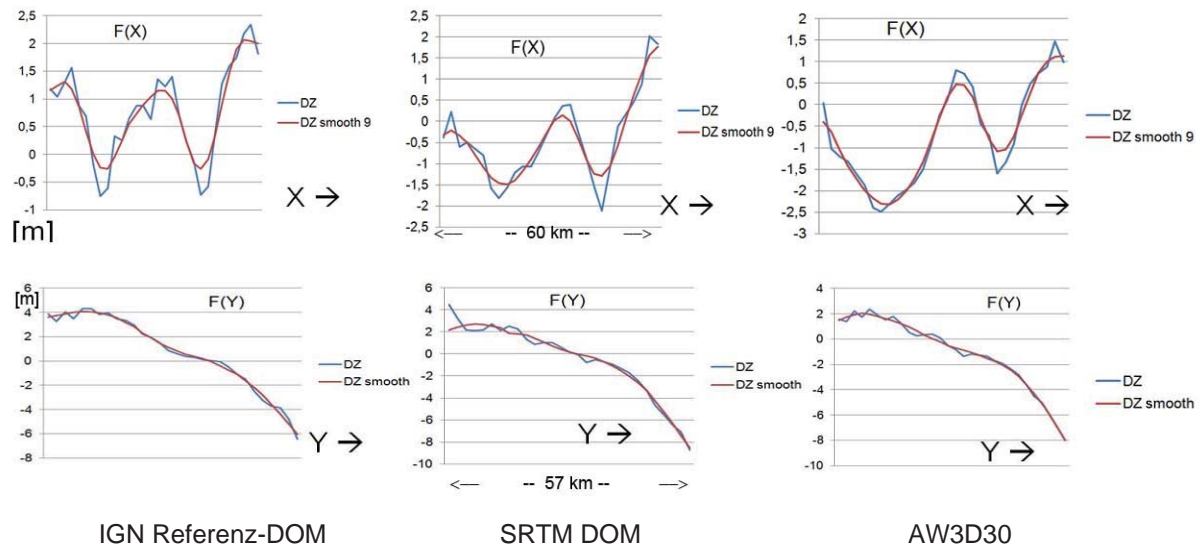


Abb. 4: Systematische Höhenfehler des Ziyuan-3-DOM gegenüber den Referenz-DOM

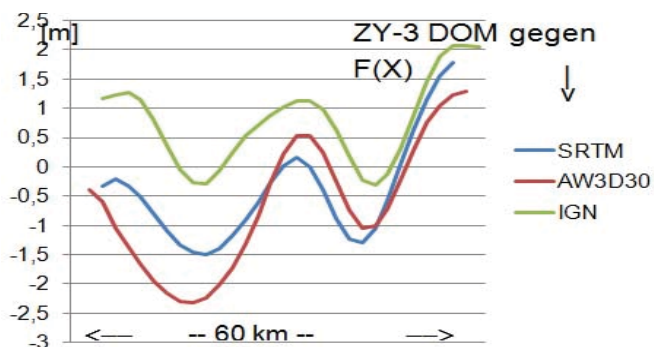


Abb. 5: Geglättete systematische Höhenfehler des Ziyuan-3 Oberflächenmodells gegenüber 3 Referenz-DOM

Während in der Bahnrichtung (Y) eine Neigung des Höhenmodells dominiert, sind in Zeilenrichtung (X) klare Undulationen erkennbar, die, wenn leichte Neigungen und ein Höhenversatz berücksichtigt werden, sehr ähnlich sind (Abb. 5).

Tab. 1: Genauigkeitsmaße für die Differenz gegenüber dem Ziyuan-3 Höhenmodell

Höhenmodell	SZ original	NMAD original	SZ nach Verbesserung	NMAD Nach Verbesserung
IGN	5,33 m	5,06 m	4,34 m	3,35 m
SRTM	5,78 m	5,32 m	4,98 m	3,99 m
AW3D30	4,90 m	4,54 m	3,75 m	3,22 m

Tabelle 1 zeigt die Standardabweichung der Höhendifferenzen (SZ) die „normalized median absolute difference“ (NMAD) der ursprünglichen und der um die systematischen Höhenfehler

verbesserten Höhenmodelle gegenüber den Referenzhöhenmodellen. SZ reduziert sich im Mittel um 19% und NMAD um 29% durch Korrektur um die systematischen Höhenfehler. Die Normalverteilung basierend auf NMAD beschreibt in allen Fällen die Häufigkeitsverteilung der Höhendifferenzen besser als die Normalverteilung basierend auf SZ (Abb. 7).

Tabelle 1 drückt die innere Genauigkeit aus, deswegen wurde das Ziyuan-3 Höhenmodell mittels der systematischen Höhenfehler von SRTM und AW3D30 verbessert und das Ergebnis mit dem IGN Referenzhöhenmodell verglichen (Tabelle 2). Die erreichte Höhengenaugkeit im Gelände mit Neigungen kleiner als 10% liegt unter der Bodenauflösung von 3,4m und ist damit sehr zufriedenstellend.

Tab. 2: Genauigkeitsmaße für das Ziyuan-3 DOM verbessert um systematische Höhenfehler gegenüber SRTM und AW3D30, verglichen mit dem IGN Referenz-DOM in offenen Gebieten und zusätzlich für die flacheren Bereiche mit Geländeneigungen < 10%

Referenz zur Verbesserung	SZ	NMAD	SZ für $\alpha < 10\%$	NMAD für $\alpha < 10\%$
SRTM	4,53 m	2,86 m	3,24 m	2,76 m
AW3D30	3,62 m	2,62 m	2,86 m	2,62 m

Durch Verbesserung mit dem AW3D30 wurden bessere Genauigkeiten als mit dem SRTM DOM erreicht. Das entspricht der höheren Genauigkeit von AW3D30 gegenüber dem SRTM DOM (Jacobsen 2016b). Der Unterschied im NMAD für flache Gebiete beträgt allerdings nur 5%.

2.2 SPOT-5

Auch das SPOT-5 Höhenmodell Zonguldak zeigt in X-Richtung ähnliche systematische Höhenfehler wie das Ziyuan-3 DOM (Abb. 6), wogegen in der Bahnrichtung nur eine geringfügige Neigung vorhanden ist.

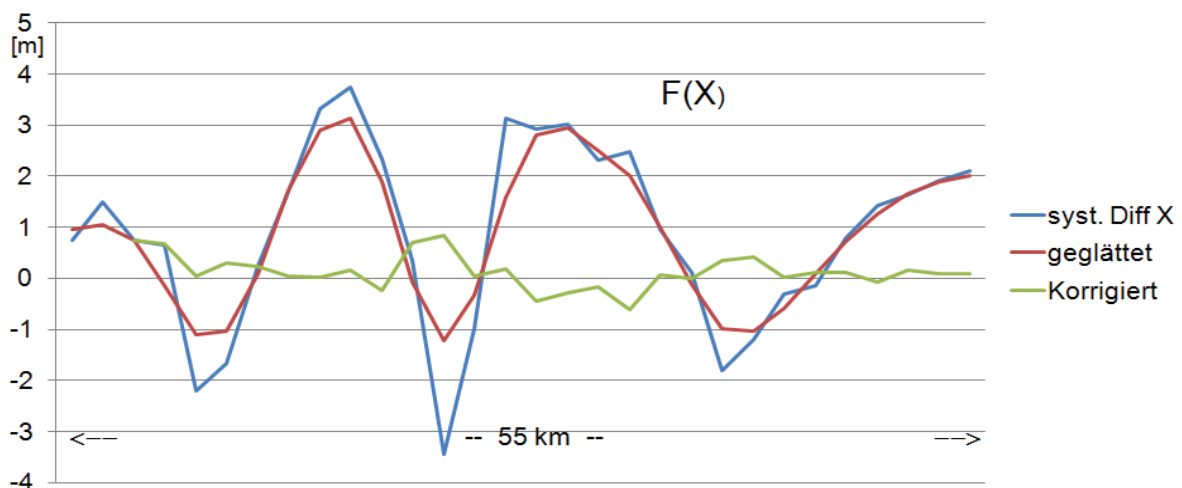


Abb. 6: Systematische Höhenfehler des SPOT-5 Höhenmodells Zonguldak gegen AW3D30 in X-Richtung

Tabelle 3: Genauigkeitsmaße des SPOT-5 DOM vor und nach Verbesserung um systematische Fehler

	SZ	NMAD	SZ für $\alpha < 10\%$	NMAD für $\alpha < 10\%$

ursprünglich	7,76 m	5,86 m	6,79 m	5,02 m
Nach Verbesserung	7,32 m	5,10 m	6,23 m	4,13 m

Die nach Korrektur der systematischen Fehler verbleibenden Restsystematiken (Abb. 6 „korrigiert“) sind klein und in Relation zu 5m Bodenauflösung vernachlässigbar. Durch die Verbesserung um die systematischen Höhenfehler reduzieren sich die Standardabweichungen um 6% bzw. 8% und NMAD um 13% bzw. 18%. Die geringere Verbesserung der Standardabweichung und der deutliche Unterschied zwischen SZ und NMAD sind durch den höheren Anteil großer Differenzen als entsprechend der Normalverteilung zu erklären (Abb. 7), die sich besonders auf die Standardabweichung auswirken. Abbildung 7 verdeutlicht auch die bessere Approximation der Häufigkeitsverteilung durch die Normalverteilung basierend auf NMAD als durch SZ. Im Mittel weichen hier die Normalverteilung basierend auf NMAD um 4,5% und SZ um 21,9% von der Häufigkeitsverteilung ab.

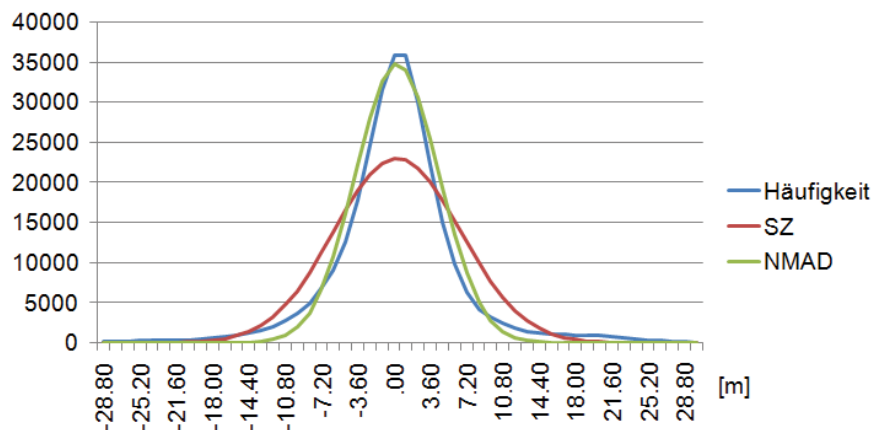


Abb. 7: Häufigkeitsverteilung der SPOT-5 Höhendifferenzen gegenüber dem Referenz-DOM mit überlagelter Normalverteilung basierend auf SZ und NMAD

2.3 WorldView-2

Ein Teil eines WorldView-2 (mit 0.5m Bodenpixel vorhanden) Oberflächenmodells wurde mit einem Laserscanner-Oberflächenmodell verglichen. Da sich einige Steinbrüche in dem Gebiet verändert hatten und sich in den bewaldeten Teilen die Oberflächendefinition zwischen Laserscanner und optischer Bildzuordnung deutlich unterscheidet, wurden die entsprechenden Flächen aus der Untersuchung herausgenommen. Auch hier zeigen sich klare Auswirkung von nicht berücksichtigten Satellitenvibrationen (Abb. 8 und Abb. 2). In Bahnrichtung (Y) ist nur eine gleichmäßige Neigung des DOM vorhanden.

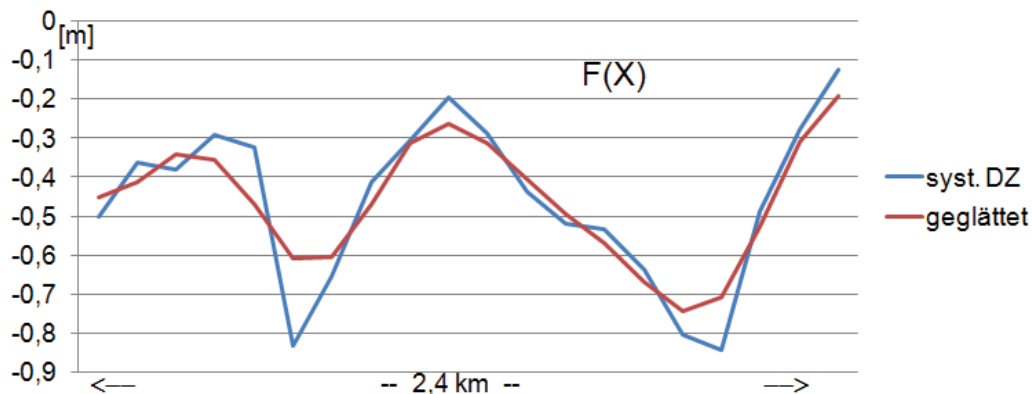


Abb. 8: Systematische Höhenfehler eines Teiles eines WorldView-2 DOM, Karaburun

Tab. 4: Genauigkeitsmaße des WorldView-2 DOM vor und nach Verbesserung um systematische Fehler

	SZ	NMAD	SZ für $\alpha < 10\%$	NMAD für $\alpha < 10\%$
ursprünglich	1,24 m	0,87 m	0,96 m	0,65 m
Nach Verbesserung	1,16 m	0,74 m	0,88 m	0,55 m

2.4 Cartosat-1

Das mittels der Stereoaufnahmen von Cartosat-1 (2,5m Bodenpixel) erstellte Oberflächenmodell in Jordanien zeigt eine gleichmäßige und keine periodische Deformation in X-Richtung (Abb. 9). Die Standardabweichungen werden durch ihre Korrektur um etwa 7% und NMAD um 15% verringert.

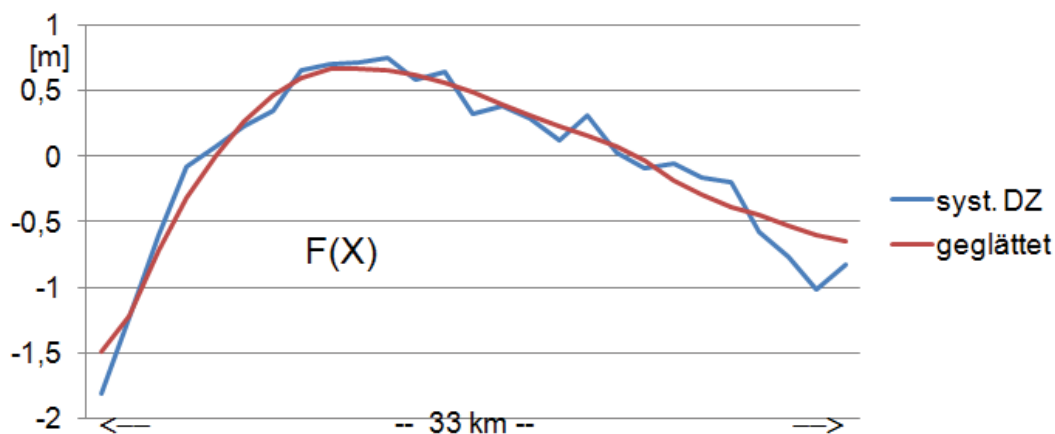


Abb. 9: Systematische Höhenfehler des Cartosat-1 DOM als Funktion von X, Jordanien

Tab. 5: Genauigkeitsmaße eines Cartosat-1 DOM vor und nach Verbesserung um systematische Fehler

	SZ	NMAD	SZ für $\alpha < 10\%$	NMAD für $\alpha < 10\%$
ursprünglich	3,11 m	2,49 m	2,55 m	2,03 m
Nach Verbesserung	2,85 m	1,99 m	2,28 m	1,52 m

3 Zusammenfassung

Die gezeigten Beispiele demonstrieren, dass nicht nur Höhenmodelle die mittels des Ziyuan-3 erstellt wurden, über dessen eingeschränkter Orientierungsgenauigkeit von chinesischer Seite berichtet wird, sondern auch Höhenmodelle anderer optischer Satelliten deformiert sein können. Grundsätzlich müssen die Deformationen in Relation zur Objektpixelgröße und damit zur erreichbaren Höhengenaugigkeit betrachtet werden. Um die geometrische Genauigkeit bestimmen zu können dürfen auch nur DOM mit Referenz-DOM oder Geländemodelle mit Referenzgeländemodellen verglichen werden und nicht, wie es in der Literatur häufig geschieht, DOM mit Referenzgeländemodellen verglichen werden. Die Vegetationshöhe bestimmt die Definition der Höhenmodelle und nicht deren Genauigkeit.

Bei den oben dargestellten Höhenmodellen treten vor allem in Zeilenrichtung höher frequente Deformationen auf, die nicht durch die innere Orientierung, sondern durch den Versatz der CCD-Einheiten (Abb. 1) in Verbindung mit Satellitenvibrationen erklärt werden können. In Bahnrichtung zeigen sich dagegen hauptsächlich Modellneigungen und großräumige Deformationen. Alle sehr hoch auflösenden optischen Satelliten haben „Transfer Delay and Integration“ (TDI) Sensoren, die die Ladung während der Aufnahme entsprechend der Vorwärtsbewegung in die Nachbarzeilen verschieben und so die Objektreflexion über eine wählbare Zeit integrieren können. Satellitenvibrationen reduzieren deswegen die Bildqualität. Wogegen eine leicht reduzierte Bildqualität bei flächenbasierten Bildzuordnungen praktisch keine Reduktion der Zuordnungsqualität bewirkt.

Durch die Korrektur der Höhenmodelldeformationen treten die in Tabelle 6 aufgezeichneten Verbesserungen ein. Der größte Anteil der Verbesserungen tritt durch die Beseitigung der Modellneigungen und nicht durch hochfrequente Anteile ein. Die Modellneigungen können allerdings auch durch eine Stereorientierung der Höhenmodelle erreicht werden (D'ANGELO 2013), wogegen sie hier entsprechend der üblichen Methode separat für die einzelnen Bilder mittels bias-korrigierter RPC durchgeführt wurde.

Tab. 6: Genauigkeitssteigerung der Höhenmodelle durch Korrektur um systematische Fehler

	Ziyuan-3	SPOT-5	WorldView-2	Cartosat-1
SZ	19%	7%	7%	9%
NMAD	29%	15%	15%	22%

NMAD wird stärker verbessert als SZ, da SZ stärker durch größere Differenzen beeinträchtigt wird und diese häufiger, als der Normalverteilung entsprechend, auftreten (Abb. 7). Auch wenn der Hauptanteil der systematischen Fehler durch die Modellneigungen hervorgerufen wird, dürfen die nicht-linearen Fehler nicht vernachlässigt werden da sie bei Ziyuan-3 30% - 60%, bei SPOT-5 60%, bei WorldView-2 30% und bei Cartosat-1 bis zu 50% erreichen. Eine Korrektur mittels SRTM-DOM oder AW3D30 stellt keinen großen Aufwand dar. Ähnliche Probleme, wenn auch etwas kleiner, wurden auch bei anderen optischen Satelliten festgestellt (Jacobsen 2016a).

4 Literaturverzeichnis

D'ANGELO, P., 2013. Automatic orientation of large multi-temporal satellite image blocks. Proceedings of International Symposium on Satellite Mapping Technology and Application 2013, 1-6

- JACOBSEN, K., 2016a: Analysis and correction of systematic height model errors. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B1, 333-339.
- JACOBSEN, K., 2016b: Ersatz des SRTM Oberflächenmodells durch AW3D30? *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 25*, Kersten, T. (Hrsg.), Tagungsband der 36. Wissenschaftlich-Technischen Jahrestagung der DGPF, Dreiländertagung vom 7.-9. Juni 2016 in Bern (auf CD), 80-88.
- JIANG, Y.H., ZHANG, G., TANG, X.M., LI, D., HUANG, W.C. & PAN, H.B., 2015a: Geometric Calibration and Accuracy Assessment of ZiYuan-3 Multispectral Images. *IEEE Transactions on Geoscience and Remote Sensing*, **52**(7), 4161-4172.
- JIANG, Y.H., ZHANG, G., TANG, X., LI, D.R., WANG, T., HUANG, W.C. & LI, L.T. 2015b: Improvement and Assessment of the Geometric Accuracy of Chinese High-Resolution Optical Satellites. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **8**(10), 4841-4852.
- PAN, H. & ZOU, Z., 2016: Penalized Spline: A General Robust Trajectory Model for Ziyuan-3 Satellite. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B1, 365-372.
- RADHADEVI, P.V., SOLANKI, S.S., AKILAN, A., JYOTHI, M.V. & NAGASUBRAMANIAN, V., 2016: Performance Assessment and Geometric Calibration of Resourcesat-2. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B1, 237-244.
- TADONO, T., ISHIDA, H., ODA, F., NAITO, S., MINAKAWA, K. & IWAMOTO, H., 2014: Precise Global DEM Generation by ALOS PRISM. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-4, 71-76.
- TAKAKU, J., TADONO, T. & TSUTSUI, K., 2014: Generation of High Resolution Global DSM from ALOS PRISM. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-4, 243-248.

Verifizierung der Fokalebene(n) (FPA) von Sentinel-4

RÜDIGER HOHN¹, CHRISTIAN WILLIGES², MAX HARLANDER¹,
JÜRGEN HINGER¹, HANNES ROSSMANN² & RALF REULKE³

Zusammenfassung: Das Instrument Sentinel-4 ist Teil des Copernicus-Programms und soll erstmalig von einem geostationären Orbit aus Daten zur Überwachung der atmosphärischen Zusammensetzung mit dem Fokus auf Luftqualität und Klima liefern. Sentinel 4 wird auf dem MTG (Meteosat Third Generation) Sounder fliegen.

Ein wesentliches Subsystem sind die zwei Fokalebene(n) (FPAs) mit den entsprechenden Spektrometern, die die Wellenlängenbereiche 305nm - 500nm (ultraviolett und sichtbares Spektrum) und 750nm – 775nm (nahes Infrarot) abdecken. Da die Systeme bei ca. 215K betrieben werden, müssen die Messungen in einer Thermal Vakuum Kammer erfolgen.

Beide Fokalebene(n) wurden im Auftrag von Airbus (Airbus Defense and Space GmbH) am Deutschen Zentrum für Luft- und Raumfahrt in Berlin Adlershof vermessen.

1 Einleitung

Dieser Beitrag beschreibt die Sentinel-4 Verifikationskampagne des Qualification Models (QM) der Focal Plane Systems (FPS).

In Kapitel 2 wird das System beschrieben. Dieses besteht neben den beiden Focal Plane Arrays (FPAs) aus jeweils einer Front End Electronic (FEEs) und insgesamt einer Front End Support Electronic (FSE). In Kapitel 3 wird auf den Testaufbau eingegangen. Es werden die Parameter der Tests angegeben und der daraus abgeleitete Aufbau in einer Thermal-Vakuum-Kammer beschrieben, sowie die verwendeten Lichtquellen erläutert. Kapitel 4 befasst sich mit der Analyse der Testergebnisse. Es werden die Testergebnisse für ausgesuchte Tests präsentiert. In Kapitel 5 werden Fazit und Ausblick auf die Verifikationskampagne der beiden Flugmuster (FM1 und FM2) gegeben.

2 Detektoren, Fokalebene(n) und Elektronik

Das Sentinel 4 Instrument verfügt über zwei Fokalebene(n). Die NIR Fokalebene umfasst den Wellenlängenbereich von 750 nm bis 775 nm, die UVVIS Fokalebene den Bereich 305 nm bis 500 nm. Die Detektoren (CCDs) sind auf jeweils einer „Optischen Bank“ montiert. Diese Optische Bank wird als „Focal Plane Assembly“ bezeichnet. Das Instrument verfügt über zwei FPAs. Um die Empfindlichkeit der Detektoren zu erhöhen und parasitäre Signale zu minimieren werden die Detektoren bei 215 K betrieben. Diese Temperatur wird während des Betriebs über „Heat-Pipes“

¹ Airbus Defense and Space, 88039 Friedrichshafen, Deutschland,
E-Mail: [ruediger.hohn, maximilian.harlander, juergen.hinger]@airbus.com

² Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Institute für Optische Sensorsysteme,
12489 Berlin, Deutschland, E-Mail: [Christian.Williges, Hannes.Rossmann]@dlr.de

³ Humboldt-Universität zu Berlin, Institut für Informatik, 10099 Berlin, Deutschland,
E-Mail: reulke@informatik.hu-berlin.de

und Radiatoren erreicht. Die elektrische Versorgung der Detektoren und das Auslesen der Detektorsignale erfolgt über eine analoge und eine digitale Elektronik. Einen Überblick über die Fokalebene und der Elektronik zeigt die Abb. 1.

2.1 S4 UVN Detektoren

Bei dem UVVIS Detektor handelt es sich um einen sogenannten Split Frame Transfer CCD mit zwei Image und Memory Zonen. Der gesamte Spektralbereich ist auf dem UVVIS Detektor in zwei Bereiche, 305 nm bis 343 nm (UVVIS I) und 343 nm bis 500 nm (UVVIS II) unterteilt. Beide Bereiche können mit unterschiedlichen Integrationszeiten betrieben werden. Dies führt zu einer Erhöhung der Empfindlichkeit im UV Bereich. Zusätzlich verfügt der UVVIS I Detektorbereich über zwei Ausgänge, die die Möglichkeit bieten, mit verschiedenen Verstärkungsfaktoren zu arbeiten.

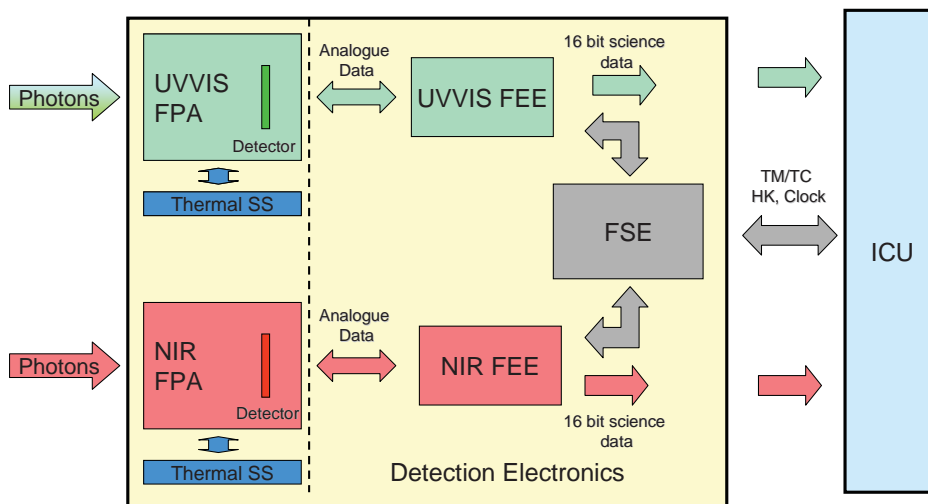


Abb. 1: FPS Block Diagramm

Bei dem NIR Detektor handelt es sich um eine Standard Frame Transfer CCD mit einer Image- und einer Memory Zone. Beide Detektoren werden bei einer Temperatur von 215 K betrieben. Die Temperaturen werden durch passive Kühlung über Heat-Pipes und Radiatoren erreicht. Zur Minimierung der parasitären externen Wärmelasten wurde eine spezielle Halterung für die Detektoren entwickelt. Zusätzlich sorgt diese Halterung für eine möglichst homogene Temperaturverteilung über den Detektor. Die nachfolgende Abbildung zeigt einen QM Detektor vorintegriert auf der Optischen Bank (FPA).

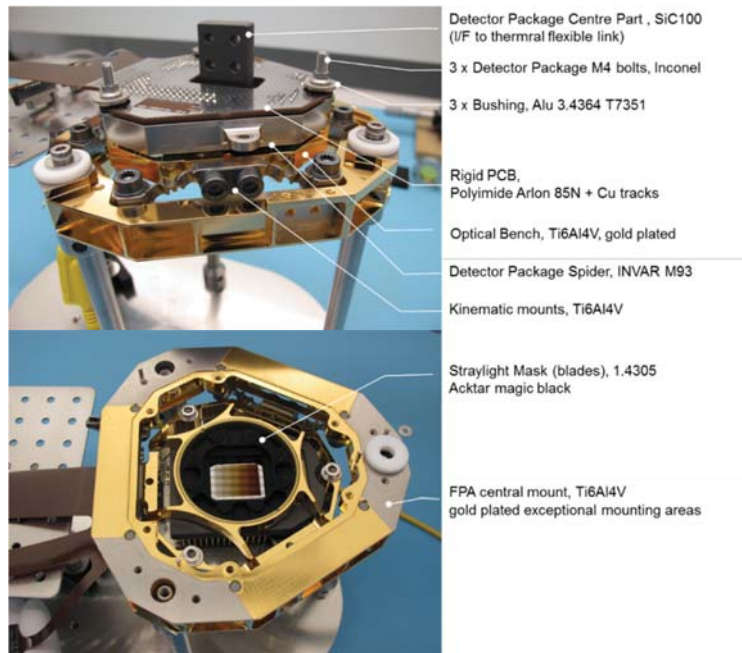


Abb. 2: QM Detektor vor-integriert auf der Optischen Bank (FPA)

2.2 S4 UVN Fokalebenen (FPA)

Das Struktur –und Thermal Design ist für beide FPAs identisch. Minimale Unterschiede liegen aufgrund unterschiedlicher Detektorgrößen im Bereich von Streulichtunterdrückungsblenden. Im Gegensatz zur nominalen Detektortemperatur von 215 K, bleibt die FPA während des Betriebes bei Raumtemperatur. Zentrale Elemente bestehen aus Titan zur Minimierung von Gewicht und zur Minimierung von parasitären Wärmelasten.

Die von den Detektoren produzierte Wärme sowie die parasitären Wärmelasten werden über einen „Kaltfinger“ aus SiC und eine gut wärmeleitende flexible Verbindung zur Heat-Pipe geleitet. Aufgrund unterschiedlicher thermischer Bedingungen im Weltraum, muß die Detektortemperatur aktiv kontrolliert, bzw stabilisiert werden. Dazu dient ein Heizer, der am Übergang Heat-Pipe /Flexible Verbindung angebracht ist.

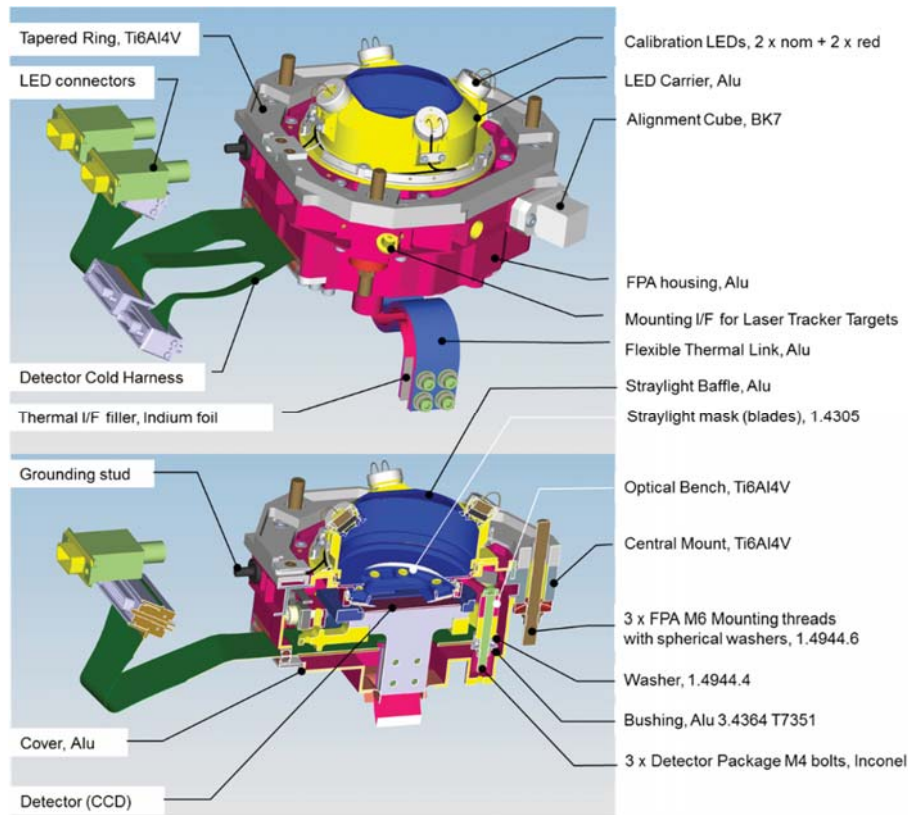


Abb. 3: UVVIS FPA design

2.3 S4 UVN Elektronik

Das Fokalebene System verfügt über einen analogen (FEE) und einen digitalen (FSE) Teil. Jeder Detektor verfügt über eine eigene individuelle FEE (siehe auch Abb. 1).

Die FSE stellt alle elektrischen Versorgungen für die beiden FEEs zur Verfügung. Die FEE verarbeitet die analogen Detektorsignale und muss damit in unmittelbarer Nähe der Detektoren platziert werden um mögliche Verluste über Leitungen zu minimieren.

Ferner sorgt die FEE für die direkte Übertragung der wissenschaftlichen Daten an die Instrument Kontrolleinheit (ICU).

3 Aufbau und Durchführung des Experiments

Die Abteilung Optik, Kalibrierung, Verifizierung (OCV) im Institut für Optische Sensorsysteme (OS) des DLR am Standort Berlin Adlershof führte in der zweiten Jahreshälfte 2016 die Validierung der Fokalebene (FPAs, Focal Plane Assemblies) des Sentinel-4 Qualifizierungsmodells (QM) durch. Nach erfolgreichem Abschluss und Auswertung der QM Testergebnisse wird OCV auch für die Verifizierung der beiden Flugmodelle (FM) verantwortlich sein.

Die Verifizierungskampagne besteht zu einem Großteil aus radiometrischen Tests, bei der eine homogene Lichtquelle (bzw. Dunkelheit) zum Einsatz kommt. Als homogene Lichtquelle dient eine Ulbrichtkugel. Zusätzlich werden ein Cross-Talk-Test mit einem 635 nm Dioden-Laser, die Kalibrierung der internen FPA-LEDs, sowie eine MTF Messung durchgeführt. Letztere findet an den baugleichen EM (Engineering Models) statt.

Tab. 1: Zusammenfassung der Messungen

Durchgeführter Test	Abgeleitete Analyse	Verwendete Lichtquelle
Linearity	Linearity	Ulbrichtkugel + Spektralfilter
	FPS gain / conversion factor	
	Full Well Capacity (linear)	
	Full Well Capacity (absolut)	
	Signal to Noise Ratio	
	Charge Transfer Efficiency (Serial)	
	Charge Transfer Efficiency (Parallel)	
Average Dark Signal	Average Dark Signal	Keine (Dunkelheit)
	Dark Signal Non-Uniformity	
	Read-out Noise	
	Dynamic Range	
Photo Response Defects	Photo Response Defects	Ulbrichtkugel + Spektralfilter
	Photo Response Non-Uniformity	
Memory Effect	Memory Effect	Ulbrichtkugel
Cross –Talk	Channel to Channel Cross Talk	635 nm Dioden-Laser
Stability over 24 h	Stability over 24 h	Keine (Dunkelheit)
Random Telegraph Signal	RTS and pixels with high dark noise	Keine (Dunkelheit)
FPA LED Calibration	FPA LED Calibration	Intrinsische LEDs der FPAs
Quanteneffizienz	QE	Ulbrichtkugel + Spektralfilter
Modulations Transfer Funktion (Durchführung am baugleichen EM)	MTF	Kollimator, Filter, Fokussieroptik

3.1 Aufbau

Zur Durchführung der radiometrischen Tests wurde eine räumlich homogene und zeitlich hochstabile Strahlungsquelle, in diesem Fall eine Ulbrichtkugel, verwendet.

Für den Cross-Talk-Test wird ein 635 nm Dioden-Laser verwendet.

3.2 Randbedingungen

Wie in der Einleitung erwähnt, sollen die Detektoren unter Einsatzbedingung untersucht werden. Dafür müssen sie, je nach Test, auf 210 K, 215 K, und 220 K temperiert und im Vakuum betrieben werden. Diese Randbedingung erfordert den Einsatz einer Thermal-Vakuum-Kammer (TVK). In der beim DLR vorhandenen TVK kann ein Druck von $5 \cdot 10^{-6}$ mbar mit Hilfe eines dreistufigen Pumpensystems erzeugt werden. Die Temperatur wird über externe Thermostate geregelt. Druck und Temperatur sind langzeitstabil über mehrere Tage.

3.3 Optischer Aufbau

Die Verwendung einer Ulbrichtkugel erfordert im Allgemeinen die Positionierung des zu vermessenden Detektors direkt in der Austrittsapertur der Kugel. Die beim DLR verwendeten Ulbrichtkugeln sind für den Einsatz unter Normalatmosphäre vorgesehen. Diese Randbedingung verhindert eine Positionierung der Detektoren direkt in der Austrittsapertur, da sie in der TVK betrieben werden müssen.

Um dennoch eine gleichmäßige Beleuchtung zu erzielen, wird auf eine Auswirkung des \cos^4 -Gesetzes zurückgegriffen: Bei hinreichend großem Abstand und entsprechendem Größenverhältnis von Detektor zu Austrittsapertur der Ulbrichtkugel wird eine homogene Beleuchtung auch von einer entfernt stehenden Ulbrichtkugel erreicht, da der IFOV (Instantaneous Field Of View) und somit der Betrachtungswinkel für sämtliche Pixel praktisch gleich ist. Das heißt, dass die Ulbrichtkugel als Quasi-Punktlichtquelle für den Detektor fungiert.

Um eine Bestrahlung mit diversen Wellenlängen zu realisieren und die Bestrahlung regeln zu können, werden zwei Filtrerräder in den Strahlengang integriert, jeweils mit acht Spektral- bzw. Neutraldichtefiltern (siehe Abb. 4).

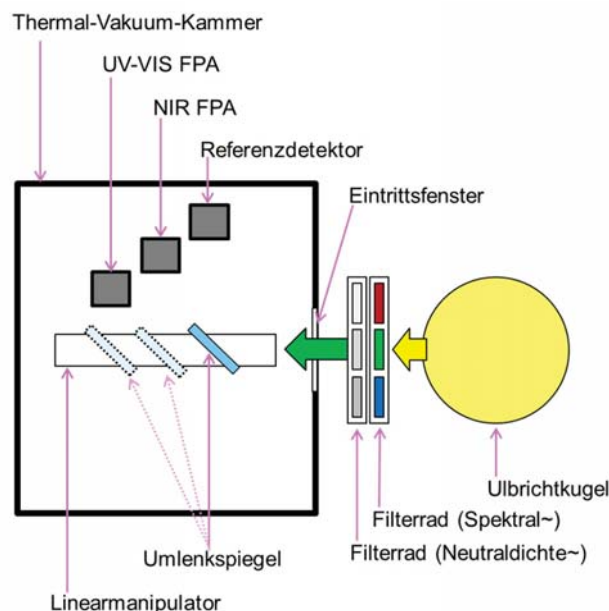


Abb. 4: Prinzipskizze des opto-mechanischen Testaufbaus: Das Licht der Ulbrichtkugel wird mit Hilfe von Filtern manipuliert, bevor es in die Thermal-Vakuum-Kammer eintritt. Über einen linear verstellbaren Umlenkspiegel erreicht das Licht mit gleicher optischer Pfadlänge den jeweiligen Detektor.

Um eine gleichmäßige Bestrahlung aller drei Detektoren zu realisieren, wird der opto-mechanische Aufbau so konzipiert, dass er das Prinzip gleicher optischer Pfadlängen ausnutzt: Über einen linear verstellbaren 90°-Umlenkspiegel werden die Detektoren abwechselnd bestrahlt. Durch räumlich versetzt angeordnete Detektoren ist die Entfernung zwischen Ulbrichtkugel und Detektor für alle drei Detektoren gleich (siehe Abb. 4).

Für den Cross-Talk-Test werden die Filterräder und die Ulbrichtkugel aus dem Strahlengang entfernt und durch einen 635 nm- Laser ersetzt. Dieser wird per Faser in einen Kollimator eingekoppelt, der den Strahl parallelisiert, über ein Teleskop wird der Strahl auf einen Durchmesser von ca. 24 mm aufgeweitet.

Durch eine Kombination zweier Linsen wird der kollimierte Strahl über eine Entfernung von ca. 70 cm auf die Detektoren fokussiert. Die Linsen sind in z-Richtung justierbar, sodass eine Fokussierung möglich ist. Der gesamte optische Aufbau des Lasermessplatzes ist auf einem x-y - Linearmanipulator montiert, der eine Positionierung auf verschiedenen Abschnitten der Detektoren zulässt.

Die theoretische Spotgröße des fokussierten Lasers ist $< 25 \mu\text{m}$. Dieser Wert wird unter realen Bedingungen aufgrund der langen optischen Pfadlänge, diversen optischer Elemente im Strahlengang, sowie thermischer und mechanischer Drifts jedoch überschritten.

3.4 Mechanischer Aufbau

Die FPAs und die zugehörige FEE besitzen eine bestimmte, nicht rechtwinklige und nicht parallele, Ausrichtung zueinander (s. Abb. 5, Position 3.1 und 3.2). Die Flexharnesse, die die Detektoren mit ihrer jeweiligen Elektronik verbinden, haben eine der Instrumentengeometrie angepasste Form und Länge. Um diese auch während der Tests in der TVK verwenden zu können, wird die Geometrie der mechanischen Halter der Satellitengeometrie nachempfunden. Dieser Aufbau ermöglicht es, verlässliche Testergebnisse zu erhalten, die unter anderem die speziellen Widerstände in den verwendeten Kabelbäumen mit berücksichtigen.

3.4.1 Kühlung

Wie oben bereits erwähnt werden die Tests an den FPAs bei Betriebstemperatur durchgeführt. Die Vorgabe ist die FPAs – je nach Test – auf 210 K, 215K sowie 220 K zu temperieren; die FEEs und FSE (front end support electronic) soll auf 293 K stabilisiert sein. Die Vorgabe war eine Stabilität von $< +/- 50 \text{ mK}$. Dafür werden zwei Kühlkreisläufe aufgebaut, deren Thermostate getrennt voneinander die FPAs und die Elektronik (FEEs und FSE) versorgen.

Die FPAs werden über Kronsteine stabilisiert, die in Reihe geschaltet sind. Aufgrund der größeren thermischen Last wird zuerst die UVVIS-FPA versorgt und anschließend die NIR-FPA (Abb. 5, Position 2.3, 3.3, 4). So können beide FPAs stabil auf der gleichen Temperatur gehalten werden. Die FEEs und die FSE sind ebenfalls über Kronsteine an eine Temperaturregelung angeschlossen und auch in Reihe geschaltet: UVVIS-FEE, NIR-FEE, FSE.

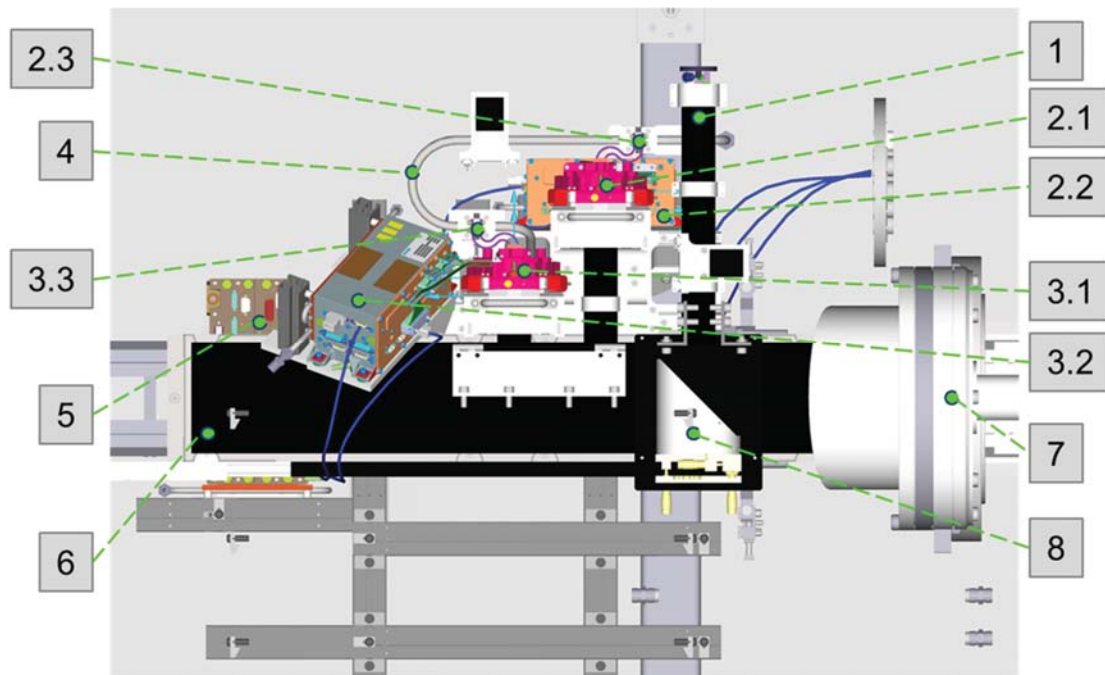


Abb. 5 Aufsicht auf eine Konstruktionszeichnung des Versuchsaufbaus.

- | | | | |
|-------|---------------------|-----|---------------------------|
| 1. | Referenzdetektor | 3.3 | NIR Kühlanschluss |
| 2. | UVVIS FPA | 4. | FPA Kühlkreislauf |
| 2.2. | UVVIS FEE | 5. | FSE |
| 2.3. | UVVIS Kühlanschluss | 6. | Linearmanipulator |
| 3.1. | NIR FPA | 7. | Optischer Eingang der TVK |
| .3.2. | NIR FEE | 8. | Umlenkspiegel |

3.5 Durchführung

Um die Vielzahl der Tests möglichst schnell durchführen zu können, ist eine weitgehende Automatisierung der Testabläufe notwendig. Dies wird über das Programm AutoIt® realisiert, mit dem die zur Verfügung stehende Matlab®-Oberfläche autonom gesteuert werden kann.

Auf diese Weise können Einstellungen am Sentinel-4-Testequipment (Integrationszeit, Anzahl der Frames, Aufnahmemodi, etc.), als auch an der GSE (Spiegelposition, Temperatur, etc.) vorgenommen werden.

4 Datenauswertung und erste Ergebnisse

Im Folgenden sollen einzelne Messungen genauer beschrieben werden:

- PRNU und Shade
- Linearität & Nichtlinearität
- Overall system gain aus der Analyse der PTC (Photon Transfer Curve)

4.1 Signalerzeugung

Eine Übersicht zu diesem Thema findet man in KOPEIKA (1998) und JANESICK (2001). Demnach lässt sich das erwartete Signal \hat{S} in DN (Digital Numbers) folgendermaßen bestimmen:

$$\hat{S}[\text{DN}] = \eta_{\text{SG}}^{\text{Det}} \left[\frac{\text{DN}}{e^-} \right] \cdot \eta_{\text{QE}}(\lambda) \left[\frac{e^-}{\text{photon}} \right] \cdot \tau_{\text{int}} [\text{s}] \cdot A_{\text{Det}} [\text{m}^2] \cdot \frac{\lambda}{hc} \left[\frac{\text{m}}{\text{Ws}^2 \text{ m/s}} \right] \cdot E \left[\frac{\text{W}}{\text{m}^2} \right] + \text{DS} \quad (1)$$

Tab. 2: Einflussgrößen zur Signalberechnung

$\eta_{\text{QE}}(\lambda)$	Quanteneffizienz	$[e^-/\text{photon}]$
$\eta_{\text{SG}}^{\text{Det}}$	Overall system gain	$[\text{DN}/e^-]$
A_{Det}	Detektorfläche	$[\text{m}^2]$
τ_{int}	Integrationszeit	$[\text{s}]$
E	Bestrahlungsstärke	$[\text{W}/\text{m}^2]$
DS	Temperaturabhängiges Dunkelsignal	$[\text{DN}]$
λ	Wellenlänge	$[m]$

Wir erwarten mindestens drei Rauschkomponenten: Photonenrauschen, Dunkelstrom und das Lese- oder Ausleserauschen (read noise). Während Dunkelstrom- und Photonenrauschen Poisson verteilt sind, wird der read-noise durch eine Normalverteilung beschrieben. Durch die Poissonverteilung des Signals und des Dunkelstroms ergibt sich für das Rauschen die folgende Relation:

$$\sigma_s^2 = \eta_{\text{SG}}^2 \cdot \hat{S} + \sigma_k^2 \quad (2)$$

Diese Gleichung wird hier zur Bestimmung des Overall System Gains genutzt (JANESICK 2007).

4.2 PRNU und Shade

Die Strahlungsverteilung vor dem Detektorchip (Abb. 6) war inhomogen (ca. 20% PV). Dieser Shade und die PRNU wurden vor der Auswertung bestimmt und korrigiert. Die Korrektur beruht auf der folgenden Überlegung:

Wir erwarten einen linearen Zusammenhang zwischen Signal \hat{S}_{ij}^k und Exposer (Bestrahlungsstärke \times Integrationszeit) $\hat{S}_{ij}^k \approx a_{ij} \cdot E_{ij}^k + b_{ij}$ bezogen auf die Pixelposition (ij) und der k -ten Messung (z.B. durch Variation der Integrationszeit). Durch den Shade bei der Bestrahlung wird Exposer orts-, bzw. pixelabhängig. Die Korrektur von PRNU und Shade können nicht unabhängig voneinander vorgenommen werden. Die Korrektur bewirkt einen Ausgleich des Anstiegs für jedes Pixels ohne Offset, bezogen auf eine homogene Referenzbestrahlung \bar{E}^k , die durch einem mittleren Anstieg c beschrieben wird:

$$c = \text{mean}(a_{ij}) \quad c_{ij} = \frac{c}{a_{ij}} \quad \tilde{S}_{ij}^k = c_{ij} \cdot (\hat{S}_{ij}^k - b_{ij}) \quad (3)$$

In den folgenden Auswertungen wird mit den korrigierten Signalwerten \tilde{S}_{ij}^k gearbeitet. Für die Charakterisierung der PRNU wurde eine medianbasierte Analyse des Anstiegs verwendet:

$$d_{ij} = \text{median}(c_{ij}) \quad \text{PRNU} = \text{stddev}(\text{abs}(d_{ij} - c_{ij})) \quad (4)$$

Bezieht man die Standardabweichung auf die maximale Dynamik des gemessenen Signals erhält man für die PRNU < 1% für alle Detektoren (Forderung war maximal 4%).

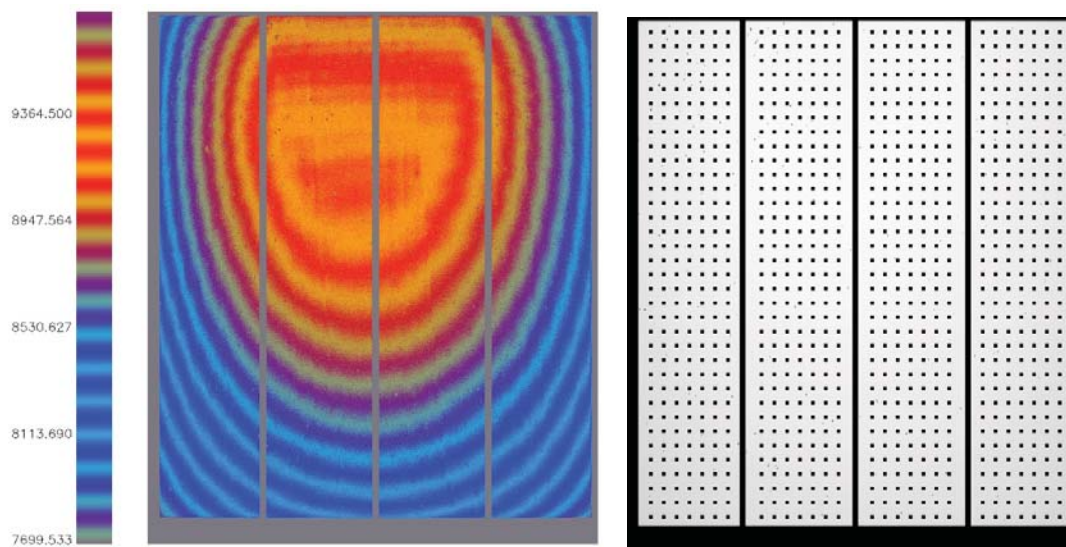


Abb. 6: VIS2 Detektor mit Shade und Referenzpunkten, für die Abhängigkeiten von der Integrationszeit separat untersucht wurden

Den Unterschied zwischen der Analyse der Linearität an den Referenzpunkten ohne und mit Korrektur ist in der Abb. 7 zu sehen. Auf Grund des Shades gehen die Signale für die einzelnen Pixel unterschiedlich in die Sättigung.

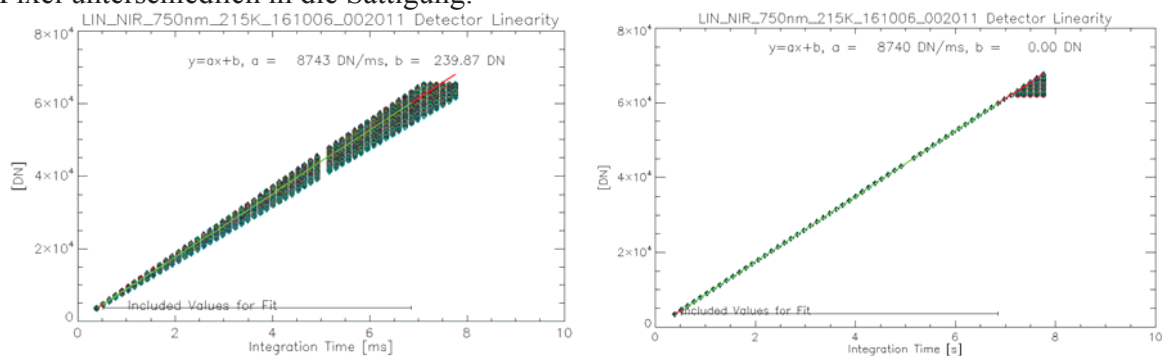


Abb. 7: Linearität an den Referenzpunkten (siehe Abb. 7). Original (links) und PRNU-korrigiert (rechts)

4.3 Nichtlinearität und Korrektur der Nichtlinearität

Der beschriebene Ansatz erlaubt gleichzeitig eine Analyse der Nichtlinearität. Hier wurde in dem interessierenden Bereich die maximale Abweichung von der Linearität analysiert. In der Abb. 8 ist die räumliche Abhängigkeit zu sehen. Die Abweichung korreliert offensichtlich mit dem Shade, größere Signale erzeugen größere Abweichungen. Aus dem Histogramm lässt sich ablesen, dass die Abweichung absolut für den hier betrachteten Detektor 40-50 DN oder relativ $< 0.1\%$ ist.

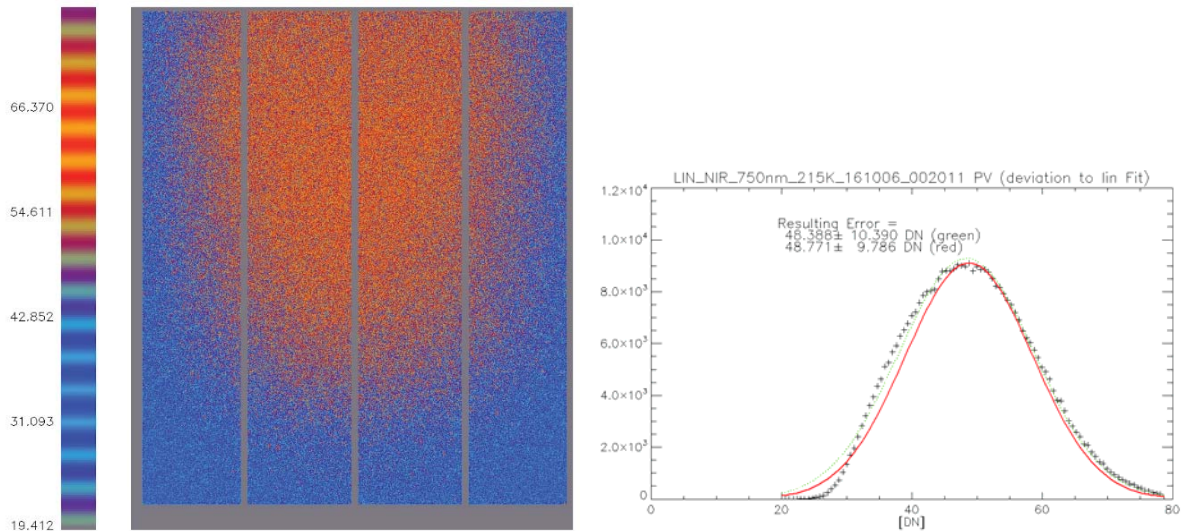


Abb. 8: Abweichung von der Linearität (PV)

In der Abb. 9 (links) ist die Abweichung von der Linearität für die Referenzpixel als Funktion der Intensität aufgetragen. Die Abweichungen sind vergleichsweise gering und lassen sich durch ein Polynom höherer Ordnung gut korrigieren (Abb. 9, rechts).

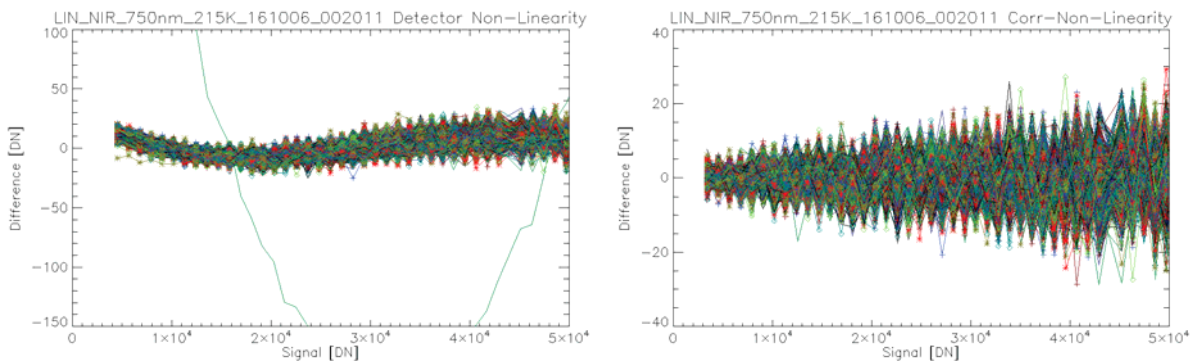


Abb. 9: Abweichung von der Linearität für die Referenzpixel (links) und Korrektur der Nichtlinearität (rechts)

4.4 Bestimmung des System Gain

Mit dem Overall System Gain η_{SG} lässt sich die erzeugte Anzahl von Ladungen in digitale Werte ohne Kenntnis der Quanteneffizienz umrechnen. Dazu wird die Photon-Transfer Methode (Janesick 2007) verwendet. Diese nutzt die Poissoneigenschaften der einfallenden Strahlung und

des Dunkelstroms. Es kann gezeigt werden, dass es einen linearen Zusammenhang zwischen Varianz und Signal gibt. Der Anstieg der Geraden ist η_{SG}^{Det} . Beispielhaft soll auf Abb. 10 verwiesen werden. Für den NIR-Detektor erhält man 0,084 DN/e oder 12 e pro DN.

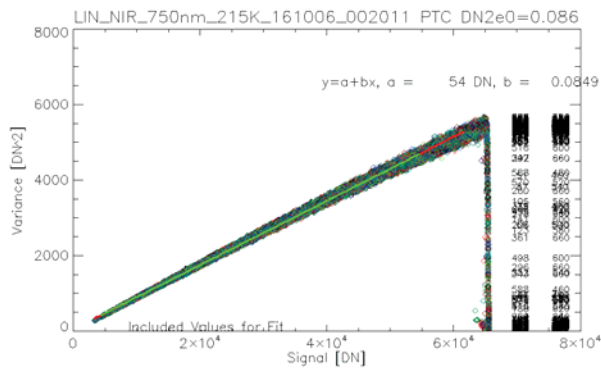


Abb. 10: Varianz als Funktion des Signals für die Referenzpixel

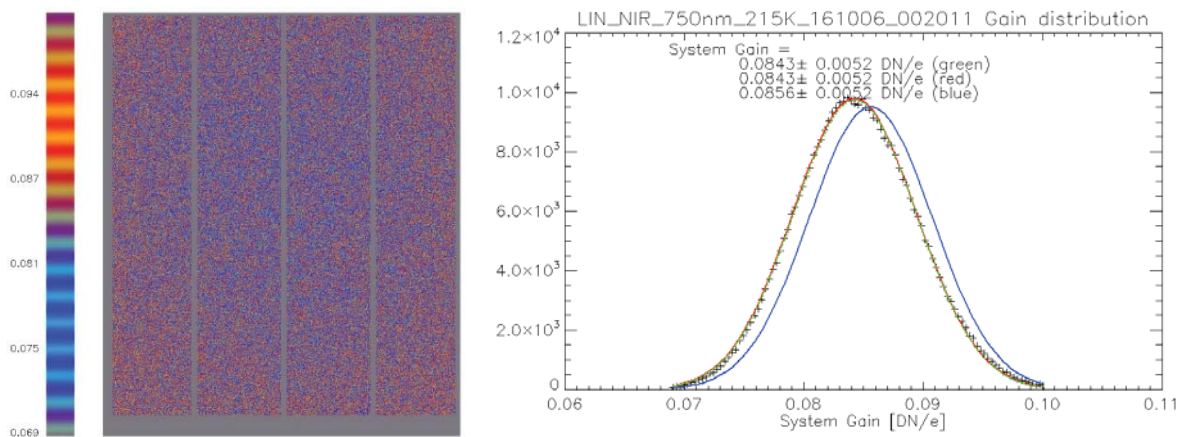


Abb. 11: Räumliche Abhängigkeit von η_{SG}^{Det} (links) und Histogramm (rechts)

5 Fazit & Ausblick

Die Sentinel-4-QM-Verifikationskampagne am DLR Institut für Optische Sensorsysteme (OS) wurde erfolgreich abgeschlossen. Die Untersuchungen ergeben, dass die Detektoren und die zugehörige Elektronik unter den spezifizierten Bedingungen hervorragend arbeiten. So sind z.B. die gemessenen Werte der PRNU deutlich unterhalb der erlaubten Obergrenzen (vgl. Tab. 3). Auch die Nichtlinearität kann sehr gut korrigiert werden und weist nach einer Korrektur Abweichungen $< 0,1\%$ auf. In Anbetracht der erhaltenen Ergebnisse wird eine Optimierung der FM-Detektoren als nicht erforderlich betrachtet.

Während der Verifikationsphase ist eine Anomalie aufgetreten: Die geplante räumlich-homogene Ausleuchtung der Detektoren weist ein deutliches Shading von ca. 20 % ... 30 % auf (Abb. 6). Eine Analyse des Versuchsaufbaus deutet darauf hin, dass eine nicht korrekt justierte Blende im Strahlengang für diese Vignettierung verantwortlich ist. Durch Anpassung des Aufbaus wird diese Anomalie bei den Flugmodellen vermieden.

Tab. 3: PRNU Vergleich der beim DLR erhobenen Daten zu den geforderten Spezifikationen der ESA

	Detektor		
	UVVIS 1	UVVIS 2	NIR
Resultat DLR: Standardabweichung (%)	0.288	0.187	0.18
Resultat DLR: Anzahl der Pixel außerhalb von 3σ (%)	0.035	0.011	0.47
Spezifizierte Obergrenze (%)	4.0	2.0 ... 1.0 (in Abhängigkeit der Wellenlänge)	4.0

6 Danksagung

Die vorliegende Arbeit wurde im Auftrag und mit Unterstützung der ESA durchgeführt. Die Autoren möchten allen Kollegen von Airbus Defence and Space, ESA und EUMETSAT, sowie allen Partnerfirmen innerhalb des Sentinel-4 Industriekonsortiums danken. Sie alle haben einen sehr wertvollen Beitrag zu diesem herausfordernden Programm geleistet.

Dieser Artikel wurde mit finanzieller Hilfe der EU produziert. Die hierin geäußerten Ansichten spiegeln nicht die offizielle Haltung der EU wieder.

7 Literaturverzeichnis

JANESICK, J.R., 2001: Scientific Charge-Coupled Devices, SPIE Press Book.

JANESICK, J.R., 2007: Photon Transfer. SPIE Press Book.

KOPEIKA, N.S., 1998: A System Engineering Approach to Imaging, SPIE Press Book.

Mit Kamera, Maßstab und Laptop – Best Practice zur Erstellung von 3D-Modellen für eine Museumsausstellung am Beispiel vietnamesischer Kulturobjekte

THOMAS P. KERSTEN¹ & MAREN LINDSTAEDT¹

Zusammenfassung: In diesem Beitrag werden Beispiele vietnamesischer Kulturgüter von unterschiedlicher Größe und Bedeutung vorgestellt, die aus Bildsequenzen 3D rekonstruiert wurden, um z.B. digitale foto-realistische Replikate für eine Ausstellung und/oder eine 3D-Visualisierung in virtuellen Museumsanwendungen herzustellen. Für die Ausstellung „Schätze der Archäologie und Kultur Vietnams“ wurden ausgewählte Kulturobjekte Vietnams durch das Labor für Photogrammetrie und Laserscanning der HafenCity Universität Hamburg, im September 2015 in verschiedenen Museen Vietnams durch Bildsequenzen mit einer digitalen Spiegelreflexkamera Nikon D800 aufgenommen und detailliert dokumentiert. Die Schätze Vietnams, die noch niemals außerhalb des Landes zu sehen waren, werden erstmals in Deutschland im Rahmen einer Ausstellung in den Städten Herne, Chemnitz und Mannheim im Zeitraum von Oktober 2016 bis Februar 2018 gezeigt wird. Der Arbeitsablauf von der 3D-Objektaufnahme über die Modellierung bis zur Visualisierung und dem 3D-Druck werden als Best Practice beschrieben. Die Ergebnisse der 3D-Modelle und deren Integration in virtuelle Touren werden präsentiert.

1 Einleitung

Kulturgüter sind bedeutende Zeugnisse der menschlichen Vergangenheit. Dieses historische Erbe ist heute aufgrund zunehmender Zerstörung durch Krieg, Terrorismus und Vandalismus sowie durch schleichende Verwitterung weltweit in großer Gefahr. Eine erhebliche Anzahl von Kulturobjekten wurde bereits in der Vergangenheit verschleppt, verkauft oder zerstört. Beispiele für die sinnlose Zerstörung stellen die großen Buddha-Statuen von Bamiyan in Afghanistan (GRÜN et al. 2002) und die archäologisch bedeutsame Stätte Palmyra in Syrien dar (WAHBEH & NEBIKER 2016). Andererseits lagern viele Objekte nicht restauriert und konserviert in Magazinen und verlieren so nach Jahren ihre Verzerrungen oder gar Form. Um den vollständigen Verlust aller Informationen dieses Kulturerbes zu vermeiden, ist die 3D-Aufnahme und Dokumentation mit modernen Messmethoden eine hervorragende Methode. Zur Hilfe kommt die stetig zunehmende Leistungsfähigkeit von Internet und Computertechnologie kombiniert mit einer rasanten Entwicklung der entsprechenden Rechenalgorithmen im Bereich der Computer Vision und Photogrammetrie. Dadurch wird eine effiziente und flexible Rekonstruktion der 3D-Geometrie von Kulturobjekten für deren Sicherung, für zukünftige Konservierungsaufgaben und für eine ortsunabhängige Präsentation möglich.

Da in den meisten Fällen die wertvollen und interessanten archäologischen Fundstücke und Exponate aufgrund ihrer hohen nationalen Bedeutung, aus rechtlichen Gründen oder auch wegen des Transportrisikos im Herkunftsland verbleiben, müssen Verfahren und Werkzeuge entwickelt

¹ HafenCity Universität Hamburg, Labor für Photogrammetrie & Laserscanning, Überseeallee 16, D-20457 Hamburg, E-Mail: [Thomas.Kersten, Maren.Lindstaedt]@hcu-hamburg.de

werden, um diese einmaligen Objekte dennoch einer interessierten Öffentlichkeit zu zeigen. So können Kulturgüter unterschiedlicher Größe (von der Münze bis zur großen Tempelanlage) mithilfe hochentwickelter Software aus Bildsequenzen kommerzieller digitaler Kameras 3D aufgenommen und modelliert werden. Kamera, Maßstab und Laptop (mit entsprechender Software) bilden damit Bestandteile sogenannter Low-Cost Systeme, die für eine Vielzahl von Anwendungsbereichen (Archäologie, Restauration, Denkmalpflege, Visualisierung, Analyse von Baukonstruktion und deren Beschädigung, etc.) eingesetzt werden können bzw. bereits eingesetzt wurden (REMONDINO et al. 2008, BARAZZETTI et al. 2009, KERSTEN & LINDSTAEDT 2012, KERSTEN & LINDSTAEDT 2014). Betrachtet man die Tatsache, dass der Mitnahme und Einführung eines Streifenprojektionssystems oder eines Laserscanners in ein fremdes Land oftmals diverse bürokratische und manchmal auch transporttechnische Hürden im Wege stehen, ist der Einsatz einer Kamera unkompliziert und ohne viel Aufwand zu realisieren. Wenn man in den vielseitigen Anwendungsbereichen bewährte Methoden einsetzt und den Workflow anhand der Aufgaben entsprechend optimiert, dann spricht man von Best Practice. Der Begriff stammt aus der angloamerikanischen Betriebswirtschaftslehre und bezeichnet bewährte, optimale bzw. vorbildliche Methoden, Praktiken oder Vorgehensweisen, die im Folgenden anhand der Aufnahme und Auswertung von vietnamesischer Kulturgüter gezeigt werden.

2 Photogrammetrische Aufnahme

In diesem Beitrag werden fünf Beispiele vietnamesischer Kulturgüter von unterschiedlicher Größe und Bedeutung (Abb. 1) vorgestellt, die aus Bildsequenzen zur Erstellung digitaler und analoger Replikate 3D rekonstruiert wurden. Für die Ausstellung «Schätze der Archäologie Vietnams» (REINECKE & MÜHLENBROCK 2016), die in Deutschland in den Städten Herne, Chemnitz und Mannheim im Zeitraum von Oktober 2016 bis Januar 2018 gezeigt wird, wurden durch das Labor für Photogrammetrie & Laserscanning der HafenCity Universität Hamburg ausgewählte Kulturobjekte Vietnams, die noch niemals außerhalb Vietnams zu sehen waren, im September 2015 in verschiedenen Museen Vietnams durch Bildreihen aufgenommen und detailliert dokumentiert. Die photogrammetrische Aufnahme der Objekte erfolgte aus verschiedenen Perspektiven und Höhen mit einer digitalen Spiegelreflexkamera Nikon D800, die einen Vollformatsensor von 36 mm × 24 mm mit einer Auflösung von 36 Mio. Pixeln aufweist. Durch eine stabile Kamerapositionierung auf einem Stativ konnten die Aufnahmen mit einer langen Brennweite (ca. 35 mm oder 80 mm) und der kleinen Blende (22) gemacht werden, die bei den teilweise ungünstigen Beleuchtungsverhältnissen eine optimale Tiefenschärfe und eine lange Belichtungszeit erlaubte. Durch die Aufnahme mit HDR (High Dynamic Range) in einer Belichtungsreihe von drei Fotos (über-, unter- und normal belichtet) konnte eine optimale Belichtung des Objektausschnittes garantiert werden, da die drei Fotos bereits in der Kamera zu einem Bild zusammengerechnet wurden. In den Objektraum wurden während der photogrammetrischen Aufnahme kalibrierte Maßstäbe mit einer Länge von 14 cm, 28 cm und 44 cm um das Objekt gelegt, um die aus den Fotos generierten 3D-Informationen präzise zu skalieren. Bei der Auswertung der verschiedenen Objekte konnte eine Genauigkeit der Maßstäbe von ca. 0,05 mm bestimmt werden.

Um ein Objekt wie z.B. den Schädel rundherum und komplett modellieren zu können, wurden in diesen Fällen zwei Bildverbände gemacht, einer der Ober- und einer der Unterseite. Bei dieser Vorgehensweise sollte beachtet werden, dass zwischen den beiden Bildverbänden ausreichend Überlappung besteht, um schließlich ein komplettes Modell berechnen zu können.



Abb. 1: Schätze der Archäologie Vietnams als Aufnahmeobjekte – v.l.n.r.: hölzerne Buddha-Figur, Schädel mit Ohrring, Goldmasken, Mukhalinga („Phallus mit Gesicht“) und Trommel

3 Bildauswertung und 3D-Modellierung

Im Folgenden werden die Auswertungen und 3D-Modellierungen der fünf wichtigsten Kulturobjekte kurz vorgestellt. Die Auswertung der Bilddaten erfolgte für alle Objekte mit der Software PhotoScan Professional Edition 1.2. von der Firma Agisoft aus St. Petersburg in folgenden Schritten. Nach dem Import der Fotos wurden die Bildorientierungen und die Kamerakalibrierung pro Bildverband berechnet, in dem alle Bilder durch automatische Bildpunktmessungen in Form einer dünnen Punktwolke miteinander zu einem Bildverband verknüpft wurden. Eine Skalierung der generierten 3D-Punktwolke erfolgte durch präzise manuelle Messung der Maßstäbe. Für den Fall mehrerer Bildverbände eines Objektes, gilt es zunächst diese Verbände alle in einem zu vereinen. Dafür wird in einer (niedrigen) Auflösung eine Vermaschung gerechnet und anhand dieser wird aus dem Objektraum zurück in den Bildraum eine Maskierung des Objektes berechnet, die nur das Objekt ohne störende Umgebung beinhaltet. Anschließend werden die Bilder der beiden Bildverbände unter der Berücksichtigung der Masken nochmals orientiert, um so alle Bilder in einem Verband vorliegen zu haben. Gelingt hierbei die Orientierung der Bilder nicht vollständig, so lassen sich zusätzlich manuelle Verknüpfungspunkte in den entsprechenden Bildern generieren. Nun lässt sich eine dichte Punktwolke des gesamten Objektes ohne fehlende Bereiche berechnen. Die Punktwolke lässt sich direkt in PhotoScan bereinigen, filtern, segmentieren und vermaschen. Oder man wählt für ausführlichere Editierungsmöglichkeiten den Export, um in Geomagic oder einem anderen Modellierungsprogramm weiterzuarbeiten. Als letzten Schritt wird das bereinigte Modell dann anhand der vorhandenen Bilder mit einer hochauflösenden Textur in PhotoScan versehen.

3.1 Der hölzerne Buddha aus Südvietsnam

Das erste Objekt war aufnahmetechnisch auch das anspruchsvollste Projekt, da die hölzerne Buddha-Statue sehr fragil und relativ flach ist, was eine Verknüpfung der beiden Seiten (Vorder- und Rückseite) bei der späteren Modellierung erschwerte. Die im März 2004 entdeckte und 1.90 m große Buddha-Figur stammt aus Gò Tháp (Mekong-Delta, Provinz Đồng Tháp, Südvietsnam) und befindet sich im Museum Đồng Tháp in Cao Lãnh. Sie ist der Óc Eo-Kultur zu zurechnen und sie

wurde im 2. Jahrhundert n. Chr. hergestellt (REINECKE 2016). Ein senkrechter Aufbau der Figur für die Aufnahme war aus Sicherheitsgründen nicht möglich. Die Statue wurde in 378 Fotos (193 Vorder- und 185 Rückseite) mit der Nikon D800 ($f=85\text{ mm}$) auf zwei Hocker liegend aufgenommen (Abb. 2), wobei aus Gründen der Belichtungsqualität von jedem Standpunkt jeweils eine HDR- und eine normale Belichtungsaufnahme gemacht wurde. Wegen der flachen Geometrie der Figur und der Aufnahmekonfiguration für die Vorder- und Rückseite gab es wenig überlappende Bereiche für die Verknüpfung der beiden Objektteile. Daher mussten korrespondierende Punkte in beiden Objektteilen manuell gemessen werden, um so eine geometrisch korrekte Verknüpfung der beiden Bildverbände zu gewährleisten. Das 3D-Modell der Figur (Abb. 2 rechts) wurde in vier verschiedenen Auflösungsstufen im Fileformat OBJ erstellt, um unterschiedlichen Anforderungen vom 3D-Druck bis zur Visualisierung gerecht zu werden: 5 Mio. Dreiecke (368 Mbyte), 2 Mio. Dreiecke (147 Mbyte), 600.000 Dreiecke (52 Mbyte) und 200.000 Dreiecke (23 Mbyte).



Abb. 2: Aufnahmeaufbau für den hölzernen Buddha (links) und 3D-Modell des Buddha-Kopfes (rechts)

3.2 Ein Schädel mit Ohrring aus Ho-Chi-Minh-Stadt

Das zweite vorgestellte Objekt ist ein Schädel mit Ohrring der Sa Huỳnh-Kultur (Abb. 1), aus dem 3.-1. Jahrhundert v. Chr. vom Gräberfeld Giồng Cá Vồ, der im Nationalmuseum in Ho-Chi-Minh-Stadt erfasst wurde. Es wurde die Ober- und Unterseite des Schädels durch zwei Bildsequenzen mit einer gesamten Anzahl von 193 Fotos um das Objekt herum und aus verschiedenen Höhen aufgenommen (Abb. 3 links). Leider waren Schädel und Unterkiefer nicht fest miteinander verbunden, sodass nach Drehung des Objektes auf die Unterseite der Unterkiefer und die Schädelunterseite nebeneinander liegend fotografiert werden mussten. Dieses Missgeschick erschwerte die Auswertung erheblich, da die exakte Position des Unterkiefers nun aus der ersten Bildserie der Oberseite auch mit rekonstruiert und eine aufgetretene Deformation im Unterkiefer korrigiert werden musste. Das texturierte 3D-Modell des Schädels (Abb. 3 rechts) besteht aus knapp 1 Mio. Dreiecken und es weist eine Datenmenge von 102 Mbyte auf.

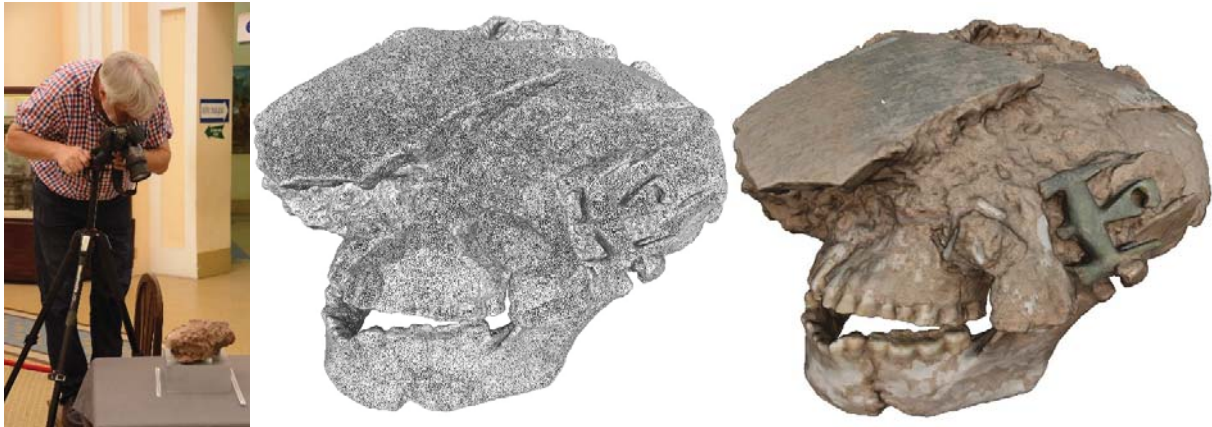


Abb. 3: Aufnahme des Schädels mit Ohrring (links), Schädel als Dreiecksvermaschung (Mitte) und als texturiertes 3D-Modell (rechts)

3.3 Goldmasken aus Südvietnam

Im Museum der Provinz Bà Rịa - Vũng Tàu wurden zwei Goldmasken (Größe – $10,9 \times 4,5$ cm und $9,7$ cm \times $6,0$ cm, Dicke – ca. 1 mm) aufgenommen (Abb. 4), die als Grabbeigaben in zwei verschiedenen Gräbern des 1. Jahrhunderts v. Chr. auf dem Fundplatz Giồng Lón in Südvietnam gefunden wurden. Die zwei Seiten der Masken wurden mit 107 bzw. 60 Fotos aus verschiedenen Positionen rundherum so gut dokumentiert, dass man alle Bearbeitungsspuren detailliert sehen kann. Aus den Bildsequenzen wurden jeweils ein Oberflächenmodell für die Vorder- und Rückseite berechnet, die anschließend zu einem 3D-Modell mit einem Datenvolumen von 20 Mbyte (Goldmaske 1: 124.249 Punkte und 248.494 Dreiecke) zusammengesetzt wurden. Das Zusammenfügen der beiden Seiten erfolgte über vier Punkte, die als Löcher in der Ecken der Goldmasken gekennzeichnet waren.



Abb. 4: Aufnahme-Setup für eine Goldmaske (links), 3D-Modell der Goldmaske in Vorder- und Rückansicht (Mitte) und detaillierter Ausschnitt der Dreiecksvermaschung (rechts)

3.4 Mukhalinga der Cham-Kultur

Im Museum Mỹ Sơn bei Đà Nẵng in Zentralvietnam (Provinz Quảng Nam), etwa 30 km südwestlich der ehemaligen Hafenstadt Hội An wurde eine Mukhalinga der Cham-Kultur (aus dem 8. Jahrhundert) mit 184 Fotos aus verschiedenen Höhen und Perspektiven aufgenommen. Die Mukhalinga („Phallus mit Gesicht“), eine frühe Erscheinungsform Shivas mit der Darstellung

seines Gesichts, hat die stattlichen Dimensionen von 1,26 m Höhe, 0,41 m Breite und 0,41 m Tiefe. Als zusätzliche Skalierung wurden zwei 2 m-Zollstöcke in horizontaler und vertikaler Aufstellung in den Objektraum platziert (Abb. 5 links). Aus den Bilddaten wurden 3D-Modelle (Abb. 5 rechts) in unterschiedlicher Auflösung erstellt, die die Grundlage für verschiedene Anwendungen, wie z. B. 3D-Druck (hohe Auflösung) und 3D-Visualisierung auf einem Standard-PC (niedrige bis sehr niedrige Auflösung), bilden.



Abb. 5: Foto (links) und 3D-Modell der Mukhalinga aus verschiedenen Blickrichtungen (rechts)

3.5 Die Bronzetrommel von Cồ Loa

Eine über 2000 Jahre alte Bronzetrommel von Cồ Loa wurde mit 165 Fotos im Museum der Stadt Hanoi aufgenommen. Die Aufnahmepositionen der Kamera und einer der beiden verwendeten Maßstäbe sind in Abb. 6 (links) dargestellt.

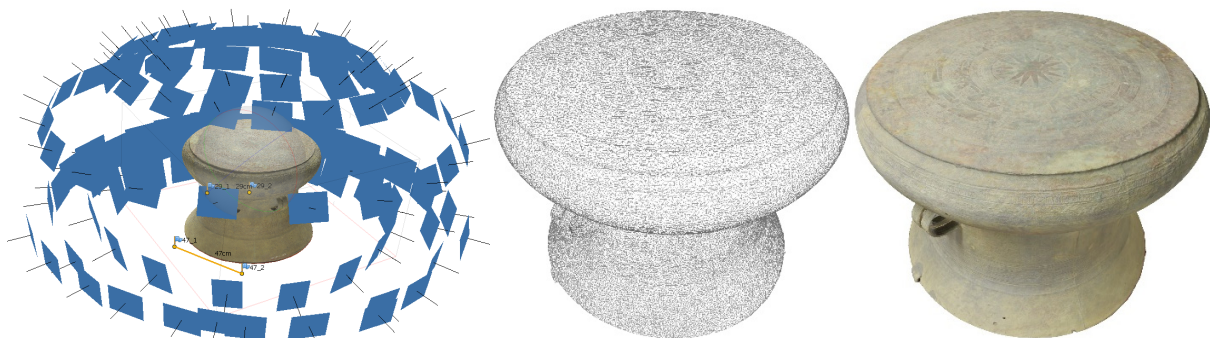


Abb. 6: Aufnahmeanordnung für die Trommel (links), Dreiecksvermaschung (Mitte) und texturiertes 3D-Modell der Trommel (rechts)

Die Bildorientierungen und die Kamerakalibrierung wurden auf einem Notebook (Schenker XMG 64bit Betriebssystem, 32 GB RAM, 2 Nvidia Geforce GT780M, CPU Intel Core i7-4940MX CPU 3.1 GHz) in 37 Minuten berechnet, während die Erstellung einer dichten Punktwolke mit 42,5 Mio. Punkten fünf Stunden und 21 Minuten dauerte. Aus diesen Daten wurden ebenso 3D-Modelle (Abb. 6) in unterschiedlicher Auflösung erstellt: a) hohe Auflösung (100%) mit 8.501.538 Dreiecken und einer Filegröße von 940 Mbyte, b) niedrige Auflösung (11%) mit 945.093 Dreiecken und einer Filegröße von 99,7 Mbyte und c) sehr niedrige Auflösung (1,2%) mit 100.000

Dreiecken und einer Filegröße von nur 12,4 Mbyte, was in diesem Fall einer Datenreduktion um den Faktor 83 entspricht. Zur Beurteilung der erstellten 3D-Modelle wurde mit der Software Geomagic Studio ein 3D-Vergleich zwischen den drei Datensätzen berechnet, der nur geringe durchschnittliche Abweichungen von kleiner als 0,1 mm bestätigte (Abb. 7).

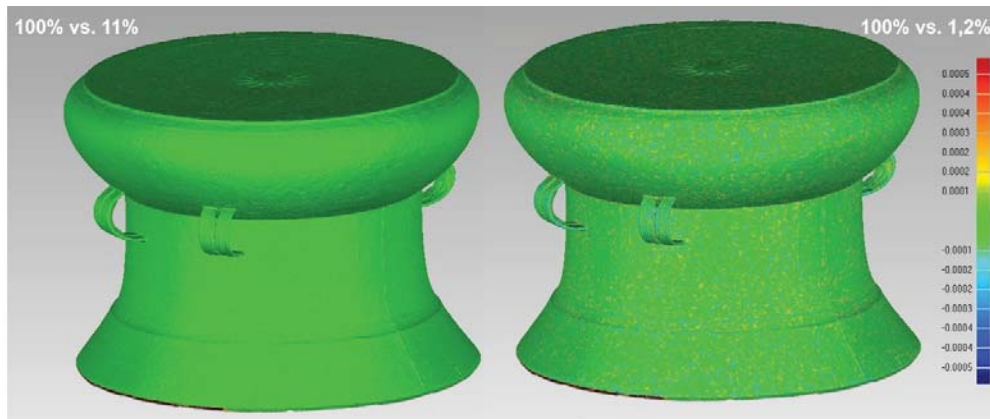


Abb. 7: 3D-Vergleich zwischen hochaufgelösten (100%) und reduzierten Datensätzen (11% und 1,2%) der Trommel (Einheit Farbskala in [m])

4 Replikate durch 3D-Druck

Für die aktuelle Ausstellung «Schätze der Archäologie Vietnams» im LWL Museum für Archäologie in Herne hat die Firma ARC-TECH GbR (Birstein, Hessen) drei Exponate (Buddha, Schädel und Mukhalinga) als 3D-Replikat angefertigt (Abb. 8).



Abb. 8: 3D-Replikate als Ausstellungsexponate im LWL-Museum für Archäologie in Herne: v.l.n.r. Buddha, Schädel und Mukhalinga

Die Replikate wurden aus den photogrammetrisch erstellten 3D-Daten der Originale als 3D-Drucke 1:1 hergestellt. Der Schädel mit Ohrring wurde monochrom (Gipsfarbe) mit einer Schichtstärke von 0,1mm im ColorJet Printing Verfahren mit einem ZPrinter ProJet 660Pro 3D-Drucker in ca. 8-9 Stunden gedruckt, indem im generativen Schichtbauverfahren PU Harz

infiltriert wurde. Der Holzbuddha und der Mukhalinga wurden im Sanddruck (dunkelgrau) mit einer Schichtung von 0,3 mm in ca. 30-35 Stunden gefertigt, indem beide Objekte zusammen in einem Druckraum bzw. Druckvorgang verarbeitet wurden. Auch hier wurde in dem Druckvorgang PU Harz infiltriert, um die Repliken robuster und widerstandsfähiger zu machen. Die Oberflächen wurden sorgfältig überarbeitet und nachträglich von Hand koloriert, was insgesamt ca. 120 Stunden in Anspruch nahm.

5 Virtuelle Touren für Weltkulturerbestätten

Für eine museale Darstellung von Weltkulturerbestätten in Vietnam wurden von fünf verschiedenen Orten in Mittel- und Nordvietnam virtuelle Touren erstellt, die aus voll-sphärischen Panoramen (360-Grad-Blick) von unterschiedlichen Standpunkten generiert wurden. Beispielhaft wird hier die virtuelle Tour von Mỹ Sơn dargestellt (Abb. 9).

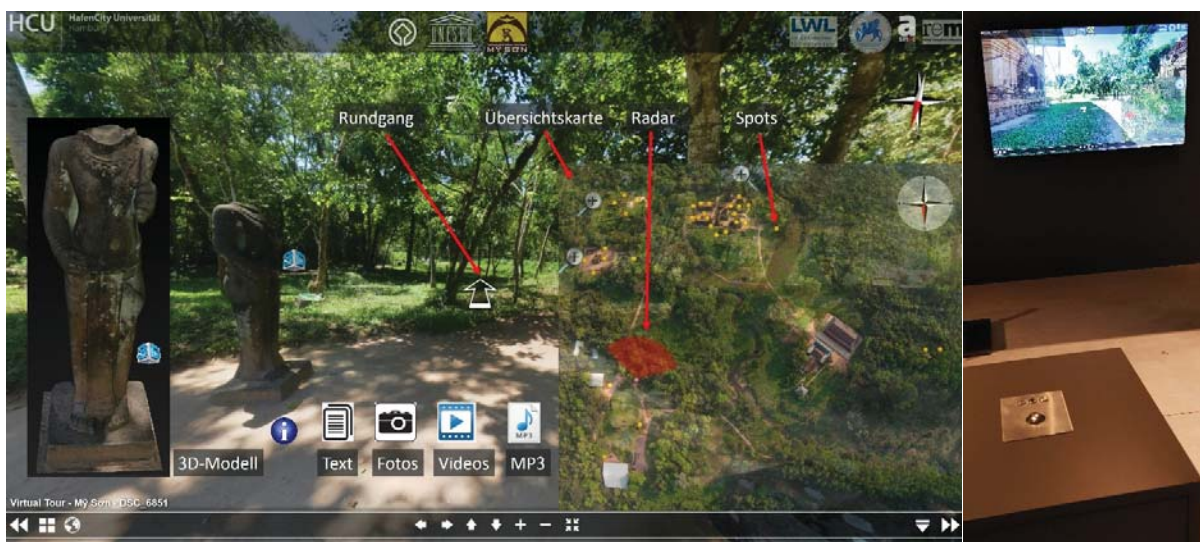


Abb. 9: Benutzeroberfläche der virtuellen Tour in Mỹ Sơn mit Erklärungen und potentiell integrierbaren Zusatzinformationen (links) und Betrachtung der virtuellen Tour in der Ausstellung (rechts)

Der interaktive Rundgang umfasst insgesamt 31 Panoramen (alle $360^\circ \times 180^\circ$) und zeigt von jedem Standpunkt ein voll-sphärisches Panorama, das dem Betrachter eine 360° Rundumsicht ähnlich wie bei Google Street View bietet. Der Betrachter sieht zum Einstieg des virtuellen Rundgangs ein Panorama und eine Luftbildübersicht der Weltkulturerbestätte (ein- und ausschaltbar), in der alle verfügbaren Standpunkte eingezeichnet sind (Abb. 9). Das Panorama kann mithilfe der Maus um 360° gedreht und hinein- und herausgezoomt werden. Durch Klick auf entsprechende Richtungspfeile im Panorama kann man dem Rundgang folgen oder man wechselt den Standpunkt durch Anklicken eines Spots in der Übersichtskarte. Ein Radar in der Übersichtskarte (Abb. 9) zeigt die aktuell gewählte Blickrichtung an und ein Kompass mit Nordpfeil liefert die wichtige Orientierung. Zusätzlich sind Hotspots z. B. als Verlinkung von Fotos, Detailaufnahmen und Zeichnungen von z. B. Fundstücken sowie Texte als Erklärungen oder Beschreibungen, Musik, Videos und andere Anwenderaktionen wie 3D-Objekte über eine XML-basierte Programmiersprache integriert. Mit dieser neuen Technologie können dem interessierten

Museumsbesucher nicht nur entfernte Orte wie Weltkulturerbestätten informativ und interaktiv näher gebracht werden, sondern auf diese Weise können auch archäologische Fundstücke in 3D mit dem originären Fundort verknüpft werden. Für die Ausstellung «Schätze der Archäologie Vietnams» im LWL-Museum für Archäologie in Herne wurden weitere virtuellen Touren von der Zitadelle Hoa Lu bei Trường Yên (Ninh Bình), der Zitadelle der Hồ-Dynastie in Thanh Hóa sowie vom Literaturtempel und der Zitadelle Thăng Long in Hanoi zusammengestellt (Abb. 10), die über einen großen Bildschirm an der Wand durch Steuerung mit einer Konsole betrachtet werden können (Abb. 9 rechts).

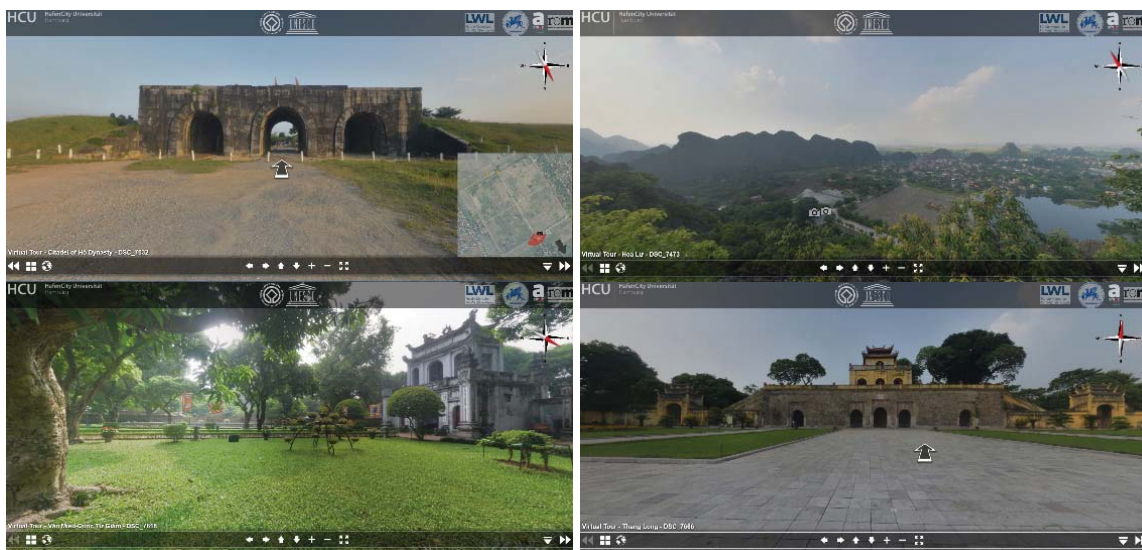


Abb. 10: Virtuelle Touren für die Weltkulturerbestätten in Nordvietnam - Zitadelle Hoa Lu bei Trường Yên (Ninh Bình) (oben links), Zitadelle der Hồ-Dynastie in Thanh Hóa (oben rechts) sowie für den Literaturtempel (unten links) und die Zitadelle Thăng Long (unten rechts) in Hanoi

6 Fazit & Ausblick

Die eingesetzten Methoden und Systeme haben sich im Sinne des Best Practice bei der Aufnahme und 3D-Modellierung vietnamesischer Kulturobjekte sehr bewährt, so dass die gewünschten Produkte in Form von gedruckten 3D-Replikaten und entsprechenden virtuellen Touren erstellt werden konnten. Eine hochauflösende digitale Spiegelreflexkamera mit der HDR-Funktion auf einem Stativ eingesetzt, bietet die entsprechend gute Ausstattung für Aufnahmen mit optimaler Beleuchtung und Tiefenschärfe für Projekte im Museumsbereich. Ergänzend dazu ermöglichen leistungsstarke tragbare Rechner (Notebooks) eine flexible und automatische Auswertung der erfassten Bildsequenzen auch schon vor Ort, was eine schnelle Qualitätskontrolle erlaubt. Durch die automatische Generierung von texturierten 3D-Modellen in unterschiedlichen Auflösungen können mehrere Anwendungen (z.B. Visualisierung, Internet, 3D-Druck) mit den entsprechenden Daten bedient werden. PhotoScan ist auch für Nicht-Photogrammeter leicht zu bedienen und es liefert in einem klar definierten Workflow automatisch Ergebnisse, allerdings ist bei der Beurteilung der Ergebnisse (Bildorientierungen, Kamerakalibrierung und Dreiecksvermaschung) Expertenwissen sinnvoll, wenn nicht sogar erforderlich.

Durch die Kombination von 3D-Modellen und virtuellen Rundgängen entstehen virtuelle Museen, die den Zugang zu entfernten Kulturen und Exponaten ermöglichen. Solche virtuelle Museen können als interaktive Lernplattformen aufgebaut werden, sodass sich interessierte Personen spielerisch Informationen erarbeiten können, um daraus einen entsprechenden Lerneffekt zu erhalten. Dabei bieten moderne Visualisierungswerkzeuge vom Smartphone und Tablet bis zu großen 3D-Leinwänden dem Betrachter eine interaktive und detaillierte Darstellung archäologischer Objekte und Exponate in 3D (TSCHIRSCHWITZ et al. 2014).

7 Literaturverzeichnis

- BARAZZETTI, L. REMONDINO, F. & SCAIONI, M., 2009: Combined use of photogrammetric and computer vision techniques for fully automated and accurate 3D modeling of terrestrial objects. *Proc. of SPIE Optics+Photonics*, Vol. 7447.
- GRÜN, A., REMONDINO, F. & ZHANG, L., 2002: Reconstruction of the great Buddha of Bamiyan, Afghanistan. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **34**(5), 363-368.
- KERSTEN, T. & LINDSTAEDT, M., 2012: Automatic 3D Object Reconstruction from Multiple Images for Architectural, Cultural Heritage and Archaeological Applications Using Open-Source Software and Web Services. *Photogrammetrie – Fernerkundung – Geoinformation*, Heft 6, 727-740.
- KERSTEN, T., MECHELKE, K., LINDSTAEDT, M., TSCHIRSCHWITZ, F., SCHREYER, K. & MAZIULL, L., 2014: Bildbasierte Low-Cost Systeme zur automatischen Generierung von 3D-Modellen archäologischer Fundstücke in Äthiopien und Katar. *Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage 2014*, Th. Luhmann/Ch. Müller (Hrsg.), Wichmann, VDE Verlag GmbH, Berlin und Offenbach, 210-222.
- REINECKE, A., 2016: Tempel, Inschriften und frühe Buddha-Figuren aus Holz von Gò Tháp. Frühgeschichte in Südvietnam – die Óc Eo-Kultur. *Schätze der Archäologie Vietnams*, Begleitband zur Sonderausstellung, Nünnerich-Asmus Verlag & Media, Mainz, 599 S.
- REINECKE, A. & MÜHLENBROCK, J., 2016: Schätze der Archäologie Vietnams – Die erste deutsche Vietnam-Ausstellung 2016/17. *Antike Welt – Zeitschrift für Archäologie und Kulturgeschichte*, Philipp von Zabern, WBG, Darmstadt, 5/2016, 8-11.
- REMONDINO, F., EL-HAKIM, S.F., GRUEN, A. & ZHANG, L., 2008: Turning images into 3-D models. *IEEE Signal Processing Magazine*, **25**(4), 55-65.
- TSCHIRSCHWITZ, F., KERSTEN, T. & ZOBEL, K., 2014: Interaktive Visualisierung und Steuerung von 3D-Modellen und Punktwolken mit Low-Cost Systemen. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, Band 23, Potsdam, Seyfert, E., Gülch, E., Heipke, C., Schiewe, J., Sester, M. (Hrsg.), Beitrag 173.
- WAHBEH, W. & NEBIKER, S., 2016: Kombination unterschiedlicher Photogrammetrischer Ansätze zur Rekonstruktion zerstörter Kulturdenkmäler. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, Band 25, T. Kersten (Hrsg.), 36. Wissenschaftlich-Technische Jahrestagung der DGPF, Dreiländertagung vom 7.-9. Juni 2016 in Bern (auf CD), 338-350.

Hyperspektrale und photogrammetrische Datenaufnahme am ehemaligen Zisterzienserkloster Wörschweiler

JENS BRAUNECK¹ & HERIBERT FELDHAUS²

Zusammenfassung: Für vergleichbare Restaurierungsmaßnahmen werden an einem überschaubaren Bauteil des Klosters Wörschweiler die Möglichkeiten der zwischenzeitlich neu entwickelten digitalen Aufnahmemethoden verglichen. Hintergrund ist darüber hinaus der Wunsch eines vollständigen Rekonstruktionsmodells des Klosters, das zwischen Bestand und Rekonstruktion unterscheiden soll und bei dem die vorhandene Bestandsaufnahme an ihre Grenzen stößt. Ausgewählt wurde die ca. 5 m aufrecht stehende Südwand des südlichen Seitenschiffs. An ihr sind seit dem 19. Jahrhundert zahlreiche mittelalterliche Grabsteine aus dem Kloster aufgestellt, die 2003 bis 2005 umfangreich restauriert wurden. Diese dienen als Experimentierfeld für hyperspektrale Untersuchungen im Nahbereich. Dabei stehen zunächst die Erfassung der aktuellen Situation sowie die Klassifizierung der verwendeten Baumaterialien aufgrund der Aufnahmen im Vordergrund.

1 Einleitung

Seit jeher beäugen historische Fachdisziplinen neue Techniken zunächst einmal mit Neugier, aber auch einer gewissen Skepsis. Das betrifft auch das Verhältnis zwischen Bauforscher und Vermessern resp. Vermessungsingenieuren. Die Mitte/Ende der 1980er Jahre aufkommenden berührungsfrei messenden Tachymeter wurden zunächst despektierlich als „moderne Zauberbesen“ bezeichnet, und doch haben sie die Bauaufnahmepraxis entscheidend verändert und bei richtigem Einsatz zu einer Qualitätssteigerung der Bauaufnahmen geführt. Besonders unversöhnlich schienen die Positionen angesichts der in den 1990er Jahren anwendungsreifen 3D-Scanner, die im Wesentlichen wohl auf einem gegenseitigen Missverstehen der Anforderungen bzw. Möglichkeiten und Grenzen beruhten. In Hinblick auf die Möglichkeiten der schnellen und vermeintlich umfassenden Datenerfassung durch 3D-Scanner forderte Manfred Schuller im Jahr 2000 „Mehr denken, statt nur messen“ (SCHULLER 2001). Wenig später bestand weitgehender Konsens darüber, dass sich auch die historische Bauforschung nicht den neuen Aufnahmeverfahren verschließen kann und sollte, wie zahlreiche in Kooperation ausgeführte Anwendungsbeispiele belegten. Die Versprechungen hinsichtlich der Qualitätssteigerung und Ökonomisierung müssen jedoch immer wieder einer kritischen Prüfung unterzogen werden. Die Frage ist nicht *ob*, sondern *wie* die Methoden in den Prozess der Bauforschung einbezogen werden können, wo sie Sinn machen und wo ihre Grenzen liegen.

Vor diesem Hintergrund werden aktuell Forschungen an der Klosterruine Wörschweiler (Abbildung 1) bei Homburg/Saar durchgeführt, die in den Jahren 2009-14 umfassend instand gesetzt wurde (LANDESDENKMALAMT SAARLAND 2015). Neue Aufnahmemöglichkeiten werden am konkreten Objekt erprobt und mit der bisherigen Praxis verglichen, um Erkenntnisse für

¹ TU Kaiserslautern, FB Bauingenieurwesen, Lehrgebiet Vermessungskunde und Geoinformation, Paul-Ehrlich-Str. 14, D-67663 Kaiserslautern, E-Mail: jens.brauneck@rhrk.uni-kl.de

² Büro für Bauaufmaß und Bauforschung, Auf Sprung 15, D-54292 Trier, E-Mail: info@bauaufmass.eu

vergleichbare Restaurierungsmaßnahmen zu erhalten. Dabei werden auch die möglichen Verknüpfungen mit der Bauforschung und Restaurierung betrachtet. Für den Methodenvergleich wurde zunächst als überschaubares Bauteil die ca. 5 m aufrecht stehende Südwand des südlichen Seitenschiffs ausgewählt. Sie weist zum einen zahlreiche Befunde des ehemals vorgelagerten Kreuzgangs auf und ist damit von besonderem Interesse für die Bauforschung. Zum anderen sind an ihr seit dem 19. Jahrhundert mittelalterliche Grabsteine aus dem Kloster aufgestellt, die 2003 bis 2005 umfangreich restauriert wurden. Diese dienen als Experimentierfeld für ebenfalls durchgeführte hyperspektrale Untersuchungen im Nahbereich. Hintergrund ist darüber hinaus die Anfrage eines Rekonstruktionsmodells der Gesamtanlage für eine museale Präsentation, das zwischen Bestand und Rekonstruktion unterscheiden soll und bei dem die vorhandene Bestandsaufnahme an ihre Grenzen stößt. Zu diesem Zweck werden mehrere Methoden miteinander kombiniert, um einen konsistenten und effizienten Arbeitsablauf zu entwickeln, der im finalen Schritt zur Erstellung des Modells per 3d-Druck führen soll.

Kloster Wörschweiler wurde im frühen 12. Jahrhundert zunächst als benediktinisches Priorat

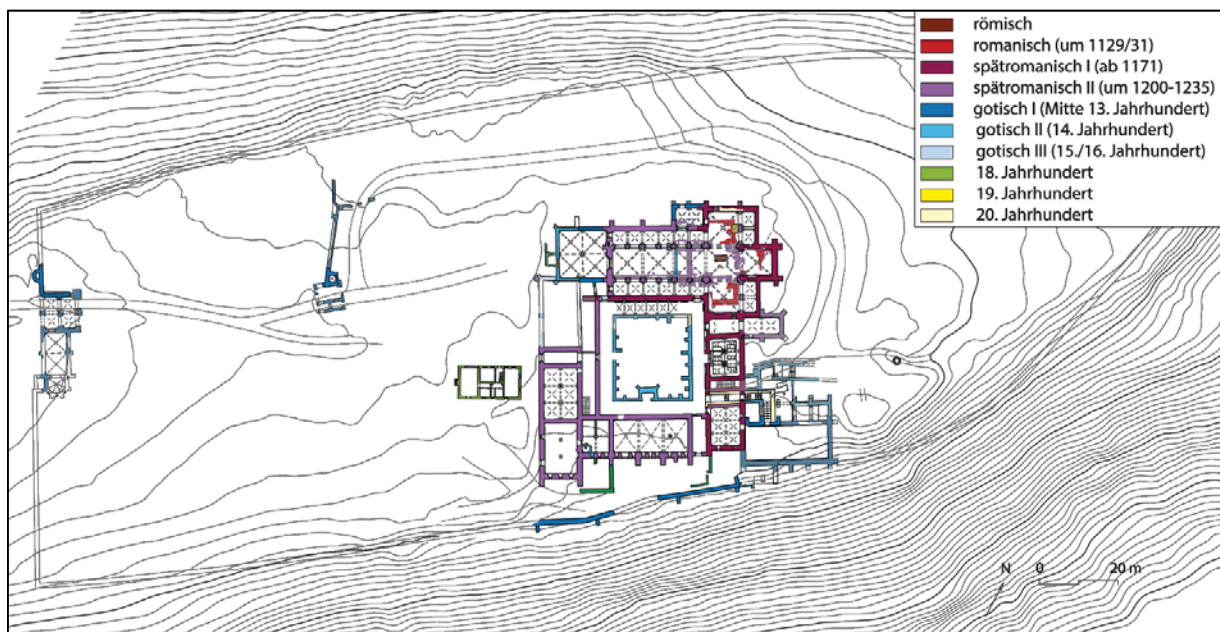


Abb. 1: Grundrissplan mit Bauphasen

gegründet und 1171 dem Orden der Zisterzienser übergeben. Die Mönche errichteten auf dem Bergsporn bis Mitte des 13. Jahrhunderts nach „Idealplan“ ein vollständig neues Kloster, das zum Zentrum des geistlichen, kulturellen und wirtschaftlichen Lebens der Region wurde. Im Zuge der Reformation wurde das Kloster 1558 aufgelöst und als Hofgut der Herzöge von Pfalz-Zweibrücken genutzt. 1614 fielen die Gebäude durch Unachtsamkeit weitgehend einem Brand zum Opfer und wurden nicht wieder aufgebaut. Ende des 19. Jahrhunderts (1889) erfolgte die erste Bestandsaufnahme der damals freiliegenden Bereiche im Maßstab 1:100. Erst durch umfangreiche archäologische Grabungen in den Jahren 1954-58 wurde der gesamte Grundriss freigelegt und ein Gesamtplan der Anlage erstellt. Nach zahlreichen punktuellen, nicht immer fachgerecht ausgeführten Sanierungsmaßnahmen, erfolgte in den Jahren 2009-2014 eine

Generalinstandsetzung. Die Maßnahmen umfassten neben statischen Sicherungen im Wesentlichen die Instandsetzung des Natursteinmauerwerks sowie die Konservierung der Putz- und Farbbefunde. Vorbereitend für die Maßnahme wurde in Kooperation des Landesdenkmalamts Saarland und der TU Kaiserslautern eine steingerechte Bestandsaufnahme der Ruine in Grundriss Ansichten und Schnitten angefertigt. Baubegleitend erfolgten archäologische Grabungen, Untersuchungen des Instituts für Steinkonservierung Mainz (IfS), restauratorische und bauhistorische Untersuchungen, d.h. es gibt eine Vielzahl an Referenzwerten für die aktuelle Forschung.

2 Material und Methoden

2.1 Photogrammetrische Aufnahmen

2.1.1 Bestandsaufnahme 2006-09

Die Aufnahme erfolgte durch angehende Architekten im Fach Denkmalpflege in Kooperation mit dem Lehrgebiet Vermessung/Kartographie der TU Kaiserslautern, vertreten durch seinen damaligen Leiter, Herrn Dr.-Ing. Klaus Trumpke. Wunsch der Landesdenkmalpflege war die steingerechte Erfassung sämtlicher Maueraufsichten sowie wesentlicher Wandansichten im Maßstab 1:50 in digitaler Form. Die gewählte Aufnahmemethode spiegelt sowohl die technischen Möglichkeiten der Zeit als auch die damalige technische Ausstattung der Universität wieder und wird daher kurz vorgestellt: Unabdingbare Voraussetzung bildete ein dauerhaft vermarktes Festpunktfeld im Ruinengelände, auf das sich die über vier Jahre verteilten Messkampagnen sowie spätere Folgemessungen beziehen konnten und das die Maßhaltigkeit der Einzelmessungen garantierte. Aus didaktischen Gründen wurden zur Aufnahme der Maueraufsichten unterschiedliche Methoden angewandt: Vom Feldpantographen, der nach dem Prinzip des Storchenschnabels unmittelbar eine maßstäbliche (1:20), jedoch stark bearbeitungsbedürftige Zeichnung liefert, über die „klassische“ Dreiecksmessung im Handaufmaß auf Karton bis hin zu Fotoentzerrungen einzelner Mauerabschnitte (Programm PhotoPlan der Firma kubit). Die zeichnerisch besten Ergebnisse wurden mit der händischen Verdichtung eines maßstäblich ausgedruckten, zuvor tachymetrisch eingemessenen Festpunktnetzes (markante Steinkanten) erzielt, da hier mit wenigen, mit dem Zollstock abgegriffenen Maßen und gesundem Augenmaß sehr schnell ausdifferenzierte Bleistiftzeichnungen möglich sind, deren Aussage bei entsprechender Erfahrung und Übung über ein reines Konturenaufmaß hinausgeht. Sehr aufwändig war jedoch die abschließend erforderliche Digitalisierung der Zeichnungen, die zudem eine weitere Interpretation bzw. Fehlerquelle darstellte.

Die Aufnahme der aufrecht stehenden Mauerscheiben erfolgte im Fall der weitgehend in einer Bildebene liegenden Seitenschiffwände mit Fotoentzerrungen, im Bereich der Grabplatten in zwei Entzerrungsebenen. Auf Grund der Tiefenstaffelung von Kirchenwestwand und Querhauswand wurden diese stereoskopisch mit einem analogen System (Zeiss Planicomp P3 und Rollei metric 6006, Format 6x6) erfasst. Für das Anbringen der Passpunkte und das Fotografieren der jeweiligen Bildpaare stand für einen Tag eine Arbeitsbühne zur Verfügung. Die Auswertung erfolgte unter CAD durch die Studierenden an der Universität in Kaiserslautern und erforderte – wie bei der Stereoskopie üblich – eine Kontrolle vor Ort und Verdichtung verschatteter Bildbereiche. Im

unmittelbaren Umfeld der Ruine wurde durch Studierende der Fachrichtung Bauingenieurwesen eine Geländeaufnahme durchgeführt. Das Landesamt für Vermessung, Geoinformation und Landentwicklung (LVermGeo Saarland) stellte für den Bergrücken die aus einer landesweiten Befliegung gewonnenen Höhenlinien zur Verfügung. Als Ergebnis lag vor Beginn der Instandsetzungsarbeiten eine vollständige, zweidimensionale Bestandsaufnahme der Ruine vor, die – angesichts der unterschiedlichen Bearbeiter nicht anders zu erwarten – in einigen Bereichen überarbeitungsbedürftig war. Überarbeitung und Ergänzungen erfolgten im Rahmen der Bauforschung.

2.1.2 Bestandsaufnahme 2016

Die aktuellen Aufnahmen erfolgten im Rahmen eines mehrtägigen Geländepraktikums von Masterstudierenden der TU Kaiserslautern/FB Bauingenieurwesen im Jahr 2016, wobei sämtliche Daten der Klosteranlage innerhalb eines Tages aufgenommen wurden.

Als Referenzmessung wurde zunächst ein Laserscan der kompletten Wand mit einer scanfähigen Leica Multistation (MS60) durchgeführt. Diese relativ neue Geräteklasse stellt die Verbindung von Tachymeter und terrestrischen Laserscannern (TLS) dar. Die Scangeschwindigkeit liegt bei vergleichsweise wenigen 1000 Punkten/Sekunde, dafür erhalten sämtliche Punkte bei zuvor erfolgter Stationierung des Geräts direkt bei der Aufnahme Koordinaten im Landeskoordinatensystem. Damit entfällt die nachträgliche Bearbeitung des sonst notwendigen meshings. Darüber hinaus kann im Feld auch nach einem Positionswechsel und erneuter Stationierung direkt erkannt werden, welche Bereiche bereits gescannt wurden. Dadurch werden weniger Flächen mehrfach erfasst, was den Arbeitsablauf zusätzlich optimiert. Zur Referenzierung der Aufnahmen wurden zusätzlich 13 digitale Markierungen (12 bit Codierung, auf festem Untergrund verklebt) an der Wand angebracht und mit der Multistation aufgemessen.

Die photogrammetrische Datenaufnahme erfolgte mit drei unterschiedlichen aber weit verbreiteten Kamerasystemen, die zwar jeweils eine eigene Geräteklasse repräsentieren, allerdings alle in einer ähnlichen Preiskategorie angesiedelt sind:

- Olympus OM-D-M1 Systemkamera mit 17 mm Festbrennweitenobjektiv
- iPhone5s Smartphone mit 4,15 mm Brennweite
- DJI Phantom 3 Professional UAV-Modell (Einstiegssegment)

Die Punktwolkenmodelle wurden per Structure-From-Motion-Verfahren in AgiSoft Photoscan erstellt und anhand der Marker auf DHDN-GKZ2 referenziert. Für einen direkten Vergleich der Punktdaten von iPhone und Olympuskamera wurden die Aufnahmen auf jeweils 100 Stück reduziert. Die Prozessierung der Daten bis zur dichten Punktwolke erfolgte per Stapelverarbeitung in Photoscan mit identischen Parametern. Die Punktwolken wurden anhand der Marker georeferenziert, in das E57-Format exportiert und in die Software CloudCompare geladen. Dabei erfolgte eine Reduzierung der Koordinatenwerte auf vier Vorkommastellen.

Da die Datenaufnahme auf die Erfassung eines flächendeckenden Modells ausgerichtet war, wurde auch eine unbemannte Flugplattform (engl. Unmanned Aerial Vehivles/UAV) aus dem Einstiegssegment eingesetzt (DJI Phantom 3 Professional). Der vermehrte Einsatz dieser Systeme hat in den letzten Jahren immer wieder das große Potential dieser Systeme aufgezeigt, insbesondere wenn es um einen Skalenmaßstab geht, der zwischen der Vermessung im Nahbereich und der satelliten- oder flugzeuggestützten Datenaufnahme liegt (NEX & REMONDINO 2014). Die

Fehleranfälligkeit bei der Modellierung aufgrund der Sensorgröße und Objektivqualität verursacht bei den kostengünstigen UAV-Modellen Abweichungen, die über die Qualität der vorher erwähnten Kameramodelle hinausgehen. Hinsichtlich des angestrebten Modellmaßstabs von 1:200 sind diese Abweichungen aber zu vernachlässigen. Darüber hinaus eignen sich diese Systeme für flächendeckende Luftbildaufnahmen und erlauben durch die direkte Steuerung des Gimbals weitgehend freie Wahl der Aufnahmegeometrie.

Das so erstellte Modell des Klosterbergs wurde zur besseren Vergleichbarkeit auf die Fläche der Klosterruine reduziert. Innerhalb von Agisoft Photoscan erfolgte zunächst eine Koordinatentransformation von WGS84 zu DHDN GK Zone 2 anhand von 10 markanten Punkten. Nach einem Export der Punktwolke in das E57-Format wurde zusätzlich in CloudCompare eine Registrierung der UAV-Daten auf den Laserscan über identische Punktpaare durchgeführt, um die Lagegenauigkeit des per Drohne erstellten Modells noch zusätzlich zu erhöhen.

2.2 Hyperspektrale Aufnahmen

Die Kombination eines photogrammetrischen mit einem multi- oder hyperspektralen Aufnahmesystems ermöglicht es, zusätzlich zu den räumlichen Koordinaten Informationen über die Oberflächenbeschaffenheit und das erfasste Material zu gewinnen. Mittels dieser zusätzlichen spektralen Signatur können unterschiedliche Objekteigenschaften abgeleitet werden, die über die rein visuelle Interpretation hinausgehen. Dieser Ansatz wird seit Jahrzehnten in der satellitengestützten Fernerkundung (z.B. LANDSAT-TM) verfolgt, um berührungslos Informationen zur Oberflächenbeschaffenheit und das erfasste Material zu gewinnen. Somit können die Klassifizierungsansätze aus diesem Bereich angepasst und genutzt werden. Eine neu entwickelte Geräteklasse bildgebender hyperspektraler Kameras ermöglicht die Aufnahme entsprechender Daten im Nahbereich, ähnlich einer konventionellen Digitalkamera. Die leichte und kompakte Bauweise ermöglicht sowohl den Betrieb auf Stativen als auch an Bord von Multikopterplattformen. Mittels der pro Pixel registrierten spektralen Signatur können unterschiedliche Objekteigenschaften abgeleitet werden, die über die rein visuelle Interpretation hinausgehen. Der Spektralbereich der eingesetzten Kamera Cubert UHD 185/UAV reicht von 450 bis 950 nm und deckt somit das sichtbare Licht und das nahe Infrarot ab. Dieser Ausschnitt des Spektrums eignet sich primär für die Erfassung und Analyse photosynthetisch aktiver Vegetation (ZARCO-TEJADA et al. 2009), allerdings weisen auch Baumaterialien und Werkstoffe nennenswerte Unterschiede im Verlauf der spektralen Signatur auf (SHABAN 2013). Insbesondere bei Untersuchungen von Beton wurde das Potential dieses Verfahrens zur Detektion von Schäden durch Carbonatisierung, biogenen Bewuchs und Aussalzung an der Materialoberfläche nachgewiesen (ARITA et al. 2001).

Es wird daher im Rahmen von Feld- und Laboruntersuchen geprüft, unter welchen Bedingungen verlässliche Aussagen aufgrund hyperspektraler Objekteigenschaften erreicht werden können. Da dieses Verfahren auf eine direkte Beleuchtungsquelle angewiesen ist, wurde die Südwand mit direkter Sonneneinstrahlung untersucht. Die Kamera wurde auf einen Weißstandardreflektor (50 %) sowie auf eine Schwarzreferenz kalibriert.

3 Ergebnisse

3.1 Photogrammetrie

In Abbildung 2 A & B sind die Ergebnisse der photogrammetrisch erstellten Modelle der Kamera und des Smartphones in Bezug zum Laserscan aufgeführt. Zur Überraschung sämtlicher Kursteilnehmer zeigte das iPhone-Modell geringfügig bessere Ergebnisse als das Olympusmodell, wobei die Abweichungen beider Modelle in aller Regel deutlich unter 2 cm liegen.

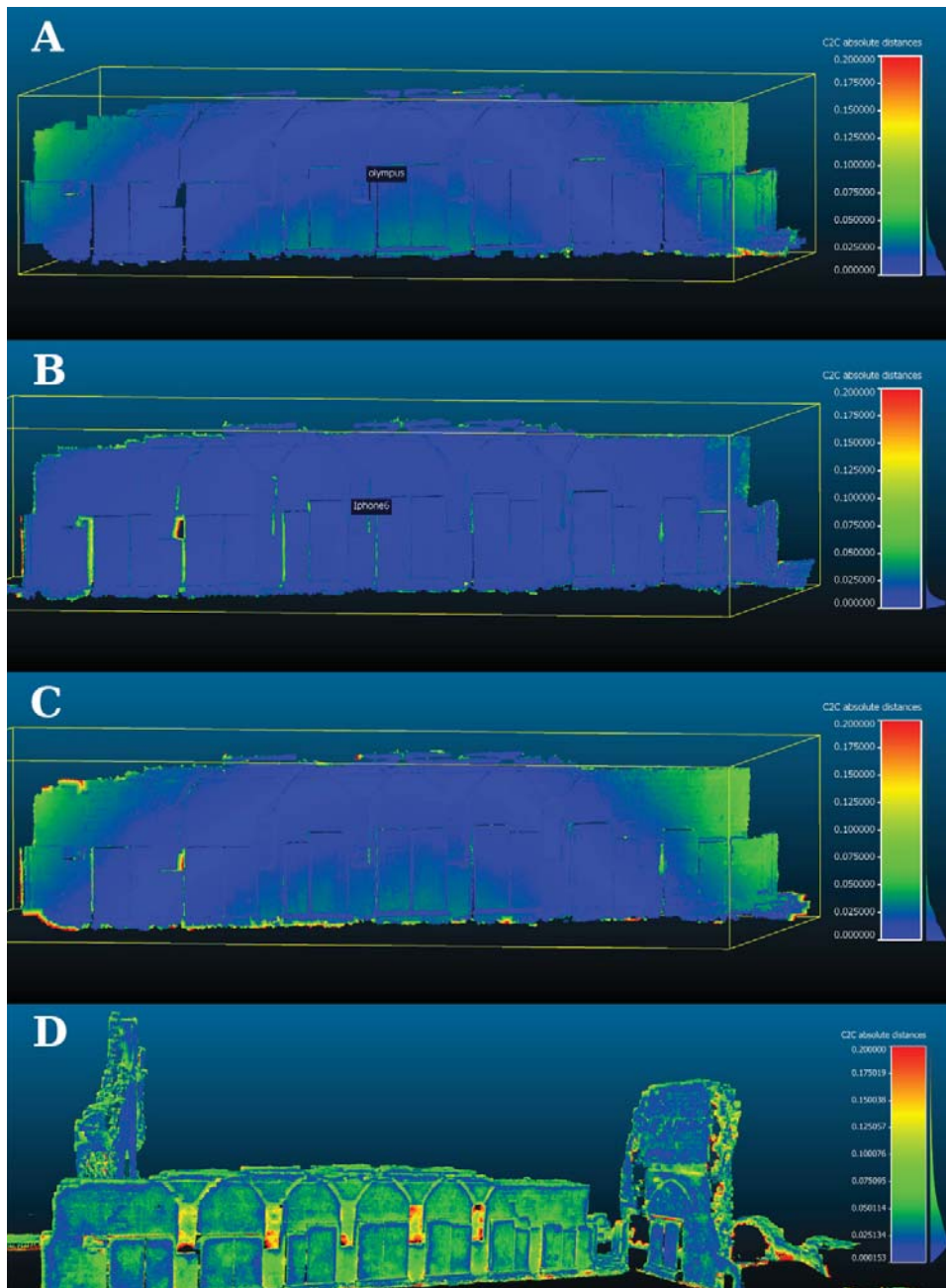


Abb. 2: Vergleich der Punktwolkenmodelle in CloudCompare (A Laserscan/Olympus, B Laserscan iPhone, C Olympus/iPhone, D Laserscan erweitert/Dji Phantom 3)

Abbildung 2C zeigt den direkten Vergleich der photogrammetrisch erstellten Modelle miteinander. Die Olympuskamera zeigt einen deutlichen radialen Versatz resp. ein im Randbereich kippendes Modell sowohl mit als auch ohne Kalibrierung der Kamera. Das Smartphone zeigt hingegen eine geringfügige Abweichung, die sich auf die gesamte Fläche verteilt.

Insgesamt können die Abweichungen aber als gering bezeichnet werden. Dies ist insbesondere für die Praxis relevant, da eine einfache und insbesondere schnelle Schadensdokumentation mit zusätzlichen Kalibrierungswerkzeugen und/oder Passpunkten auch mithilfe eines Smartphones durchgeführt werden kann.

Abbildung 2D zeigt im Detail die Abweichung des UAV-Punktwolkenmodells vom Laserscan auch in Bezug zu anderen Mauerteilen. Dabei zeigt sich wie erwartet eine insgesamt größere Abweichung als bei den zuvor aufgeführten Kameramodellen. Insbesondere sind kleinere Details weniger gut dargestellt, flächige Elemente zeigen Fehler, die unterhalb von 10 cm liegen.

3.2 Hyperspektrale Daten

Aufgrund der unvollständigen Datengrundlage, insb. vergleichbarer spektraler Signaturen, können die bisher durchgeführten hyperspektralen Aufnahmen nicht direkt zur Bestimmung der im Bild erfassten Materialien dienen. Allerdings zeigen die in Abbildung 4 exemplarisch dargestellten Spektren das Potential dieses Verfahrens. Der Kurvenverlauf der photosynthetisch aktiven Vegetation erscheint nahezu ideal. Sowohl der typische Knick im red edge Bereich als auch das Plateau im nahen Infrarot sind auch mit einer eher ungünstigen Aufnahmegeometrie und ohne weitere radiometrische Korrektur deutlich zu erkennen.

Die Differenzierung des angewitterten (historischen) Sandsteins von den restaurierten Bereichen ist im RGB-Bild deutlich auszumachen und dementsprechend weist auch die spektrale Signatur deutliche Unterschiede auf.



Abb. 3: Hyperspektrale Aufnahme in RGB-Darstellung mit punktuellen Abfragen (links) der Spektren ausgewählter Oberflächen (rechts)

4 Möglichkeiten für die historische Bauforschung

Was bedeutet dies für die Praxis des Bauforschers? Die historische Bauforschung – historisch in Abgrenzung zur ingenieurwissenschaftlichen Materialprüfung – versucht, unter Einbeziehung von Archivalien möglichst umfassend die Baugeschichte eines Gebäudes, seine „Entstehungs- und Lebensgeschichte“ zu rekonstruieren. Sie gleicht einer kriminalistischen Untersuchung, bei der das Bauwerk selbst die Quelle ist. Grundlage bildet seit jeher die maßstäbliche Bauaufnahme des Gebäudes. Weit über die reine wirklichkeitsgetreue Darstellung des Gebäudes hinausgehend ist sie zugleich Erkenntnisprozess: Es geht um das Verständnis räumlicher Gefüge und Zusammenhänge, das Nachvollziehen der Bau- und Tragwerksstruktur, das Erkennen von Fugen und Unstimmigkeiten, die Unterscheidung von Steinmetzzeichen und Steinbeschädigungen etc. Bauforschung ist zwingend Feldforschung am Objekt. Da jede Bauaufnahme eine Interpretation des Bestands darstellt, muss der Bearbeiter über entsprechende bauhistorische Kenntnisse verfügen. Ein weiteres Problem: Ansprüche an Genauigkeit und Detaillierung werden je nach Befunden erst im Laufe der Bearbeitung deutlich. Vor diesem Hintergrund werden die Möglichkeiten und Grenzen digitaler, berührungsfreier Aufnahmemethoden deutlich, die nicht zuletzt auch eine Frage der Angemessenheit ist. In einem kleinteiligen, möblierten Bauernhaus macht ein herkömmlicher 3d-Scan nur wenig Sinn. Die Nachbearbeitung übersteigt bei Weitem den Zeitgewinn vor Ort, zumal der Scan zwar ein detailgenaues Abbild des Gebäudes bietet, aber nicht „hinter“ die Oberfläche, in die Konstruktion schaut.

Bei großen Objekten können die „neuen“ Aufnahmemöglichkeiten hingegen eine verlässliche Grundlage liefern, mit der die unterschiedlichen Fachdisziplinen (Architekt, Tragwerksplaner, Bauforscher, Restaurator, didaktische/museale Aufarbeitung) weiter arbeiten. Der Vergleich in Wörschweiler zwischen der „traditionellen“ Bestandsaufnahme 2006-09 und der „neuen“ Aufnahme 2016 zeigt eindrücklich, mit welchem geringem Aufwand eine erste Gesamterfassung der Ruine möglich ist. Die dreidimensionale Betrachtung eines Bauwerks und seiner Entstehung kann neue räumliche Zusammenhänge geben. Die Möglichkeiten der Veranschaulichung nicht nur für den Laien sind unbestritten, wenngleich die Unterscheidung zwischen Bestand und Rekonstruktion erkennbar bleiben muss, gerade für den Laien. Grundlegend ist, dass Ziele und Anforderungen einer Bestandsaufnahme zuvor geklärt werden. Nicht alles was möglich ist, macht Sinn. Ebenso wichtig sind die Fragen des Datenaustauschs, der technischen Möglichkeiten der Datenbearbeitung und nicht zuletzt der langfristigen Speicherung.

Die ersten Beispiele der Hyperspektralaufnahmen zeigen, dass sich die Möglichkeiten der berührungsfreien Untersuchungen erweitern könnten. Bei Schadenskartierungen ist an die Differenzierung biogenen Bewuchses und das Erkennen von nachträglichen Reparaturmaßnahmen, etwa eine Unterscheidung von Zement- zu historischen Kalkmörteln zu denken. Wie weit die Ausdifferenzierung möglich ist, etwa die Unterscheidung unterschiedlicher Sandsteinvarietäten innerhalb eines Mauerabschnitts, müssen weitere Versuche zeigen. Die Ergebnisse wären ein weiteres Puzzleteil in der Gesamtbetrachtung des Bauwerks.

5 Fazit & Ausblick

Während seit dem 19. Jahrhundert Bauaufnahme und Bauforschung im wahrsten Wortsinn in einer Hand lagen, hat die zunehmende Verlagerung auf digitale Aufnahmesysteme seit spätestens den 1990er Jahren zwangsläufig zu einer personellen Trennung der Disziplinen geführt. Die Bestandsaufnahme, die unweigerlich eine Interpretation des Bestands darstellt, ist immer häufiger nicht mehr Teil des Erkenntnisprozesses. Berührungsfreie Aufnahmemethoden eröffnen zweifellos neue Möglichkeiten, die jedoch objektbezogen sinnvoll eingesetzt werden müssen. Sie können die Arbeit des Bauhistorikers nicht ersetzen, sondern seine Arbeit unterstützen. Voraussetzung ist ein gegenseitiges Verstehen der technischen Machbarkeiten und Anforderungen. Auf längere Sicht wird sich wohl auch die Praxis der Bauforschung verändern, so wie sich bereits seit längerem die Planungspraxis in der Architektur vom jahrhundertlang praktizierten 2D zum 3D ändert. Mit keiner Technik ersetzbar ist jedoch die Untersuchung des Bauwerks mit allen Sinnen vor Ort.

Natürlich wird die Wahl des Aufnahmesystems, neben der verfügbaren apparativen Ausstattung, in Hinblick auf die erforderliche Genauigkeit erfolgen müssen. Grundsätzlich zeigen die hier getesteten photogrammetrischen Aufnahmesysteme mitunter Schwächen, die allerdings durch die Datenintegration aus unterschiedlichen Quellen beseitigt werden können. Kombiniert man beispielsweise die Punktwolken eines hochgenauen aber eher stationären Scanners mit den Ergebnissen eines mobilen photogrammetrischen Aufnahmesystems, erhält man in kurzer Zeit ein flächendeckendes 3d-Modell, dessen Detailgrad man selbst bestimmen kann. Ein großer Vorteil der so erzeugten digitalen Daten ist die Möglichkeit, auf diese Modelle ebenfalls digitale Rekonstruktionen anzupassen. Als Endprodukt aus diesen Arbeitsschritten soll ein per 3d-Druck erstelltes Modell entstehen, welches die aktuelle und die rekonstruierte historische Situation des Klosters differenziert darstellen soll.

Die erste Auswertung der Hyperspektralaufnahmen zeigt das Potential dieses Verfahrens. Da bei Aufnahmen im Nahbereich der Anteil der Mischpixel deutlich reduziert werden kann, sollte somit unter günstigen Umständen die Unterscheidung von Materialien im VIS-NIR-Bereich möglich sein. Dazu werden aktuell an der TU Kaiserslautern Laboruntersuchungen an unterschiedlichen Werkstoffen durchgeführt. Eine Erweiterung des Spektrums, beispielsweise über ein weiteres hyperspektrales Aufnahmesystem im SWIR (Kurzwelleninfrarot) bzw. IR-B (Nahinfrarot 1450 bis 3000 nm), dürfte die Wahrscheinlichkeit zur eindeutigen Bestimmung der Materialien verbessern. Chancen dieser Methode im Themenfeld Bauingenieurwesen und Architektur werden insbesondere im Monitoring von Bauwerken, in der Aufnahme von Bestandsdaten und deren Verknüpfung mit Datenbanken gesehen (z.B. BIM). Langfristig soll das Potential hyperspektraler Aufnahmen für die Schadensdokumentation und Veränderungsanalyse genauer überprüft werden. Ein großer Vorteil dieser Methodenkombination liegt in der objektiven Datenerfassung und Dokumentation sowohl der Objektgeometrien als auch der Oberflächeninformation.

6 Literaturverzeichnis

- ARITA, J., SASAKI, K. I., ENDO, T. & YASUOKA, Y., 2001: Assessment of concrete degradation with hyper-spectral remote sensing. 22nd Asian Conference on Remote Sensing **5**, 5 p.
- LANDESDENKMALAMT SAARLAND, 2015: Baudenkmalpflege 2005-2014 (Denkmalpflege im Saarland, Bd. **5**), Saarbrücken.
- NEX, F. & REMONDINO, F., 2014: UAV for 3D mapping applications: a review. *Applied Geomatics* **6**(1), 1-15.
- SCHULLER, M., 2001: Mehr Denken statt nur Messen. Von Handaufmaß bis High Tech – Aufnahmeverfahren in der historischen Bauforschung (Kolloquium TU Cottbus 2000), Weferling, U., Heine, K. & Wulf, U. (Hrsg.), Verlag Philipp von Zabern, 213-226.
- SHABAN, A., 2013: Determination of Concrete Properties Using Hyperspectral Imaging Technology: A Review. *Science Journal of Civil Engineering and Architecture*, Volume **2013**, Article ID sjcea-102, 11 p.
- ZARCO-TEJADA, P.J., BERNI, J.A. J., SUÁREZ, L., SEPULCRE-CANTÓ, G., MORALES, F. & MILLER, J.R., 2009: Imaging chlorophyll fluorescence with an airborne narrow-band multispectral camera for vegetation stress detection. *Remote Sensing of Environment* **113**(6), 1262-1275.

Erfassung und Analyse dreidimensionaler Deformationen an Gemälden zur Unterstützung des Restaurierungsprozesses

FERDINAND MAIWALD¹ & DANILO SCHNEIDER¹

Zusammenfassung: Der Beitrag befasst sich mit der photogrammetrischen Erfassung und Analyse dreidimensionaler Deformationen an Gemälden. Im Zuge einer Restaurierung durch die Staatlichen Kunstsammlungen Dresdens wurde in drei Epochen ein historisches Gemälde von der Firma fokus GmbH Leipzig aus verschiedenen Positionen fotografiert. Mittels Structure-from-Motion wurden aus diesen Aufnahmen Punktwolken erzeugt und epochenweise miteinander verglichen. So konnte eine maximale Wölbung von bis zu 6,3 cm festgestellt werden aber auch kleinere Deformationen und Risse konnten gut sichtbar gemacht werden. Verschiedene Visualisierungen können eine hilfreiche Ergänzung zu klassischen Schadenskartierungen sein. Insgesamt wurde nachgewiesen, dass photogrammetrische Methoden sehr gut zur Überwachung eines Restaurierungsprozesses geeignet sind.

1 Einleitung

Dieser Beitrag befasst sich mit der photogrammetrischen Erfassung und Analyse dreidimensionaler Deformationen an Gemälden. Das untersuchte Gemälde von Lukas Cranach dem Älteren aus dem Jahr 1506 wurde im Zuge der Restaurierung, durchgeführt von den Staatlichen Kunstsammlungen Dresden, von der Firma fokus GmbH Leipzig mehrfach mit einer Nikon D800 aufgenommen. Das Gemälde weist einige Besonderheiten auf, die für die Analyse der photogrammetrischen Ergebnisse relevant sind. Bei dem Bild handelt es sich um die Innenseite des linken Flügels des Katharinenaltars von Lukas Cranach dem Älteren und es zeigt drei Heilige. Dieser Flügel war ursprünglich beidseitig bemalt, wurde jedoch wahrscheinlich im Zuge einer Restaurierung im 18. Jahrhundert in der Mitte vertikal auseinandergesägt (MARX 1996). Die daraus resultierende Stärke der Holztafel beträgt aus diesem Grund nur noch 2,6 – 4,0 mm. Diese geringe Dicke erschwert zum einen die Restaurierung und lässt zum anderen schneller temperatur- sowie luftdruckbedingte Deformationen zu.

Die Ölfarbe des Gemäldes ist direkt auf sieben miteinander verklebten Lindenholtzbrettern aufgebracht, die mit Hilfe eines Stützsystems auf der Rückseite (die sogenannte Parkettierung) stabilisiert wurden. Dieses Stützsystem aus Quer- und Längsleisten (Flachparkett) sollte ursprünglich eine Bewegung der Holztafel vermeiden (NICOLAUS 2003). Durch die feste Gitterstruktur dieser Parkettierung entstanden allerdings Spannungen in dem Gemälde und infolgedessen Risse in der Holztafel. Deshalb sollte das Gemälde von den Staatlichen Kunstsammlungen Dresden erneut restauriert werden. Ziel der Arbeit war eine dreidimensionale photogrammetrische Rekonstruktion des Gemäldes sowie eine Abschätzung der Größenordnung von Verformungen und die Beurteilung der Deformationen in verschiedenen Epochen. Alle drei Epochen wurden mehrfach in Form eines konvergenten Bildverbandes fotografiert:

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

- In Epoche 1 war auf der Rückseite der Holztafel noch das alte nahezu starre Stützsystem aufgebracht, welches Spannungen auf das Gemälde ausübte, die sich in Rissen und kleinen Wellen zeigten.
- In Epoche 2 wurde die Holztafel ohne Stützsystem in liegender Position aufgenommen (Abb. 1). In diesem Zustand konnte sich das Gemälde ohne mechanischen Zwang deformieren, was zu einer großflächigen Wölbung führte.
- In Epoche 3 wurde durch die Restauratoren ein neues Stützsystem auf der Rückseite angebracht. Dieses soll zum einen eine Wölbung des Gemäldes minimieren und zum anderen starke Spannungen in der Holztafel und damit die Bildung von Rissen vermeiden. Es soll damit einen Kompromiss zwischen den Zuständen aus Epoche 1 (starres Stützsystem) und aus Epoche 2 (kein Stützsystem) darstellen.

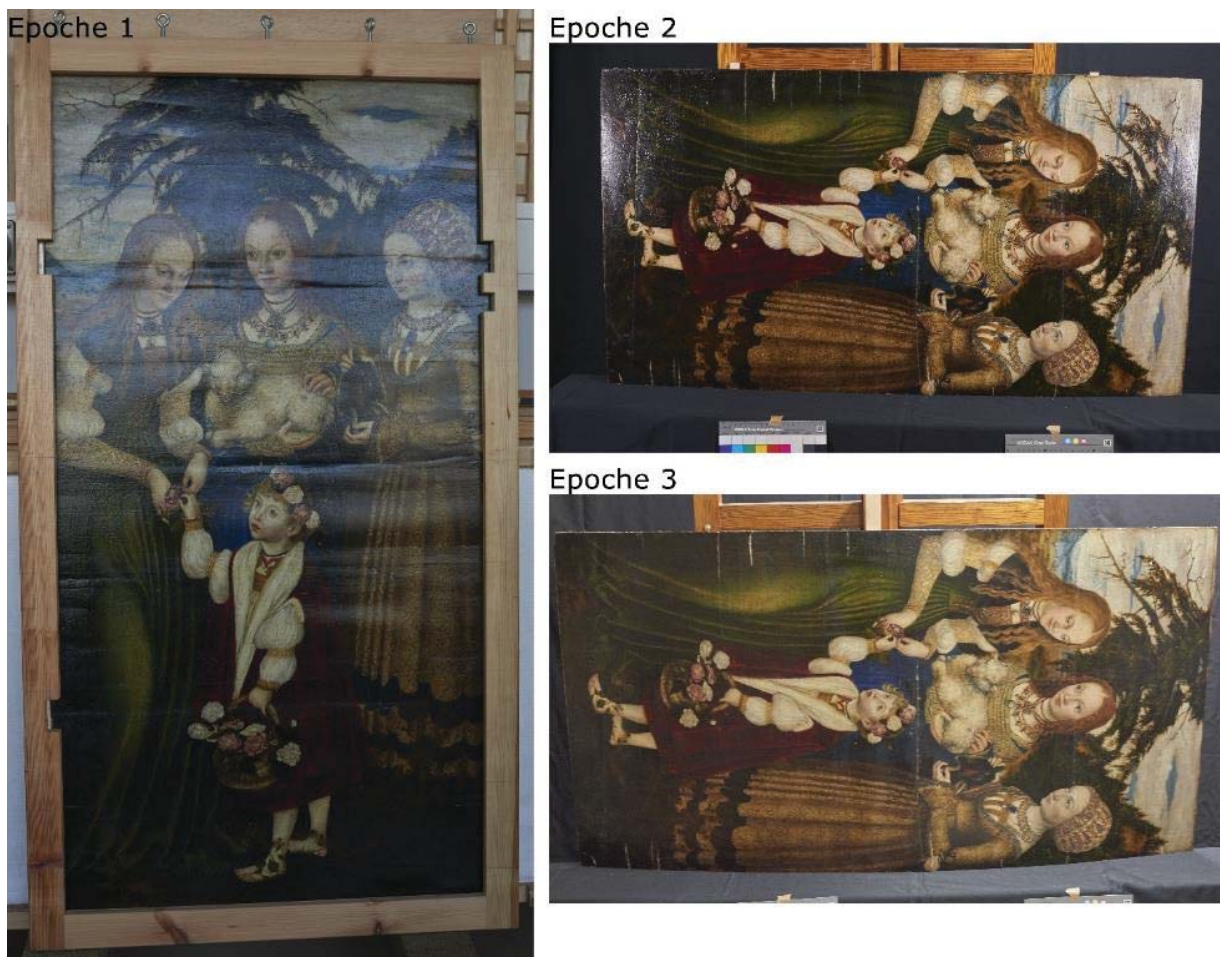


Abb. 1: Lage des Gemäldes zum Aufnahmezeitpunkt in den verschiedenen Epochen

Die dreidimensionale Aufnahme des Gemäldes sollte berührungslos erfolgen. Deshalb kamen ausschließlich optische (bildbasierte und scannende) Verfahren in Frage (LUHMANN 2010). An einem Testobjekt mit ähnlichen Eigenschaften wie beim Cranach-Gemälde wurden deshalb die Methoden Structure-from-Motion (SfM) und Streifenlichtscanning miteinander verglichen. Am Cranach-Gemälde kam nur SfM zum Einsatz. Insbesondere sollte geprüft werden, ob trotz der teilweise sehr kontrastarmen Textur ein zuverlässiges Bildmatching durchgeführt werden kann (LOWE 2004).

2 Datenakquisition

In den verschiedenen Epochen wurden Aufnahmen von der fokus GmbH Leipzig gemacht. In der ersten Epoche wurden 23 Fotos, in der zweiten 22 Fotos und in dritten Epoche 26 Fotos des Gemäldes mit einer Nikon D800 unter Verwendung eines Polarisationsfilters aufgenommen. Der Polarisationsfilter diente der Minimierung von Reflexionen, welche durch die Ölfarbe hervorgerufen wird. Nach der SfM-Prozessierung stellte sich heraus, dass bei der Bildaufnahme im Hinblick auf eine bessere Qualität der Ergebnisse noch Optimierungspotenzial bestand. Insbesondere sind eine größere Anzahl von Fotos und eine bessere Aufnahmekonfiguration, insbesondere größere Schnittwinkel homologer Bildstrahlen im Objektraum wünschenswert, sofern die Genauigkeit und Zuverlässigkeit der Bildpunktzuordnung nicht darunter leidet (SOOT et al. 2015). Durch gute Konvergenzwinkel der Aufnahmeachsen werden systematische Abweichungen aufgrund von ungenauen Schätzungen der inneren Orientierung der Kamera vermieden. Diese Abweichungen können sich beispielsweise als sog. Dome-Effekt äußern, bei dem sich in der Mitte der entstandenen Punktwolke eine systematische Wölbung bildet. Diese Verformung entsteht vor allem auf Grund der ungenauen Schätzung der radial-symmetrischen Verzerrung, kann aber durch konvergente Aufnahmen vermieden werden (WACKROW et al. 2008).

Für die Erstellung der Punktwolken mittels Structure-from-Motion lieferte die Software Agisoft Photoscan für das Testobjekt sowie für das Cranach-Gemälde die vollständigsten Ergebnisse. Softwarelösungen wie VisualSFM, SURE und Apero (Tab. 1) konnten keine lückenlose Punktwolke des Gemäldes erstellen.

Tab. 1: Auflistung der verwendeten Programme zur Berechnung der Punktwolke für das Testobjekt

Programm	Algorithmus	Lizenz	Webseite
Agisoft Photoscan	Firmeneigen	Kostenpflichtig	http://www.agisoft.com/
VisualSFM	PMVS/CMVS	Open-source	http://ccwu.me/vsfm/
SURE	Semi Global Matching	14-tägige Testversion	http://www.nframes.com/
Apero	Micmac	Open-source	http://logiciels.ign.fr/?Micmac

Oft wurden vor allem texturarme Bereiche nicht rekonstruiert. Es ist anzunehmen, dass die inhomogenen Ergebnisse durch unterschiedliche Algorithmen zur Punktwolkenberechnung in den Programmen hervorgerufen wurden (KERSTEN & LINDSTAEDT 2012). Da die verwendete Kamera vor der Messung nicht kalibriert wurde, wurden die Kameraparameter simultan in allen verwendeten Programmen mitgeschätzt. Nach der Erstellung der Punktwolke wurde diese abgespeichert und zur weiteren Verarbeitung in CloudCompare im „Polygon File Format“ (.ply) importiert. Anschließend erfolgte die manuelle Bereinigung der Punktwolke. Die Skalierung der Punktwolken erfolgte über die Kenntnis der Seitenlängen des Gemäldes von 121,3 cm × 63,6 cm (MARX 1996). Besser wären hierbei absolute Messungen mit übergeordneter Genauigkeit gewesen, um letztendlich auch absolute Deformationen bestimmen zu können. Denkbar wäre die Einrichtung eines Messtisches oder die Anbringung von Zielmarken, die zusätzlich zum Objekt in den Aufnahmen zu sehen sind. Allerdings sprachen praktische Bedingungen, die durch die Restauratoren vorgegeben wurden, gegen diese Maßnahmen. Die unbearbeitete und die bearbeitete Punktwolke sind in Abbildung 2 gegenübergestellt. Vor der Bearbeitung zeigt die erstellte Punktwolke noch die Staffelei sowie Teile des Tisches. Diese Punkte waren für weitere Berechnungen nicht notwendig und wurden manuell entfernt. Zusätzlich wurde die Punktwolke skaliert und so gedreht, so dass die z-Achse des Koordinatensystems in die Bildebene zeigt.

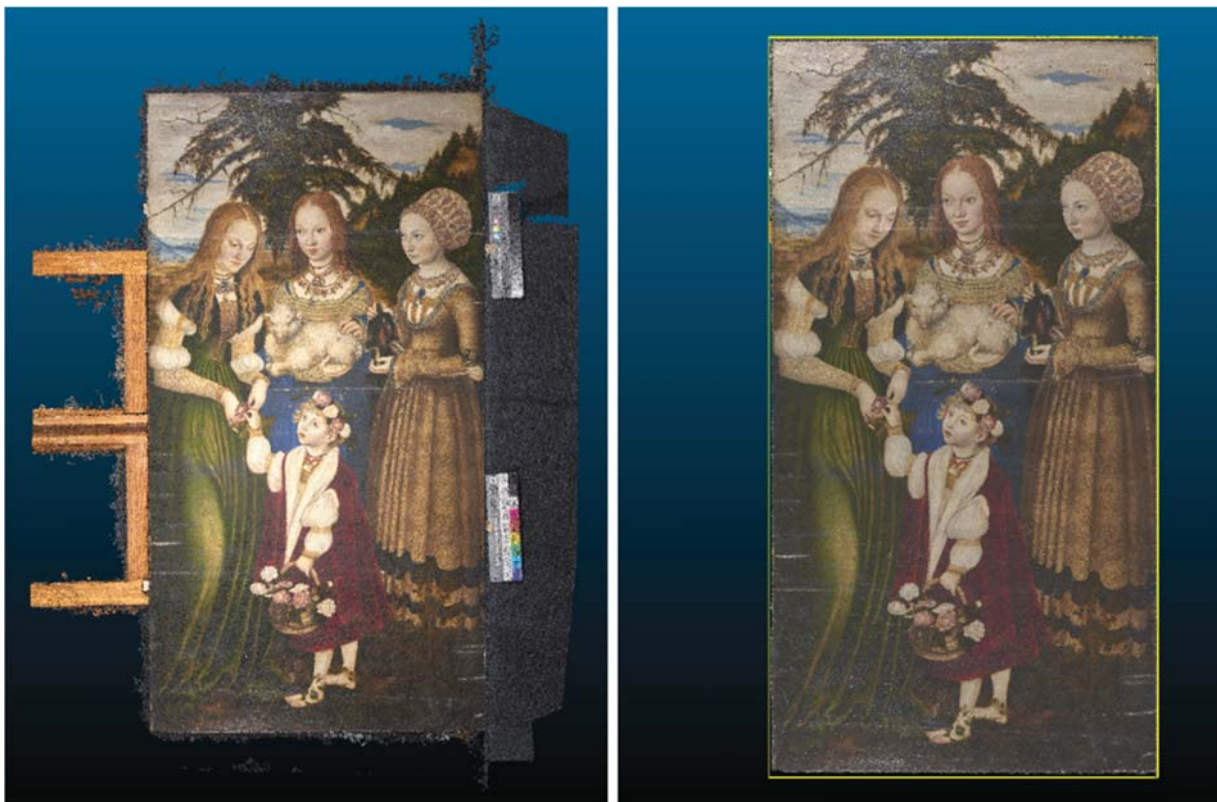


Abb. 2: Vergleich der unbearbeiteten Punktwolke aus Agisoft Photoscan (links) mit der bearbeiteten Punktwolke aus CloudCompare (rechts)

3 Untersuchungen am Cranach-Gemälde

Nach der Bereinigung der Punktwolken wurden nach Absprache mit den Staatlichen Kunstsammlungen Dresden verschiedene Ansätze zur Visualisierung der Deformationen verfolgt. Das Interesse der Restauratoren lag vor allem an der gezielten Betrachtung der einzelnen Zustände (Epochen) und der Darstellung von zwei verschiedenen Eigenschaften. Die erste Eigenschaft ist die Wölbung des Gemäldes, d.h. die Abweichungen der Holztafel zu einer festgelegten Bezugsebene. Die zweite Eigenschaft ist die sog. Binnenstruktur und stellt die deutlich kleineren Abweichungen des Gemäldes zu einer mathematisch festgelegten Fläche dar. Als mathematische Fläche wurde eine sog. Quadrik, d.h. eine spezielle Polynomfläche, gewählt und in die Punktwolke eingepasst (Abb. 3). Die Berechnung und Analyse von Wölbung und Binnenstruktur soll in den nächsten Abschnitten genauer beschrieben werden.

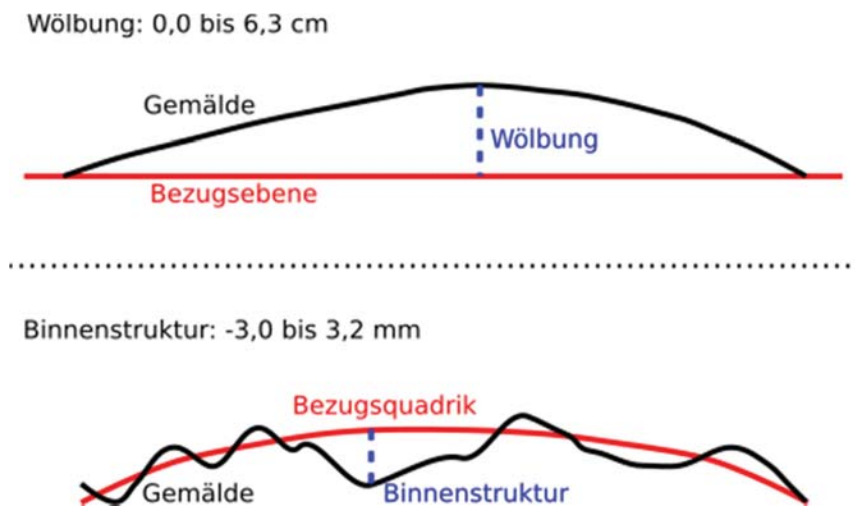


Abb. 3: Schematische Darstellung der Wölbung und der Binnenstruktur und deren Größenordnung

3.1 Analyse der Wölbung

Die Wölbung wurde durch den orthogonalen Abstand jedes Punktes der mittels SfM erzeugten Punktwolke zu einer Bezugsebene berechnet. Optimal wäre die Festlegung dieser Bezugsebene durch einen Messtisch bzw. ein absolutes Koordinatensystem. Aus praktischen Gründen war diese Lösung nicht umsetzbar. Es blieben somit verschiedene Möglichkeiten die Bezugsebene anhand der Messdaten selbst in CloudCompare festzulegen. Ansätze waren (a) die Definition der Ebene durch alle Punkte der Punktwolke als mittlere Ebene, (b) durch die vier Ecken des Gemäldes oder (c) durch drei Punkte. Eine mittlere Ebene durch alle Punkte wurde aufgrund von Ausreißerpunkten vor allem unterhalb der Punktwolke ausgeschlossen. Wenn nur drei einzelne Punkte zur Ebenenfestlegung verwendet werden, wird die mögliche Deformation einer Ecke nicht erfasst. Somit fiel die Wahl auf die Berechnung einer mittleren Ebene durch die vier Ecken des Gemäldes, womit zumindest eine geringe Redundanz erreicht wurde. Für die Berechnung der Eckpunkte wurden in den vier Ecken des Gemäldes Quadrate in der Größe von $0,5 \text{ cm}^2$ selektiert,

die von den Randbereichen der Punktwolke etwa einen Abstand von 1,0 cm haben. Ein solches Quadrat umfasst ca. 100 Einzelpunkte. Mittels des Statistical-Outlier-Removal-Filters (SOR) wurden pro Quadrat zuerst ca. 5-10 Ausreißerpunkte entfernt. Aus allen verbliebenen Punkten wurde jeweils der Schwerpunkt berechnet und als Eckpunkt definiert. Aus den vier so bestimmten Eckpunkten wurde in CloudCompare eine mittlere Ebene berechnet. Durch die geringe Überbestimmung wurde für jeden der Eckpunkte der verbleibende senkrechte Abstand zur jeweiligen mittleren Ebene bestimmt und als Indiz für eine mögliche übermäßige Deformation einer der 4 Ecken herangezogen. Tabelle 2 zeigt die mittleren quadratischen Abstände zur Bezugsebene. Aufgrund der kleinen Werte (nur bei Epoche 2 geringfügig größer als 1 mm) kann davon ausgegangen werden, dass die beschriebene Vorgehensweise zur Festlegung der Bezugsebene für alle 3 Epochen berechtigt ist.

Tab. 2: RMS der Abstände der vier Eckpunkte zur jeweiligen Bezugsebene

	Epoche 1	Epoche 2	Epoche 3
RMS der Abstände (mm)	0,069	1,219	0,809

Für die Berechnung der Wölbung wurde die auf diese Weise bestimmte Bezugsebene als Referenz und die SfM-Punktwolke als zu vergleichende Punktwolke festgelegt. Für die Berechnung wurde die Funktion „Cloud-to-Mesh“ verwendet, wobei die mittlere Ebene seitens CloudCompare als dreiecksvermaschtes Netz definiert wird (LEE & SCHACHTER 1980). Daraufhin wurden die Distanzen der Punktwolke zur mittleren Ebene in Normalenrichtung der Ebene berechnet und nach dem Eliminieren von Ausreißern farbskaliert dargestellt (Abb. 4).

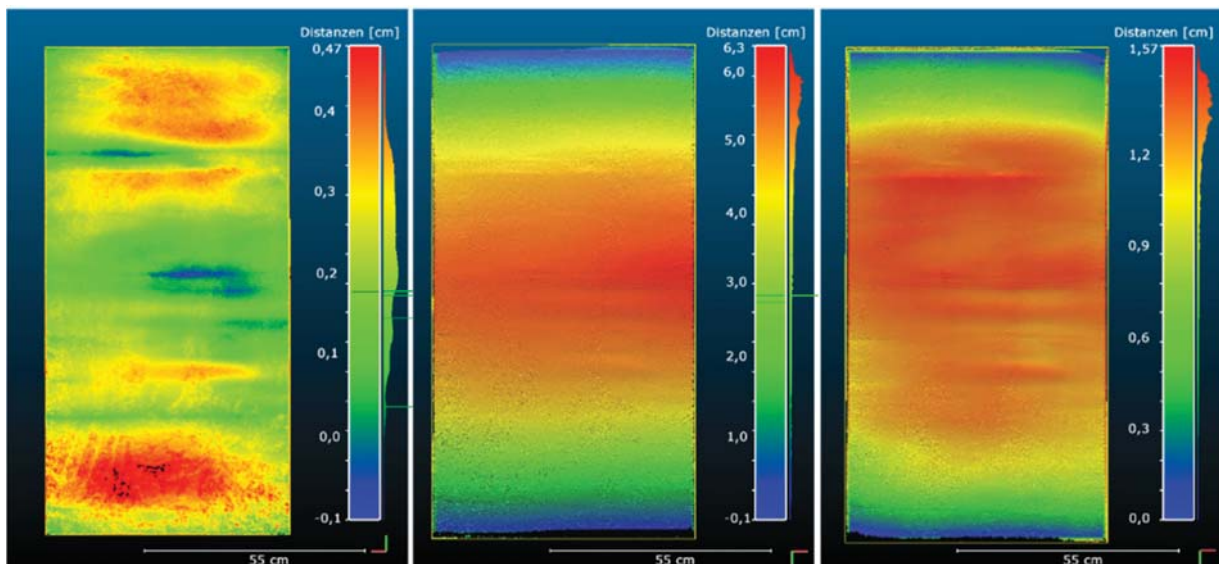


Abb. 4: Distanzen der mittels SfM erzeugten Punktwolke zu der jeweiligen Bezugsebene in allen Epochen; Aufsteigende Werte von Blau über Grün nach Rot; Links: Epoche 1 mit Distanzen von -0,1 cm bis 0,5 cm; Mitte: Epoche 2 mit Distanzen von -0,1 cm bis 6,3 cm; Rechts: Epoche 3 mit Distanzen von 0,0 cm bis 1,6 cm

Die Abbildung bestätigt die Erwartungen der Restauratoren und zeigt die Größenverhältnisse der Wölbung auf. In Epoche 1 war das Gemälde nicht bzw. kaum gewölbt, weshalb die Punktwolke sowohl unter als auch über der Bezugsebene lag. Die Holztafel glich in diesem Zustand einer starren Ebene mit den bereits angedeuteten Rissen in der Gemäldefläche. Diese sind bereits in der Darstellung der Wölbung in der oberen Bildhälfte als blaue horizontale Linien sichtbar. Die maximale Abweichung von der Bezugsebene trat am unteren Ende der Tafel mit einer Größenordnung von 0,47 cm auf.

In Epoche 2 lag die Tafel ohne Stützsystem auf der Seite und zeigte eine sehr starke Wölbung, die in der Mitte des Gemäldes maximal 6,30 cm aufweist. Epoche 3 zeigt einen ähnlichen Farbverlauf wie Epoche 2, allerdings sind die maximalen Abweichungen in der Bildmitte mit 1,57 cm deutlich geringer. Das neue Stützsystem der Staatlichen Kunstsammlungen Dresden sollte eine starke Wölbung vermeiden und das Gemälde gleichzeitig stabilisieren. Der geglättete Farbverlauf (im Vergleich zu Epoche 1) und die geringere Wölbung (im Vergleich zu Epoche 3) bestätigt die mit der neuen Restaurierungsmethode beabsichtigte Wirkung.

Um eine Vergleichbarkeit der verschiedenen Epochen herzustellen, wurden zusätzlich Längsprofile über die Gemäldefläche erstellt. Für die Berechnung wurden die eingefärbten Punktwolken reduziert. Die ermittelten Distanzen zur Bezugsebene wurden hierbei auf ein Gitter in z-Richtung (in die Bildebene) projiziert, wobei die Rastergröße des Gitters mit 1 cm² festgelegt wurde. Jeder Zelle des Gitters wurde der Mittelwert der projizierten Distanzen als Wert zugeordnet. Auf diese Weise konnten 64 Längsprofile über das 64 cm breite Gemälde erstellt werden. Die weitere Berechnung und Visualisierung der Profile wurde für jede Epoche in MATLAB realisiert (Abb. 5).

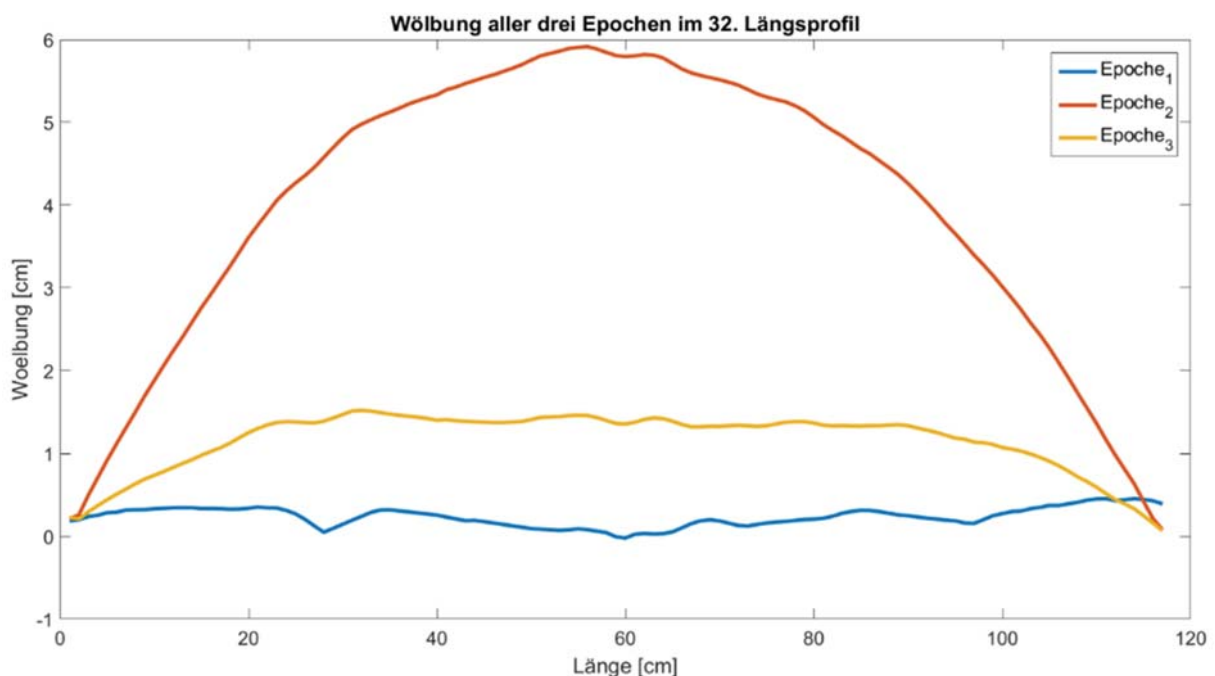


Abb. 5: Ausgewähltes Längsprofil (ca. in der Mitte der Holztafel) in allen drei Epochen

Auch in dieser Darstellung sieht man die bereits beschriebenen Eigenheiten der verschiedenen Epochen sehr gut. Epoche 2 zeigt die größte Wölbung. Epoche 1 gleicht einer Ebene und es sind bereits kleine Einschnitte (z.B. bei ca. 27 cm) in der Holztafel zu erkennen. Epoche 3 zeigt nur eine geringe Wölbung sowie keine Unstetigkeiten im Profilverlauf und deshalb vermutlich keine Risse.

3.2 Analyse der Binnenstruktur

Die Binnenstruktur des Gemäldes wurde aus dem Abstand der mittels SfM generierten Punktwolke und einer mathematischen definierten Fläche berechnet. Die mathematische Fläche sollte hierbei die Wölbung möglichst gut repräsentieren und die im Maßstab deutlich kleineren Deformationen aufzeigen.

Für die Untersuchungen wurde eine Methode gewählt, die sowohl die komplette Punktwolke berücksichtigt als auch eine konsistente Darstellung der Ergebnisse erlaubt. Deshalb wurde die in CloudCompare vorhandene Funktion „Fit Quadric“ verwendet, die eine Einpassung einer Quadrik, d.h. einer zweidimensionalen Polynomfläche in die Punktwolke erlaubt. In CloudCompare wird eine 2,5D-Quadrik der folgenden Form verwendet:

$$a + bx + cy + dx^2 + exy + fy^2 = z \quad (1)$$

Typische Formen für 2,5D-Quadriken sind elliptische bzw. hyperbolische Paraboloiden. Im Falle der Annäherung an die Gemäldeform (bzw. die Wölbung) beschreibt die Quadrik immer ein elliptisches Paraboloid. Lediglich in den Randbereichen (oberer und unterer Rand) des Gemäldes kam es zu Abweichungen, da sich die Holztafel in diesen Teilen stärker krümmt.

Nach der Berechnung der Quadrik für jede Epoche konnte wie schon bei der Wölbung die Differenz zwischen Bezugsfläche und SfM-Punktwolke berechnet werden. Da die Deformationen der Binnenstruktur alle ungefähr in der gleichen Größenordnung liegen, konnte eine einheitliche Farbskala für alle drei Epochen festgelegt werden (Abb. 6).

Die Abweichungen zur jeweiligen Bezugsquadrik liegen in jeder Epoche im Wertebereich von - 3,0 mm bis 3,2 mm. Für Epoche 1 ist die berechnete Bezugsquadrik fast identisch mit der Bezugsebene (siehe Kap. 3.1), da das Gemälde hier nahezu keine Wölbung aufweist, weshalb sich ein ähnliches Deformationsbild wie in Abbildung 4 ergibt.

Epoche 2 und 3 zeigen wie schon bei der Analyse der Wölbung einen ähnlichen Farbverlauf. Feine Risse in der Holztafel sind hier nicht zu erkennen und die Farbverläufe sind insgesamt glatter. Auffällig war der große rote Bereich im oberen Drittel an dessen Stelle in Epoche 1 noch ein Riss vorhanden war. Dieser rote Bereich konnte nach Angaben der Restauratoren durch eine Verklebung der Holztafel und eine damit verbundene Aufwölbung der Klebeschicht erklärt werden. Generell glich der Farbverlauf in Epoche 3 (neues Stützsystem) dem Verlauf in Epoche 2 (ohne stützende Parkettierung), wobei jedoch in Epoche 3 die Wölbung deutlich reduziert werden konnte (vgl. Kap. 3.1).

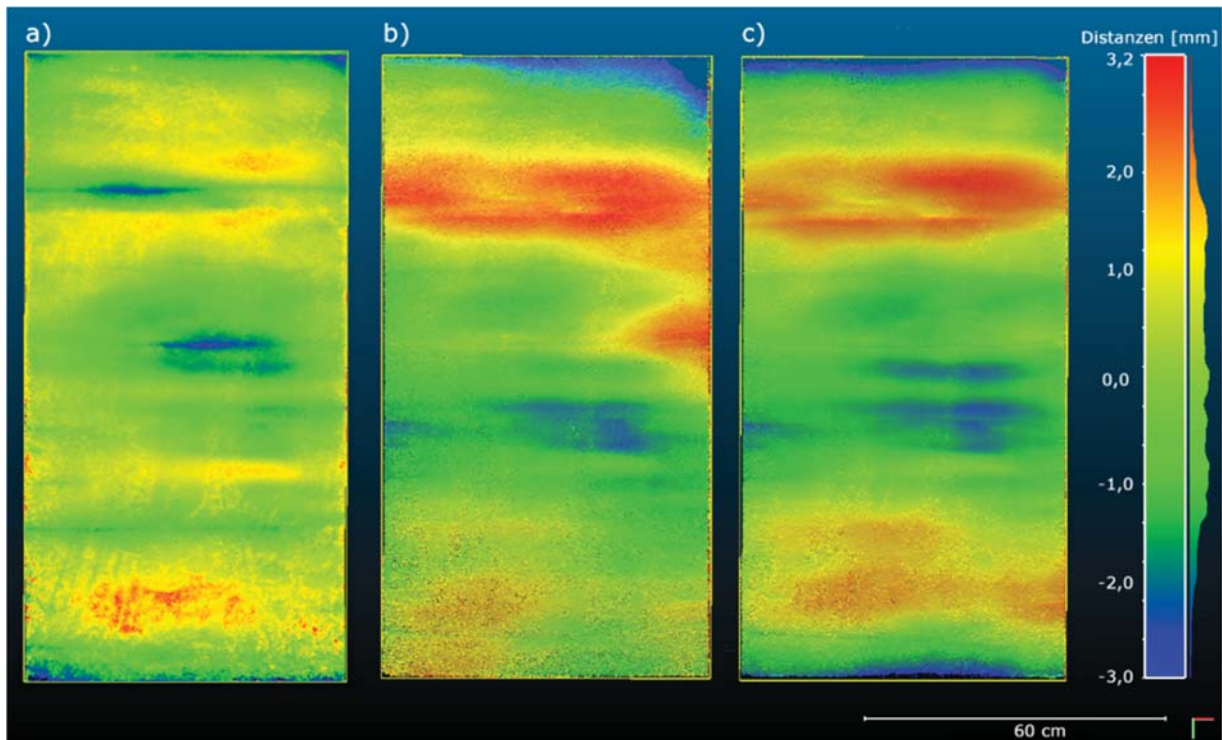


Abb. 6: Distanzen der mittels SfM erzeugten Punktwolke zu der jeweiligen Bezugsquadrik in allen Epochen; Aufsteigende Werte von Blau über Grün nach Rot von -3,0 mm bis 3,2 mm; a) Epoche 1; b) Epoche 2; c) Epoche 3

Zusätzlich zu dieser Visualisierung wurden auch für die Binnenstruktur alle Punktwolken auf ein gleichmäßiges Raster projiziert, um anschließend Profildarstellungen erzeugen zu können. Da die Abweichungen der Binnenstruktur einen deutlich geringeren Wertebereich besitzen, wurden diese zunächst auf Plausibilität überprüft. Hierfür wurde eine vor der Restaurierung händisch erzeugte Schadenskartierung der Restauratoren herangezogen. Risse, die in dieser Kartierung auftauchen, sollten auch in einem Längsprofil über die farbkodierte Darstellung der Gemäldefläche sichtbar werden (Abb. 7).

Die Abbildung zeigt, dass die Schadenskartierung mit den erzeugten Längsprofilen übereinstimmt. Risse und weitere kleine Unstetigkeiten in der Oberflächenstruktur werden in der Profildarstellung sichtbar. Nach Informationen der Restauratoren durchtrennten schmale Risse die komplette Holztafel und im Profil müsste somit eine Lücke an der Stelle eines Risses sichtbar sein. Zum einen aufgrund des flächenbasierten Bildzuordnungsverfahren und zusätzlich aufgrund der Projektion der Punktwolke auf ein gleichmäßiges Raster mit großer Rasterweite wurde für diese Stellen stattdessen lediglich ein sehr niedriger Wert berechnet, der zumindest auf einen Riss hindeutet.

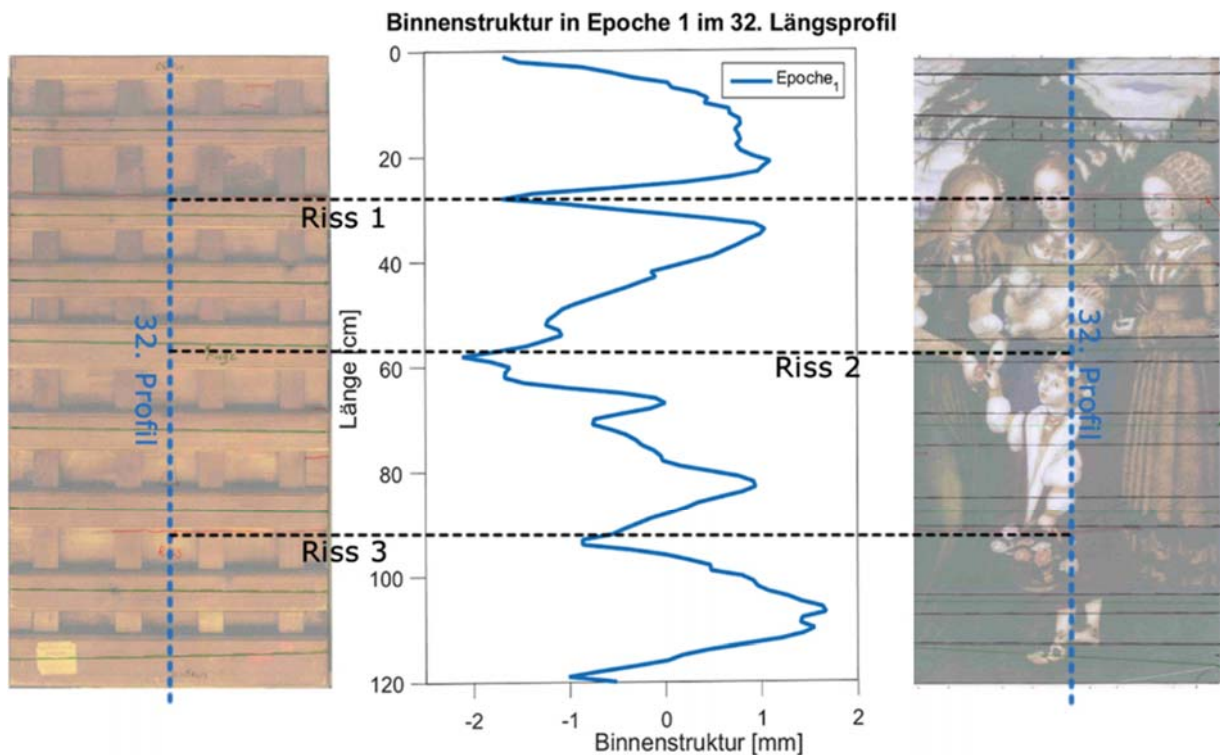


Abb. 7: Vergleich eines Längsprofils über die Binnenstruktur mit der Schadenskartierung der Vorder- und Rückseite (Michael Schweiger, Staatliche Kunstsammlungen Dresden) des Gemäldes vor der Restaurierung; Die Abbildung zeigt drei Risse, die sich auch im Längsprofil deutlich widerspiegeln.

Eine weitere Möglichkeit der Visualisierung bot die Darstellung der ursprünglichen Schadenskartierung mit der darübergelegten farbkodierten transparenten Punktwolke (Abb. 8). In dieser Abbildung sind nicht nur die einzelnen Risse und Besonderheiten des ursprünglichen Gemäldes sichtbar. Man kann zudem erkennen wie weit sich die Deformationen ausbreiten und an welchen Stellen sie ineinander übergehen. Eine exakte Lokalisierung der größten Schäden und der weniger deformierten Bereiche ist auf einen Blick zu erfassen.

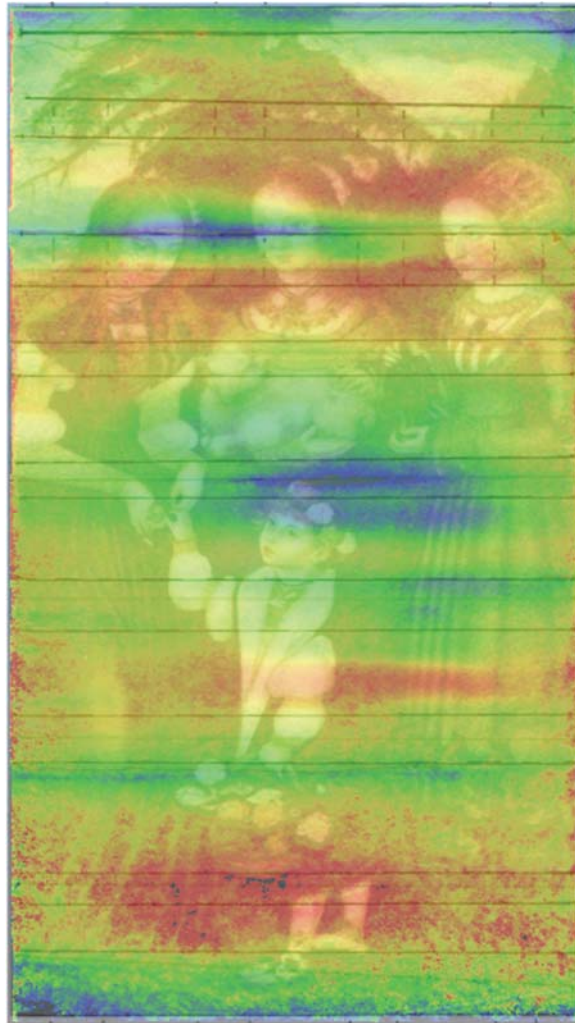


Abb. 8: Schadenskartierung der Vorderseite (Michael Schweiger, Staatliche Kunstsammlungen Dresden) überlagert mit der semi-transparenten Binnenstrukturdarstellung von Epoche 1

4 Fazit & Ausblick

Die Arbeit zeigt, dass photogrammetrische Methoden für die Beurteilung der Wirkung verschiedener Restaurierungsverfahren bzw. zur Überwachung des Restaurierungsvorganges geeignet sind. Für die Anwendung kamen ausschließlich berührungslose Techniken in Frage, wobei vor allem Structure-from-Motion (SfM) als Werkzeug für die Deformationsanalyse zum Einsatz kam. Durch vergleichende Messungen mit einem Streifenlichtscanner an einem Testobjekt zusätzlich zu SfM konnte nachgewiesen werden, dass die Genauigkeitsanforderungen im Millimeter- und Submillimeterbereich erreicht wurden.

Von den in den Arbeiten verwendeten Programmen zur Erstellung einer Punktwolke aus mehreren Aufnahmen, hat sich bei einem flachen Gemälde vor allem Agisoft Photoscan als zuverlässig erwiesen. Punktwolken, die mit diesem Programm berechnet wurden, konnten das Gemälde fast

vollständig repräsentieren. Aufnahmen des Cranach-Gemäldes wurden in drei verschiedenen Epochen mit altem Stützsystem (Parkettierung), ohne Stützsystem und mit neuem Stützsystem gemacht, wobei für jede Epoche Wölbung und Binnenstruktur analysiert wurden. Die Ergebnisse bestätigten die Erwartungen der Restauratoren und konnten zum Verständnis von Prozessen die während der Restaurierung ablaufen, sowie zur messtechnischen und objektiven Beurteilung der Effekte verschiedener Restaurierungsmethoden beitragen. Visualisierungen der produzierten dreidimensionalen Daten beispielsweise eine Profildarstellung oder eine Überlagerung einer farbkodierten Punktwolke können die Schadenskartierung vor der Restaurierung sinnvoll unterstützen. Zudem können Restaurierungs- und Zwischenergebnisse besser erfasst und dokumentiert werden. Der gleiche Arbeitsablauf soll auch in Zukunft angewandt werden, um zu späteren Zeitpunkten weitere Zustände des Gemäldes zu untersuchen. Denkbar wäre auch die abgewandelte Anwendung auf Leinwandgemälde oder die Automatisierung der Visualisierung um schneller Ergebnisse einer Restaurierung abschätzen zu können.

5 Literaturverzeichnis

- HORN, B., 1987: Closed-form solutions of absolute orientation using unit quaternions. *Journal of the Optical Society of America A* **4**(4), 629-642.
- KERSTEN, T. & LINDSTAEDT, M., 2012: Generierung von 3D-Punktwolken durch kamera-basierte low-cost Systeme – Workflow und praktische Beispiele. *Terrestrisches Laserscanning 2012 (TLS 2012)*, Schriftenreihe des DVW, Band 69, Beiträge zum 121. DVW-Seminar am 13. und 14. Dezember 2012 in Fulda, Wißner-Verlag, Augsburg, 25-46.
- LEE, D. T. & SCHACHTER B. J., 2004: Two Algorithms for Constructing a Delauney Triangulation. *International Journal of Computer Vision* **9**(3), 219-242.
- LOWE, D. G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* **60**(2), 91-110.
- LUHMANN, T., 2010: *Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen*. Wichmann, VDE Verlag GmbH, Berlin/Offenbach.
- MARX, H., 1996: *Lukas Cranach der Ältere. Der linke Flügel (Innenseite) des Katharinenaltars von 1506: Staatliche Kunstsammlungen Dresden, Gemäldegalerie Alter Meister. Kulturstiftung der Länder (Hrsg.)*.
- NICOLAUS, K., 2003: *DuMonts Handbuch der Gemäldekunde: Gemälde erkennen und bestimmen*. DuMont Buchverlag, Köln.
- SOOT, M., SCHULZE, M. & MULSOW, C., 2015: Untersuchung zur Genauigkeit der Kamerakalibrierung über Merkmalspunkten in SfM-Werkzeugen. *Photogrammetrie, Laserscanning, Optische 3D-Messtechnik – Beiträge der Oldenburger Tage 2015*, Luhmann, T. & Müller, C. (Hrsg.), Wichmann, VDE Verlag GmbH, Berlin/Offenbach, 269-277.
- WACKROW, R., CHANDLER, J. H. & GARDNER, T., 2008: Minimising systematic errors in DEMs caused by an inaccurate lens model. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **37**, B5, 1-6.

Web-Visualisierung komplexer 3D-Stadtmodelle mit pyRT

MARTIN CHRISTEN¹ & MARKUS FEHR¹

Zusammenfassung: Das interaktive Rendern von komplexen 3D-Stadtmodellen mit hochauflösten Texturen und einer hohen Anzahl an Geometrien ist im Webbrowser eine Herausforderung. In diesem Beitrag wird gezeigt, wie sehr grosse texturierte CityGML Datensätze unter Verwendung von GPUs in der cloud gerechnet und im Webbrowser effizient dargestellt werden können. Dabei wird das neu entwickelte Open-Source Rendering-Tool «pyRT» eingesetzt, welches unter Verwendung der Programmiersprache Python 3D-Szenen rendern kann. Im Beitrag wird auch gezeigt wie mit pyRT globale Beleuchtungsmodelle approximiert werden, um 3D-Stadtmodelle in hoher Qualität darzustellen.

Die physikalisch plausiblen berechneten 3D Ansichten werden in der Cloud berechnet und auf dem Webclient dargestellt. Mittels hochperformanten GPU-unterstützten Ray-Tracing können somit auch sehr komplexe 3D-Stadtmodelle mit einer sehr grossen Anzahl von Geometrien effizient dargestellt werden.

1 Einleitung

Die Visualisierung von 3D-Stadtmodellen im Webbrowser und auf mobilen Geräten hat in den letzten Jahren an Bedeutung gewonnen. Dank „out of core-Rendern“-Ansätzen und Level of Detail lassen sich sehr grosse 3D-Stadtmodelle mit weit über 100 Millionen Dreiecken und 12 GB Speicherplatz für die Geometrie darstellen (JIANG et al. 2014). Durch die Verwendung von Level of Detail und dem meist langsamen Aufbau der 3D-Modelle müssen jedoch Kompromisse eingegangen werden, was immer auch reduzierte Qualität bedeutet. Zudem werden bei der Darstellung nach wie vor sehr einfache Beleuchtungsmodelle eingesetzt.

Um qualitativ hochwertige 3D-Stadtmodelle im Webbrowser interaktiv zu rendern müssen neue Ansätze gefunden werden. Eine Möglichkeit dazu ist die Verwendung von interaktiven orthogonaler 3D-Karten, welche in der Cloud gerechnet werden und als G-Buffer (SAITO 1990). im Client dargestellt werden (CHRISTEN 2016). In diesem Beitrag wird dieses Verfahren erweitert, um qualitativ hochwertige 3D-Ansichten von untexturierten 3D-Stadtmodellen in Echtzeit zu erstellen, ohne dabei die Originalgeometrie zu reduzieren. Dabei wird ein neu entwickelter Ray-Tracer vorgestellt, welcher GPU und CPU optimiert neben dem G-Buffer auch eine Ambient Occlusion Map berechnen kann. Dieser ist optimiert auf Berechnungen von Schrägansichten aufgeteilt im Quadtrees und berechnet in der Cloud (CHRISTEN 2016).

2 Stand der Technik

2.1 GPU unterstütztes Ray-Tracing

Der erste Versuch von Raytracing auf einer GPU zu beschleunigen wurde im Jahr 2002 mit der Ray Engine von CARR et al. (2002) gemacht. Die Ray Engine konnte nur die Strahl-Dreieck

¹ Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: [martin.christen, markus.fehr]@fhnw.ch

Intersektion auf der GPU berechnen. Das Übertragen der Geometrie auf die GPU war zu dieser Zeit die grösste Schwierigkeit. PURCELL et al. (2002) und PURCELL (2004) konnten alle Berechnungen auf die GPU auslagern, wie zum Beispiel die Primärstrahlen, die Beschleunigungsstrukturen, Schattierung, Strahl-Dreieck Intersektion. Dieses Verfahren war die Basis einiger anderer Implementationen, wie von KARLSSON (2004), ERNST et al. (2004) und CHRISTEN (2005). Das Problem dieser Verfahren war, dass sie mangels einer geeigneten Hochsprache nur über Umwege durch Verwendung von Vertex- und Fragmentshadern implementiert werden konnten.

Eine weitere Schwierigkeit war das traversieren von hierarchischen Beschleunigungsstrukturen auf der GPU. FOLEY & SUGERMAN (2005) implementierten zwei Verfahren des kd-Baums auf der GPU. Weiter entwickelten HORN et al. (2007) und POPOV et al. (2007) je einen Raytracer basierend auf der kd-Tree Beschleunigungsstruktur. Auch BVH Beschleunigungsstrukturen wurden entwickelt, so zum Beispiel von THRANE & SIMONSEN (2005), CARR et al. (2006) oder von GÜNTHER et al. (2007).

Kürzlich AMD hat das Projekt GPUOpen veröffentlicht - ein Open Source Produkt, welches verschiedene Libraries, Tools und SDKs beinhaltet, um bessere Kontrolle über die GPU erhalten (THIBIEROZ 2016).

Eine dieser SDKs ist die Radeon Rays SDK (ehemals FireRays). Radeon Rays ist eine high-performance Raytracing Library, optimiert auf GPU beschleunigte Intersektionsberechnung. Sie beinhaltet eine C++ API und basiert auf OpenCL. Das heisst sie unterstützt alle Plattformen basierend auf den OpenCL 1.2 Standard (AMD, 2016a). Die Radeon Rays SDK ist nicht limitiert auf AMD Hardware, sondern ist Hardware unabhängig. Lizenziert ist Radeon Rays unter der MIT Lizenz (AMD, 2016a).

Auch CPU-optimierte Ray-Tracer haben in den letzten Jahren an grosser Bedeutung gewonnen, so hat Intel mit Embree einen Raytracing-Kernel entwickelt, welcher sich für fotorealistisches rendern auf den neusten Intel Prozessoren eignet (INTEL CORPORATION 2016b). Embree ist ebenfalls ein Open Source Produkt, veröffentlicht unter der Apache 2.0 Lizenz. Ray Tracing Kernel verfügen normalerweise nur über basis-Funktionalitäten wie Intersektionstest

2.2 Ambient Occlusion

Ambient Occlusion ist insbesondere in der Filmbranche ein Standardverfahren geworden für qualitativ hochwertiges Rendering (LAINE & KARRAS 2010). Erstmals vorgestellt wurde es von ZHUKOV et al. (1998) und danach in der Rendering Community aufgenommen (LANDIS 2002).

Ambient Occlusion wurde von ZHUKOV et al. (1998) folgendermassen definiert:

$$W(p, n) = \frac{1}{\pi} \int_{\Omega} \rho(D(p, \omega)) (\omega \cdot n) d\omega$$

W enthält die gewichtete Anzahl Verdeckungen; p ist der Punkt der Oberfläche, welcher verdeckt wird und n dessen Normale. Das Integral läuft dabei über die Hemisphäre, welche mittels n orientiert wird.

$D(p, \omega)$ misst die Distanz zur am nächsten liegenden Verdeckung in Richtung ω und mit ρ wird ein Faktor bestimmt, welcher die Distanz zur Verdeckung in einen Wert zwischen 0 und 1 konvertiert und deshalb auch eine maximale Distanz benötigt. Üblicherweise wird das Integral mittels Monte Carlo-Methoden approximiert, so dass nur eine vordefinierte Anzahl zufälliger

Strahlen in der Hemisphäre definiert werden. Typischerweise sind dies im Bereich $N = 200$ bis $N = 1000$ Strahlen pro Punkt. Ambient Occlusion kann somit folgendermassen vereinfacht werden:

$$W(p, n) \approx \frac{1}{N} \sum_{i=1}^N \rho(D(p, \omega_i))$$

Später wurde Ambient Occlusion auch für Real-Time Applikationen erweitert wie beispielsweise mittels Screen Space Ambient Occlusion (BAVOIL 2008) oder Screen Space Far-Ambient Obscuration (TIMONEN 2014), Abbildung 1. Qualitativ ist dies jedoch weniger gut als die Originallösung.



Abb. 1: Screen-Space Far-Field Ambient Obscuration (TIMONEN 2014)

2.3 Serverseitiges Vorrendern von 3D-Geodaten

In letzter Zeit wurden Methoden vorgestellt, um Geodaten durch vorgänges rendern effizient darzustellen. HILDEBRANT et al. (2011) haben zum Beispiel eine Architektur zum serverseitigen Rendering von 3D-Panoramen mittels Cube-Maps vorgestellt. Bei dieser Lösung werden aus dem 3D-Modell verschiedene Panoramen vorgerechnet und mit einem Panorama-Viewer dargestellt. Eine andere Lösung, bei der komplexe 3D-Modelle dargestellt wurden, ist das Zusammensetzen von Schrägansichten aus Kacheln, welche es ermöglicht diese im Webbrowser und auf mobilen Geräten darzustellen (KLIMKE et al. 2014). Auch erwähnenswert ist das HuMoRS system (Huge models Mobile Rendering System), bei dem die 3D-Modelle interaktiv über das Netzwerk gestreamt werden, und es ermöglicht wird, sich auf einem mobilen Gerät interaktiv durch eine Szene zu bewegen (RODRIGEZ 2014). Das Open Geospatial Consortium (OGC) hat sich in letzter

Zeit auch mit der Interoperabilität von 3D-Geodaten für die Visualisierung auseinandergesetzt, insbesondere auch für den Austausch von komplexeren Szenen (OGC 2015).

In der Arbeit von CHRISTEN & NEBIKER (2015) wurden 3D-Geodaten mittels Ray-Tracing gerechnet und der resultierende G-Buffer in einer Quadtree-Struktur gespeichert um 3D Karten ähnlich wie 2D-Karten zu behandeln, wie in Abbildung 2 zu sehen ist.



Abb. 2: Beispiele für Cloudgerenderte, gekachelte Schrägansichten (CHRISTEN & NEBIKER 2015)

3 Rendern von Schrägansichten mit pyRT

Das Projekt pyRT wurde gestartet um einen neuen Raytracer zu erstellen, welcher die Resultate in einem einfach erweiterbaren G-Buffer liefert. So ist es beispielsweise möglich Tiefenkarten mit mehr als 64-bit pro Pixel zu generieren, oder Objekt-ID maps mit direkter Referenz auf das Objekt. Mit herkömmlichen Open-Source Ray-Tracern wie POV-Ray (POV-RAY 2017) oder RENDERMAN (2017) ist dies nur über ineffiziente Umwege möglich (LINDENBERGER 2016).

In diesem Beitrag wird jedoch nicht auf diese eingegangen, sondern auf die Ambient Occlusion map, welche zusätzlich generiert werden kann, um die Qualität zu verbessern.

pyRT ist in Python geschrieben, benutzt aber Bindings zu Embree und AMD FireRays API, um auch moderne CPUs und GPUs direkt zu nutzen. Es werden nun verschiedene Szenen angesehen und verglichen. Texturen interessieren bei der Ambient Occlusion nicht, da nur die Ambiente Beleuchtung gerechnet wird. In einem nächsten Prozess kann die Textur mit der Ambient Occlusion kombiniert werden. Die Berechnung des G-Buffers erfolgt dabei im selben Rendering Schritt, es müssen dafür keine doppelten oder dreifachen Berechnungen durchgeführt werden wie z.B. bei der Verwendung von Renderman.

Tab. 1: Verwendete 3D-Modelle für die Tests

Modell	Anzahl Dreiecke	Dateigrösse (Wavefront OBJ Format)
Sponza	66'447	5.53 MB
Basel	128'993	11.90 MB
Drache	871'308	62.59 MB

3.1 Sponza Modell

Die Ambient Occlusion Map der Sponza Szene wurde als erstes gerendert. In der Abbildung 3 sind drei Ansichten (S1 bis S3) mit verschiedenen Maximaldistanzen dargestellt (FEHR 2017). Wird die Distanz zu kurz gewählt, wird das Bild heller und hat teilweise unschöne Übergänge. Wird die Distanz zu gross gewählt verdunkelt sich das Bild zunehmend.

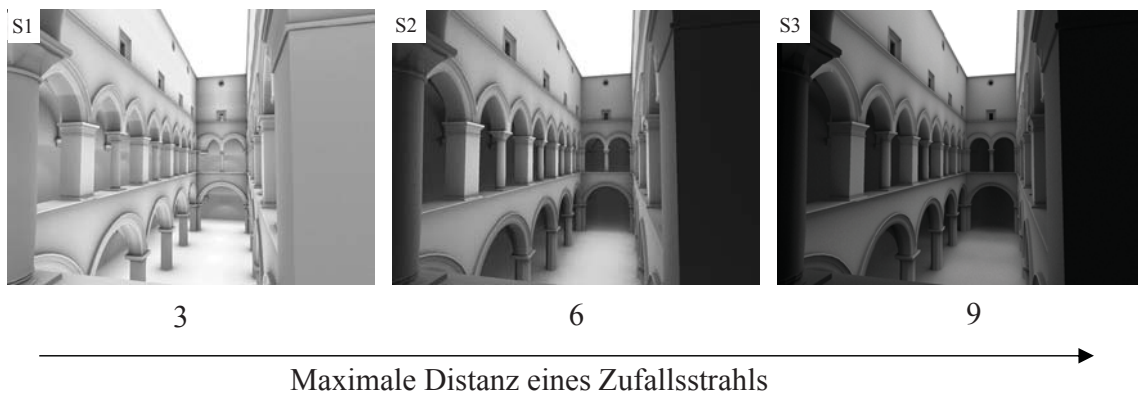


Abb. 3: Einfluss der Maximalen Distanz auf die Ambient Occlusion (FEHR 2017)

Auch die Anzahl Strahlen hat einen Einfluss auf die Qualität. Werden zu wenig Strahlen gewählt (S4, S5), so wirkt das Resultat eher pixelig, wie in Abbildung 4 zu sehen ist.

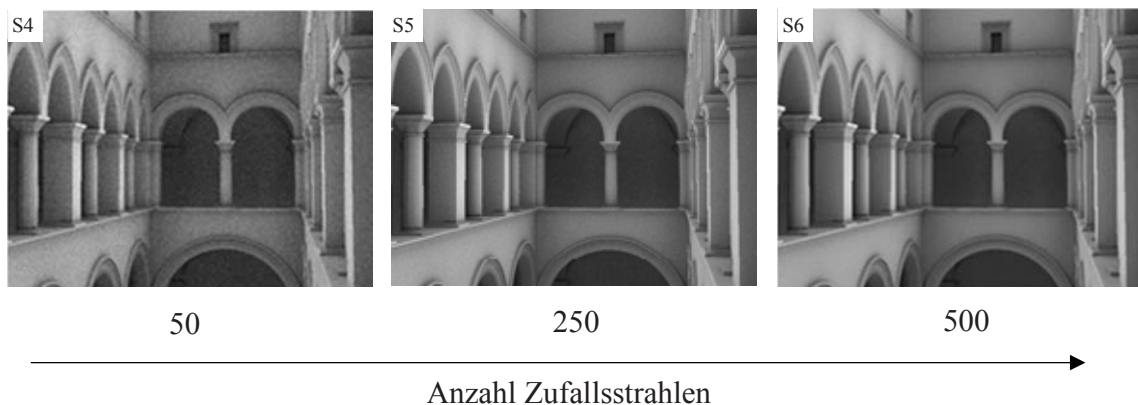


Abb. 4: Einfluss der Anzahl Strahlen auf die Szene (FEHR 2017).

Noch als Vergleich soll die Szene mit dem Blinn-Phong Beleuchtungsmodell (PHONG 1975) berechnet werden. Dabei wird es einmal mit Schatten (S8) und ohne Schatten (S7) gerendert. Intuitiv kann festgestellt werden, dass die visuelle Qualität signifikant schlechter ist, wie in Abbildung 5 zu sehen ist.

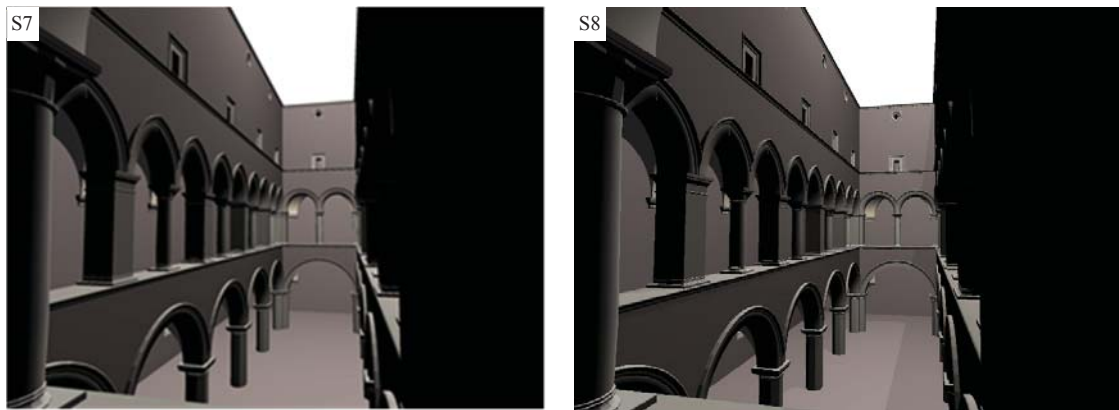


Abb. 5: Die Szene mit dem Blinn-Phong Modell gerechnet, mit und ohne Schatten (FEHR 2017)

3.2 Modell von Basel

Das 3D-Stadtmodell Basel vom Grundbuch und Vermessungsamt Basel Stadt wurde verwendet um die Ambient Occlusion zu rechnen. Dieses wurde wiederum mit Blinn Phong (BS1) und Ambient Occlusion berechnet und ist in Abbildung 6 zu sehen.

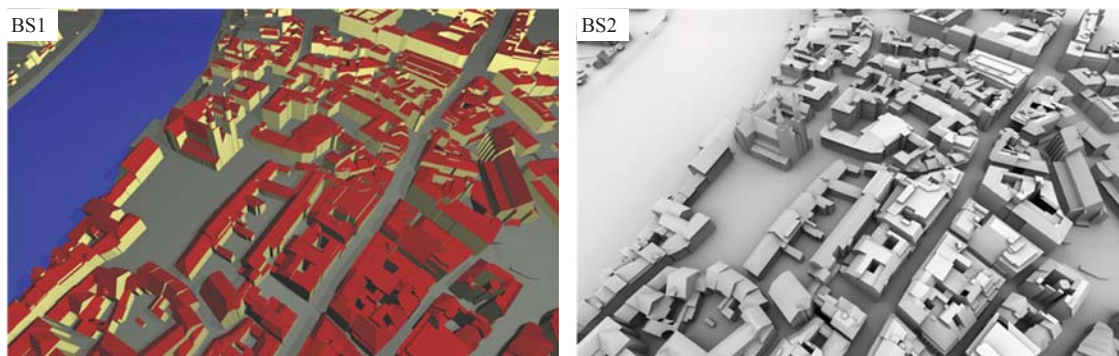


Abb. 6: Die Basel Szene mit dem Blinn-Phong (links) und Ambient Occlusion (rechts) (FEHR 2017)

3.3 Drachenmodell

Um auch mit einem Modell mit lokal höherer Anzahl Dreiecke zu testen, wurde das bekannte Drachenmodell mit Phong und mit Ambient Occlusion gerendert, wie in Abbildung 7 zu sehen ist.

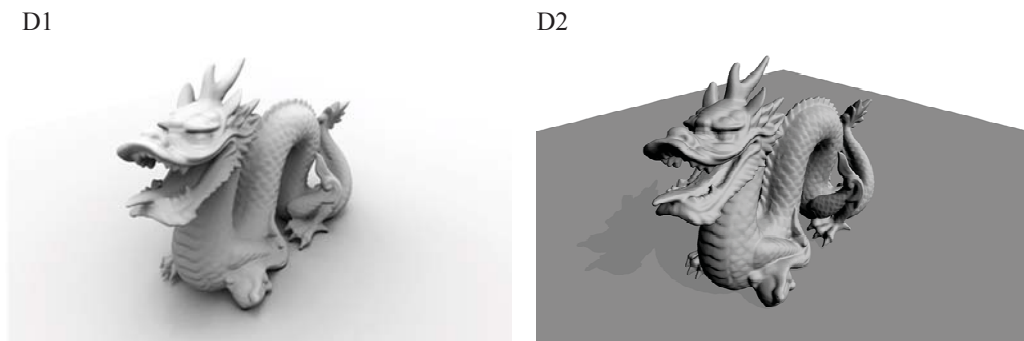


Abb. 7: Drachenszene mit Ambient Occlusion (links) und Phong (rechts)

4 Fazit und Ausblick

Es konnte gezeigt werden, dass mit pyRT ein Ray Tracer entwickelt wurde, welcher in einem einzigen Rendering-Schritt einen G-Buffer inklusive optionaler Ambient Occlusion Map generieren kann. Dabei können zur Beschleunigung Raytracing Kernel von AMD und Intel verwendet werden. Der Renderer eignet sich dazu, serverseitig 3D-Ansichten vorzurechnen, um diese dann interaktiv in einem (Web-)Client darzustellen. pyRT unterstützt im Moment allerdings noch keine Texturen, deren Support soll dann in einer späteren Version auch noch implementiert werden.

5 Literaturverzeichnis

- ÁFRA A.T., WALD I., BENTHIN C. & WOOP S., 2016: Embree ray tracing kernels: overview and new features. ACM SIGGRAPH 2016 Talks (SIGGRAPH '16), ACM, New York, USA.
- AMD, 2016a: Introducing the Radeon Rays SDK.
- AMD, 2016b: Radeon™ R7 Series Graphics Cards. AMD. <http://www.amd.com/en-us/products/graphics/notebook/r7-m200#>, letzter Zugriff am 22.12.2016.
- BARRINGER, R. & AKENINE-MÖLLER, T., 2014: Dynamic ray stream traversal. ACM Transactions on Graphics (TOG) **33**(4), article 151, 9 p.
- CARR, N. A.; HALL, J. D. & HART, J. C., 2002: The Ray Engine. Proceedings of Graphics Hardware **2**, 37-46.
- CARR, N. A., HOBEROCK, J. & CRANE, K., 2006: Fast GPU Ray Tracing of Dynamic Meshes using Geometry Images. Proceedings of Graphics Interface 2006, Canadian Information Processing Society, 203-209.
- CHRISTEN, M., 2005: Implementing Ray Tracing on the GPU. ShaderX4 - Advanced Rendering Techniques, W. Engel (ed.) Charles River Media, 413-424.
- CHRISTEN, M., NEBIKER, S. & LOESCH, B., 2012: Web-Based Large-Scale 3D-Geovisualisation Using WebGL. International Journal of 3-D Information Modeling, **1**(3), 16-25.
- CHRISTEN, M., HÜRBI, K. & NEBIKER, S., 2014: OpenWebGlobe: 3D-Visualisierung und Caching von globalen Stadtmodellen aus OpenStreetMap mittels Cloud-basiertem Framework. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation eV, Band 23, Potsdam, Seyfert, E., Gülch, E., Heipke, C., Schiewe, J., Sester, M. (Hrsg.), Beitrag 250.
- CHRISTEN, M. & NEBIKER S., 2015: Visualisation of Complex 3D City Models on Mobile Webrowsers using Cloud-based Image Provisioning, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-3/W5.
- CHRISTEN, M., 2016: Visualisierung von 3D-Geodaten im (mobilen) Webbrowser. Geoinformationssysteme 2016, Kolbe, T. H.; Bill, R.; Donaubaue, A. (Hrsg.), Wichmann, 2016, ISBN 978-3-87907-610-9
- EVANS, A., ROMEO, M., BAHREHMANN, A., AGENJO, J., & BLAT, J., 2014: 3D graphics on the web: A survey. Computers & Graphics **41**(0), 43-61.
- ERNST, M., 2009: Ray Tracing Techniques for Hybrid and Photorealistic Rendering. Dissertation, Friedrich Alexander Universität, Erlangen-Nürnberg.

- FEHR, M., 2017: Beleuchtungsmodelle für 3D-Stadtmodelle mit GPU-optimiertem Rendering in der Cloud. Master-Thesis, Institut Vermessung und Geoinformation, Fachhochschule Nordwestschweiz
- FOLEY, T. & SUGERMAN, J., 2005: KD-tree Acceleration Structures for a GPU Raytracer. ACM SIGGRAPH/EUROGRAPHICS conference on Graphics hardware, 15-22.
- GÜNTHER, J., POPOV, S. & SEIDEL, H.-P., 2007: Realtime Ray Tracing on GPU with BVH-based Packet Traversal. IEEE Symposium on Interactive Ray Tracing, Ulm, 113-118.
- HILDEBRANDT, D., KLIMKE, J., HAGEDORN, B. & DÖLLNER, J., 2011. Service-oriented interactive 3D visualization of massive 3D city models on thin clients. Proceedings of the 2nd International Conference on Computing for Geospatial Research & Applications (COM.Geo '11). ACM, New York, NY, USA
- INTEL CORPORATION, 2016a: A Ray Tracing Based Rendering Engine for High-Fidelity Visualization. OSPRay. <http://www.ospray.org/>, letzter Zugriff am 26.09.2016.
- INTEL CORPORATION, 2016b: Embree Overview. Embree - High Performance Ray Tracing Kernels. <https://embree.github.io/>, letzter Zugriff am 26.09.2016.
- JIANG, S., SAJADI, B., IHLER, A. & GOPI, M., 2014: Optimizing redundant-data clustering for interactive walkthrough applications. The Visual Computer, **30**(6-8), 637-647.
- KAJALIN, V., 2009: Screen-Space Ambient Occlusion. ShaderX7 - Advanced Rendering Techniques, W. Engel (ed.), Charles River Media, 413-424.
- KARLSSON, F., 2004: Ray tracing fully implemented on programmable graphics hardware. Master Thesis, Chalmers University of Technology, Göteborg.
- KLIMKE, J., HAGEDORN, B. & DÖLLNER, J., 2014: Scalable Multi-Platform Distribution of Spatial 3D Contents. International Journal of 3-D Information Modeling (IJ3DIM) **3**(3), 35-49.
- LAINE, S. & KARRAS, T., 2010: Two methods for fast ray-cast ambient occlusion. Computer Graphics Forum **29**(4), Blackwell Publishing Ltd., 1325-1333.
- LANDIS, H., 2002: RenderMan in Production. ACM SIGGRAPH 2002, Course 16.
- LINDENBERGER, C., 2015: Rendertechniken für komplexe Cloud-Basierte 3D-Stadtmodelle. Master-Thesis, Institut Vermessung und Geoinformation, Fachhochschule Nordwestschweiz.
- MONGODB 2015: The MongoDB 3.0 Manual. <http://docs.mongodb.org/manual/>, last access November 20, 2015.
- NEBIKER, S., BLEISCH, S. & CHRISTEN, M., 2010: Rich point clouds in virtual globes – A new paradigm in city modeling? Computers, Environment and Urban Systems **34**(6), 508-517.
- OPEN GEOSPATIAL CONSORTIUM (OGC), 2015: OGC seeks public comment on candidate 3D Portrayal Service Standard. OGC Press Release, last access November 19, 2015 <http://www.opengeospatial.org/pressroom/pressreleases/2165>
- PHONG, B.T., 1975: Illumination for Computer Generated Pictures. Communications of the ACM **18**(6), 311-317.
- POVRAY, 2017: Persistence of Vision Renderer „POV-Ray“. <http://povray.org>, last access January 5, 2017.
- PURCELL, T. J., 2004: Ray Tracing on a Stream Processor. PhD thesis, Stanford University.

- PURCELL, T. J., BUCK, I. & MARK, W. R., 2002: Ray Tracing on Programmable Graphics Hardware. Proceedings of ACM SIGGRAPH, ACM (3), New York, 703-712.
- RENDERMAN, 2017: Pixar Renderman Pro Server. <http://renderman.pixar.com>, last access January 10, 2017.
- RODRÍGUEZ, M. B., AGUS, M., MARTON, F. & GOBBETTI, E., 2014: HuMoRS: huge models mobile rendering system. Proceedings of the Nineteenth International ACM Conference on 3D Web Technologies (Web3D '14), ACM, New York, USA.
- SAITO, T. & TAKAHASHI, T., 1990: Comprehensible rendering of 3-D shapes. In Proceedings of the 17th annual conference on Computer graphics and interactive techniques (SIGGRAPH '90). ACM, New York, NY, USA, 197-206.
- SCHAUB, M., 2014: Das Bronzmodell von Augusta Raurica: vom Stadtplan zum Stadtmodell. Augusta Raurica Magazin, **2014**(1), 7-9.
- THIBIEROZ, N., 2016: It's Time to Open up the GPU. GPUOpen. <http://gpuopen.com/welcometogpuopen/>, letzter Zugriff am 26.09.2016.
- THRANE, N. & SIMONSEN, L. O., 2005: A Comparison of Acceleration Structures for GPU Assisted Ray Tracing. Master Thesis, University of Aarhus.
- TIMONEN, V., 2013: Screen-space far-field ambient obscurance. In Proceedings of the 5th High-Performance Graphics Conference (HPG '13). ACM, New York, NY, USA, 33-43.
- ZHUKOV, S., INOES, A. & KRONIN, G. 1998: An Ambient Light Illumination Model. Rendering Techniques '98, Springer-Verlag Wien New York, G. Drettakis & N. Max (Eds.), Eurographics, 45-56.

Geoinformatik und Archäologie – GIS basierte Kulturraumstudie antiker und frühmittelalterlicher Besiedlungsmuster am mittleren Main

ARMIN VOLKMANN¹

Zusammenfassung: Seit einigen Jahren schon werden Methoden der Geoinformatik auch in den archäologischen Wissenschaften angewandt. Beispielsweise hat der Einsatz von GIS die Archäologie hinsichtlich der Verwaltung, Visualisierung und Analyse von Fachdaten seit den 1990er Jahren beflügelt. Dabei wurden die GIS meist nur als bestehende „Fremdentwicklungen“ in der Archäologie genutzt, woraus eine gewisse Stagnation in der Verwendbarkeit dieser Informationssysteme resultierte. Einhergehend mit der Entwicklung von stabilen und quelloffenen Open Source GIS entstanden in den letzten Jahren auch einige Anwendungen, die primär aus der direkten Zusammenarbeit von Geoinformatik und Archäologie resultieren und die damit zielgerichteter, anstatt reiner Adaption, auf die Bedürfnisse kultur- und geisteswissenschaftlicher Forschung eingehen. So kann gerade die sogenannte Imperfektion und Heterogenität von archäologischen Fachdaten eine für beide Disziplinen hoch interessante Herausforderung sein. Das im Folgenden vorgestellte Projekt versucht genau diesen Gedanken nachzugehen. Dies beginnt in den vorgestellten Beispielen im Aufbau eines archäologischen Informationssystems (AIS) und mündet in die systematische Analyse von groß-skaligen LiDAR-Daten hinsichtlich kulturgeschichtlicher Befunde sowie in die Rekonstruktion von Paläowegen im Netzwerk prähistorischer Siedlungen.

1 Einleitung

Im Untersuchungsgebiet am mittleren Main liegt ein vergleichsweise hoher Arbeitsstand in der Aufarbeitung archäologischer Fundstellen der Römischen Kaiserzeit bis zum Frühmittelalter, d.h. aus den ersten acht Jahrhunderten n.Chr., vor, der in einigen monographischen Publikationen seinen Niederschlag gefunden hat (z.B. HOFFMANN 2004; OBST 2012; PESCHECK 1978). Gleichzeitig sind die Bodendenkmalisten des regional zuständigen Bayerischen Landesamts für Denkmalpflege bereits digital aufbereitet verfügbar, so dass in Kombination mit den Publikationen auf eine fundierte Quellenbasis zugegriffen werden kann. Im basalen Arbeitsschritt unserer Studie wurde ein relationales SQL und XML-Datenbanksystem konzipiert und aufgebaut sowie im Zusammenspiel mit der Studierendenausbildung in der Lehre weiter entwickelt und sukzessive mit Daten gefüllt, das die Grundlage für die folgenden Analysen darstellt (s. Abb. 1). Das Anliegen war dabei die Studierenden an konkreten Problemstellungen der Forschung im Umgang mit GIS auszubilden und durch kollaboratives Arbeiten die verschiedenen Forschungswissensstände aus retrodigitalisierten Monographien und Fachdatenbanken zusammenzubringen. Dabei ergaben sich besondere Herausforderungen mit den Studierenden der Geistes- und Kulturwissenschaften, da diese meist nicht über eine fundierte Ausbildung in der Geoinformatik verfügen (siehe beispielsweise die Handbücher: KAPPAS 2012;

¹ Universität Heidelberg, Nachwuchsforschungsgruppe Digital Humanities and Cultural Heritage, "Cluster of Excellence Asia and Europe in a Global Context", Karl Jaspers Centre, Voßstr. 2, Gebäude 4400, D-69115 Heidelberg, E-Mail: armin.volkmann@asia-europe.uni-heidelberg.de

DE LANGE 2013), und somit auf eine grundlegend verschiedene Wissensbasis aufgebaut werden musste. Dabei wurde grundlegend die Handhabbarkeit verschiedener GIS-Software Anwendungen verglichen und eruiert (VOLKMANN et al. 2012) und auf diese spezifischen Bedürfnisse angepasste Tutorials, wie Beispielsweise zur Georeferenzierung von historischen Kartenwerken oder archäologische Fundstellenkartierung mit Open Source GIS-Software, erstellt (vgl. VOLKMANN 2014; 2015b).

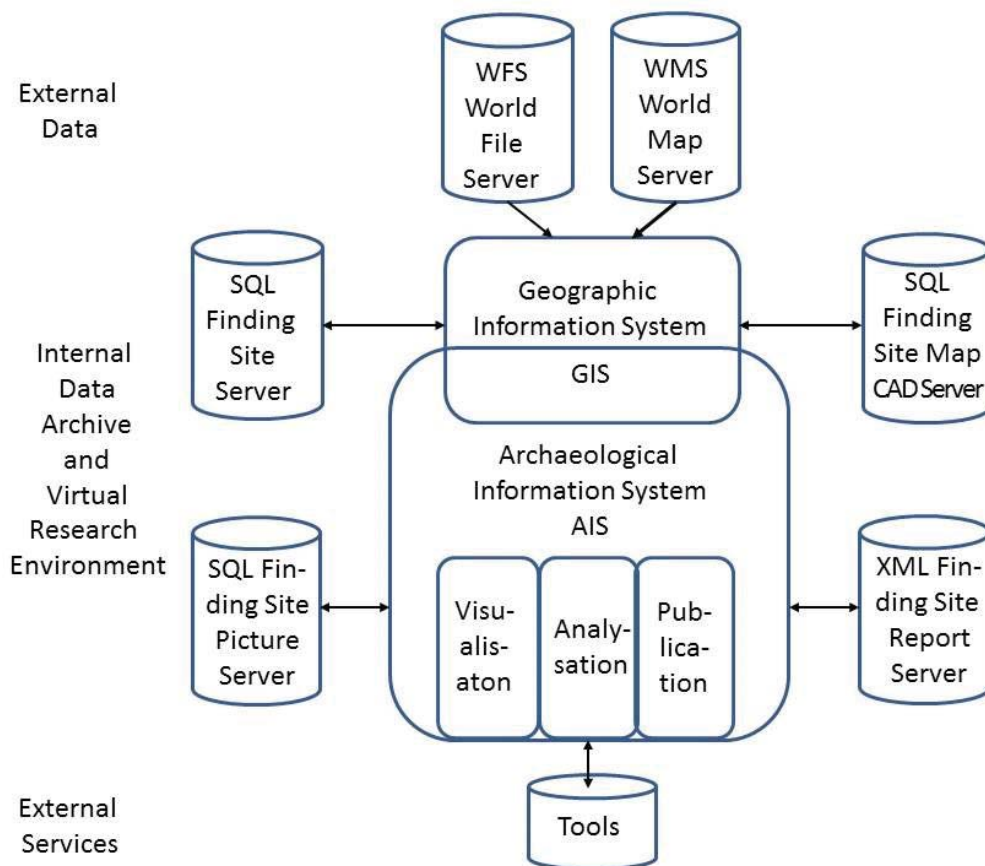


Abb. 1: Modell unseres Archäologischen Informationssystems (AIS). Das Geografische Informationssystem (GIS) ist in die virtuelle Forschungsumgebung des AIS integriert und ermöglicht den Zugriff auf externe und interne Datenbanken, d.h. via Internet und Intranet. Das AIS dient der Daten-Visualisierung, -Analyse und letztlich -Publikation heterogener archäologischer Fachdaten (Fundstellenkataster, Grabungspläne, Fotodokumentation und Grabungsberichte) unter Einbindung von weiteren (Open Source) Tools. Weiterführend zu Forschungsdaten-Infrastrukturen und virtuellen Forschungsumgebungen für die Archäologie Volkmann 2014b; Grafik Verfasser

Im analytischen Teil der Studie werden bisher kaum erforschte prähistorische Siedlungsraumkonzepte fokussiert. Dies erfolgte durch die Konzeption und den Aufbau eines archäologisch orientierten GIS (CONOLLY & LAKE 2008), dessen basierendes Datenbanksystem für den Umgang mit heterogenen Datenquellen modifiziert wurde, um neben Messdaten (Laserscanner und Totalstation) aus Feld-Surveys (vgl. VOLKMANN et al. 2016) auch mit heterogenen Text- und Bildquellen aus Retrodigitalisaten (Ausgrabungsberichte, Publikationen,

Grabungsfotografien, handgezeichneten Plänen et cetera) umgehen zu können. Ausgehend von einem GIS entstand dabei ein generisches AIS (Archäologisches Informationssystem), das diese speziellen Bedürfnisse der archäologischen Wissenschaften explizit berücksichtigt (s. Abb. 1) und grundlegend auf Open Source GIS-Software und eingebundene Hardware-Komponenten der verschiedenen Fundstellen-, Bild-, Karten- und Messdaten-Server (teils via WMS und WFS) besteht (VOLKMANN 2014b; 2016). Verwendet wurde dabei die Open Source Quantum GIS, die durch die äußerst aktive internationale Nutzerschaft mit mittlerweile hunderten sogenannten Pugins individuell anpassbar und stark erweiterbar ist (<http://www.qgis.org>). Bei der Programmierung individueller Pugins mit teils eingehender Konzeption von Datenbanken und entsprechenden Datenmodellen wurden international gebräuchliche Standards für Geodaten verwendet, die aber nicht immer oder ohne Weiteres auf geistes- und kulturwissenschaftliche Daten übernommen werden können, da diese Daten oft als imperfekte Daten mit partiell unsicheren Angaben zu Raum und Zeit vorliegen. Hieraus ergeben sich besondere Herausforderungen in der beiderseitig ergiebigen Zusammenarbeit von Geoinformatik und Archäologie, die dahingehend berücksichtigt werden müssen (VOLKMANN 2017).

2 Problemstellung

Seit den letzten Jahren haben verstärkt Methoden der Geoinformatik auch in den archäologischen Wissenschaften ihren Niederschlag in einigen hochambitionierten Projekten gefunden. Hier sind besonders die Aktivitäten der AG Computeranwendungen & quantitative Methoden in der Archäologie (CAA) hervorzuheben, die in ihren Jahrestagungen besonders das hohe Potential in der Zusammenarbeit von Geoinformatik und Archäologie deutlich machen (vgl. z.B. den Tagungsband des DEPARTMENT OF HISTORY AND CULTURAL HERITAGE 2015). In der vorliegenden Studie wurden die archäologischen Fundstellen einer intensiven Quellenkritik unterzogen, um den in den archäologischen Fachdaten weit verbreiteten Problemen der teils nicht gesicherten räumlichen und zeitlichen Eindeutigkeit effektiv zu begegnen. Dabei wurden Modelle der „Fuzzyifizierung“ bzw. der Fuzzy-Logic im AIS erprobt und angewandt, um neue Perspektiven in der archäologischen Fundstellenarchivierung, -kartierung und folgenden -analyse zu eruieren (grundlegend dazu: KAINZ 2002). Aufgrund des oft nur unscharf datierenden Fundmaterials werden darüber hinaus zeitliche Gewichtungen durch Aoristic-Analysis im AIS auf deren Handhabbarkeit im Kontext der archäologischen Fachdaten untersucht, um eben diese Fundstellen besser zeitlich zu klassifizieren (RATCLIFFE 2000). Ebenso weisen Netzwerkanalysen einen eher experimentellen Charakter auf (BRUGHMANS 2013), da neben der oft nicht gesicherten Gleichzeitigkeit der Fundstellen deren prähistorische Interaktion nur aufgrund des stark fragmentiert überlieferten Fundmaterials der einstigen archäologischen Kultur in Form von bei Ausgrabungen erkannten Befunden und Funden rekonstruiert werden kann. Als Annäherung an antike Besiedlungsmuster werden Voronoi-Diagramme aus Thiessen Polygonen im AIS für jede relevante Zeitstufe der Römischen Kaiserzeit, Völkerwanderungszeit und Merowingerzeit berechnet, die den Vergleich hypothetischer Siedlungskammern mit ihren ganz eigenen Raummustern im zeitlichen Verlauf erlauben (vgl. KNAPPETT 2013; VOLKMANN 2015a). GIS-basierte Delaunay-Triangulation für Next-Neighbour-Beziehungen der räumlichen Lage der Fundstellen zeigen mögliche Verbindungen der Siedlungen zueinander auf. Mit Least-Cost-Path-

Analysen werden plausible Paläowegverbindungen im GIS berechnet, die basierend auf der algorithmischen Auswertung des Digital Surface Modells DGM1 der Bayerischen Vermessungsverwaltung sehr detailliert rekonstruiert werden (HERZOG 2012). Kumulative Least-Cost-Path-Analysen ermöglichen die flächige Berechnung von Distanzzonen bzw. Erreichbarkeitszonen, die als gewichtete Information aus dem GIS in die Datenbank zurückfließen und folgend tiefergehend untersucht werden (siehe VOLKMANN im Druck-1). Viewshed-Analysen ermöglichen darüber hinaus Sichtbarkeitsmerkmale der Siedlungen zueinander adäquat zu erforschen. Anknüpfend an vorherige geoarchäologische Studien in der Oderregion werden die Methoden der Site Catchment-, bzw. der Site Location-Analysen methodisch weiterentwickelt, um diese vergleichend auch am mittleren Main mit strukturell abweichenden Datensätzen (wie thematischen Kartenwerken zum Bodentyp et cetera) und geökologischen Voraussetzungen anzuwenden (vgl. VOLKMANN im Druck 2).

2.1 Fallbeispiel I: Forschungsdaten-Gateway zur intelligenten Visualisierung und Auswertung groß-skaliger LiDAR-Daten für die Archäologie

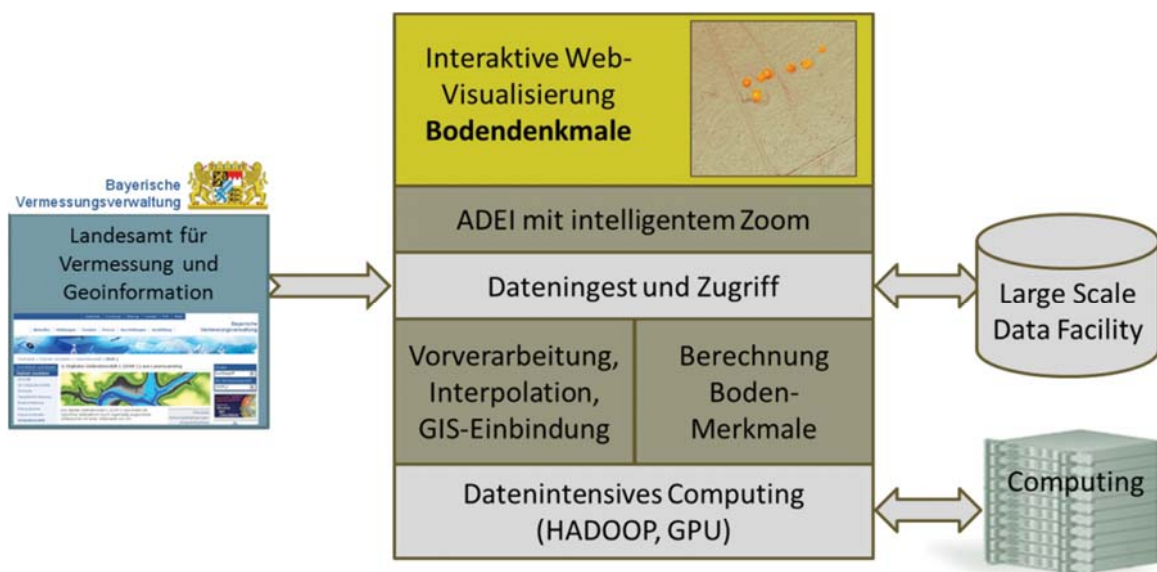


Abb. 2: Konzeption unserer Forschungsdaten-Gateway-Infrastruktur in Form des interaktiven Web-GIS zur intelligenten Visualisierung, Auswertung und Informationsextraktion groß-skaliger LiDAR-Daten für die Archäologie am Karlsruher Institut für Technologie (KIT) mit der Large Scale Data Management Group; ADEI Advanced Data Extraction Infrastructure; Apache HADOOP Software Framework; GPU Graphics Processing Unit; Grafik KIT.

Gleichzeitig ergeben sich aber neben den oben genannten Chancen auch zahlreiche Problemstellungen im praktischen Umgang beispielsweise mit Daten-LiDAR durch deren beachtliche Datengröße. Die uns im Rahmen der Studie für den gesamten Regierungsbezirk Unterfranken flächendeckend für fast 9000km² vorliegenden Airborne Laserscandaten weisen insgesamt eine Speichergröße von fast 2,5 Terabyte auf (VOLKMANN & RAUN 2015). So konnten

neben dem Bayerischen Landesamt für Denkmalpflege, das Bayerische Landesamt für Umwelt und Landesamt für Vermessung und Geoinformation (LVG) als Kooperationspartner des Projektes gewonnen werden, wobei letzteres uns die LiDAR-Daten durch eine Kooperationsvereinbarung bereitstellen konnte. Die LiDAR-Daten des LVG Bayern bestehen aus den ungefilterten Rohdaten des First Pulse mit Vegetation und Last Pulse ohne Vegetation sowie dem gefilterten DGM1 mit einer 1-Meter-Auflösung (bzw. Messgitterweite). So liegen uns im Rahmen der Untersuchung insgesamt zirka 79.000 einzelnen Dateien vor, die jeweils meist 1km² Geländeabdeckung aufweisen.

Für die Untersuchungsfläche mussten die LiDAR-Datensätze der Dateien interpoliert, georeferenziert in ein geografisches Informationssystem (GIS) eingebunden und ausgewertet werden. Die reine Vorverarbeitung der Interpolation der DGM1-Daten zu GeoTiff-Geländemodellen (mit dem Skript OPALS) benötigte eine manuelle Bearbeitungszeit von ca. 60 Minuten pro 10km², sodass ca. fast 113 Arbeitstage (a 8h Arbeitszeit) reine Rechenzeit alleine für den Bereich Unterfrankens erforderlich waren. Wenn man dies auf die Fläche Deutschlands mit einer Datengröße des Gesamt-DGM1 von ca.100 TB hochrechnet, wären ca. 4500 Arbeitstage reine Interpolationsrechenarbeit eines Mitarbeiters erforderlich, um die Geländemodelle zu erstellen. Moderne handelsübliche Computer können Datensätze bis zu einer Größe von einem Gigabyte noch effektiv in einer annehmbaren Berechnungszeit kalkulieren. Dies entspricht dabei einer Fläche von zirka 100km² mit 100 GeoTiffs im Arbeitsspeicher unseres AIS, im dem diese optisch nach anthropogenen Strukturen, wie Hügelgräber, Wege oder Burgwallbefunde, gesichtet werden. Dies ist manuell nicht mehr zu bewältigen, da es zum einen zu zeitintensiv ist, aber zum anderen die qualitativen Ergebnisse sehr stark vom auswertenden Individuum abhängen, das selektiv (mehr oder weniger systematisch und standardisiert) Bodendenkmale erkennt oder eben nicht erkennt. Die Auswertung erfolgt daher standardisiert anhand eines Vergleichskataloges typischer Bodendenkmale und diskriminierender Merkmalsausprägungen moderner Strukturen, wie Bebauung und Straßen, manuell und benötigt aufgrund der Komplexität zusätzlich ca. einen Bearbeitungstag pro 100km², sodass nun an semi-automatisierten Ansätzen zur effektiveren Eingrenzung archäologischer Verdachtsflächen anhand der LiDAR-Befunde gearbeitet wird. Dazu werden Technologien für ein interaktives Forschungsdaten-Gateway entwickelt, mit dem groß-skalige 3D-Daten in Echtzeit prozessiert, mit Daten aus anderen Quellen fusioniert und intelligent visualisiert werden können. Angewandt auf Laserscandaten sollen damit innerhalb der Landschaftsarchäologie besonders in Waldgebieten bisher unzugängliche und unbekannte Bodendenkmale hocheffektiv entdeckt werden. Die fokussierten Technologien (s. Abb. 2) können Teile der Vorverarbeitung und der Merkmalsextraktion direkt bei der Einspeisung der Daten in das Archiv-Gateway automatisch auf einem HADOOP-Cluster durchführen. Beschleunigungen um den Faktor 100, gegenüber der herkömmlichen Bearbeitung, sind auf einem solchen System realistisch. Die Auswertung der Daten erfolgt interaktiv durch den archäologischen Bearbeiter, wobei aber aufgrund der intelligenten Visualisierung interessante Stellen auch bei sehr großen Skalen in Echtzeit dargestellt werden können. So können Gebiete von mehreren Hundert km² pro Tag erschlossen und mit anderen Karten im AIS abgeglichen werden. Hierzu wird innerhalb der Large Scale Data Management Group am Karlsruher Institut für Technologie (KIT) an genau solchen Lösungsansätzen gearbeitet, die auf einem interaktiven Web-GIS beruhen, wobei dafür bereits

bestehende Open Source Technologien (wie die Javascript Libraries Leaflet und OpenLayers) auf unsere Bedürfnisse zielgerichtet angepasst werden.

2.2 Fallbeispiel II: Siedlungsstrukturen und Paläowegverbindungen in den Jahrzehnten um die Zeitenwende – Landschaftsarchäologie im AIS

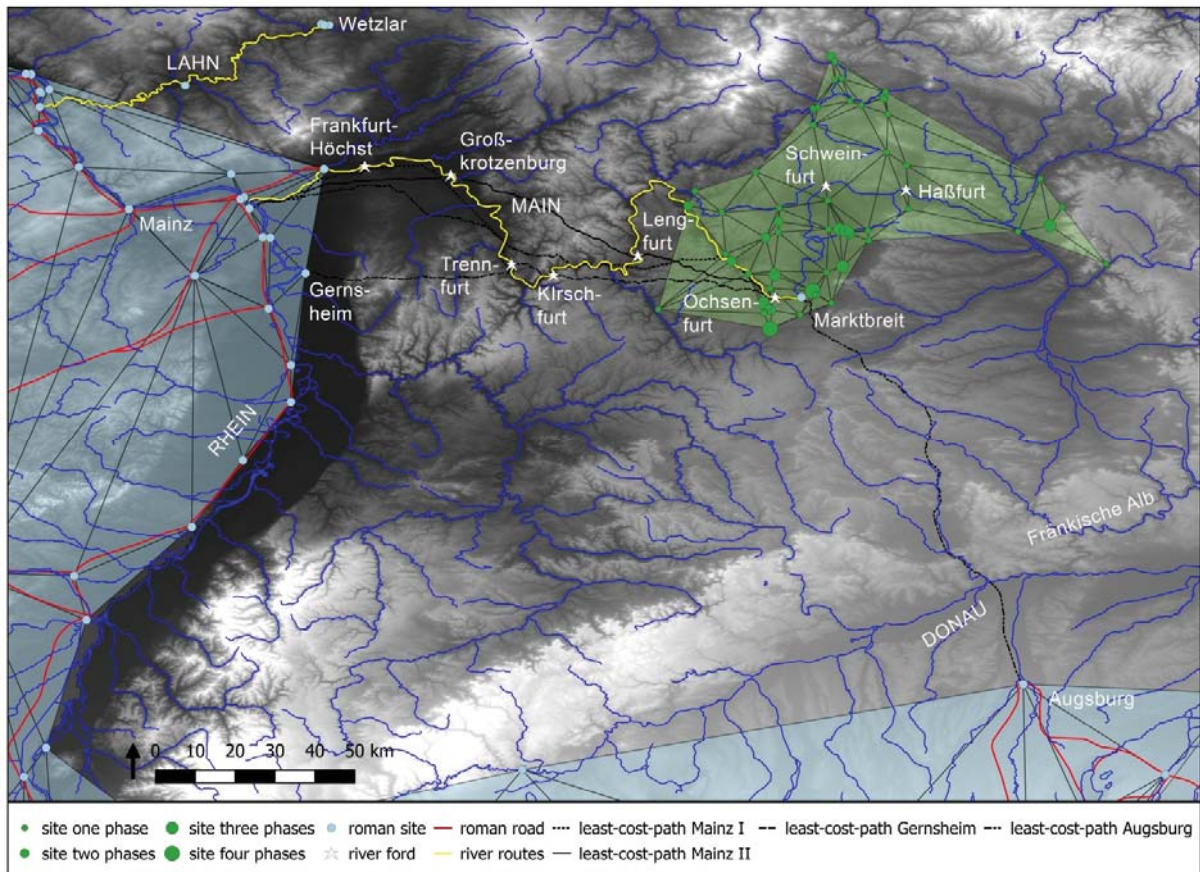


Abb. 3: Untersuchungsregion am mittleren Main zur Zeitenwende (in den Jahren 15 v. Chr. bis 70 n. Chr.). Die germanischen Fundstellen im Barbaricum außerhalb des Limes werden nach Länge der Besiedlungsandauer als entsprechend groß gewichtete grüne Punkte dargestellt, die wahrscheinlich der unterschiedlichen Bedeutung der Siedlungsplätze entsprechen. Im Westen und Süden bestanden im Römischen Reich schon einige gut ausgebaute Straßenverbindungen, die im Barbaricum in dieser Form nicht vorhanden waren. Mit gelben Linien sind Exkursionsrouten der Römer ins Barbaricum auf den Flüssen Lahn und Main hervorgehoben. Vier potentielle Landrouten zum römischen Kastell von Marktbreit im inneren Barbaricum sind mit schwarzen Linien des Weiteren hervorgehoben; Grafik Verfasser

Ausgangslage war hierbei die seit längerem bestehende Fragestellung, ob das römische Kastell von Marktbreit, das weit außerhalb des römischen Limes in mitten einer germanischen Siedlungskammer zielgerichtet angelegt wurde, von Süden her vom römischen Augsburg (Augusta Vindelicorum) oder von Westen her vom römischen Mainz (Mogontiacum) erschlossen wurde. Diesem Desiderat wurden systematisch nachgegangen, indem im AIS plausible Erschließungsrouten kalkuliert und verglichen wurden. Dazu wurden auf dem digitalen

Geländemodell DGM1 im GIS Least-Cost-Path Analysen durch die Einbindung von bestehenden GASS-GIS Algorithmen durchgeführt, die in der Anwendung trainiert werden konnten und deren standardisierter Arbeitsablauf an spezifische Bedürfnisse, wie maximale Steigung, Geschwindigkeit, Passierbarkeit von Gewässern, bzw. Barrieren u.a., anpassbar ist (<https://grass.osgeo.org>).

Die germanische Siedlungskammer am mittleren Main, in deren Zentrum im frühen 1. Jahrhundert das römische Kastell von Marktbreit kurzfristig errichtet wurde, wird durch die Höhenzüge der Mittelgebirge im Norden begrenzt. In der zugehörigen Karte des Höhenreliefs sind diese als helle Bereiche dargestellt (s. Abb. 3), und es wird deutlich, dass sie keinerlei Besiedlung aufwiesen. Die frühgermanischen Siedlungen liegen ausnahmslos in der Niederungszone des Flusslaufes des Mains mit einmündenden Bächen in den Seitentälern. Im frühen 1. Jahrhundert n.Chr. befinden sich der römische Einflussbereich am mittleren Rhein und die germanische Siedlungskammer am mittleren Main über 100km weit voneinander entfernt, wobei beide durch eine unbewohnte Zone, die Berge des Spessarts, getrennt werden. Dieser unbewohnte Grenzsäum stellte durch die Höhenlage und für den Ackerbau ungünstige geökologische Bedingungen eine natürliche Barriere dar.

Der Verlauf bedeutender Fernhandelsrouten konnte innerhalb des Römischen Reiches aus den Angaben der schriftlichen Quellen im Zusammenspiel mit den archäologischen Befunden von Römerstraßen rekonstruiert werden, wenn auch hier oft nicht die exakten Routen, sondern nur die jeweiligen Stationen bzw. Etappen oder Ziele erfasst wurden (vgl. Digital Atlas of Roman and Medieval Civilisation online). Außerhalb des Römischen Reiches ist dies sehr viel schwieriger, da die schriftlichen Quellen, wenn überhaupt, nur sehr spärliche Angaben zum Verlauf der Routen enthalten und entsprechende archäologische Befunde fehlen. Im Rahmen der Least-Cost-Path-Analysen wurden die Schwierigkeitsgrade für den Reisenden bei verschiedenen Wegführungen grundlegend aufgrund der geomorphologischen Geländeoberfläche verglichen. Dies erfolgte anhand des ins AIS eingebundenen, hochauflösenden digitalen Geländemodells (DGM1), an dem zahlreiche hypothetische Verbindungen berechnet wurden, um plausible Wege der Römer im Zuge der Anlage des Kastells von Marktbreit aufzeigen zu können. Das Legionslager wurde weit östlich und nördlich außerhalb des Römischen Reiches während eines Feldzuges gegen die noch weiter östlich ansässigen Markomannen im Barbaricum errichtet und war nur für sehr kurze Zeit, in den Jahren 5 bis 9 n. Chr., in Nutzung, wie die archäologischen Befunde verdeutlichen. Mit einer gelben Farbsignatur ist in der Karte (Abb. 3) als Wasserroute der mittlere Main markiert, der den Römern, ebenso wie die schwarz markierten Landwege, als potenzieller Weg zur Anlage des Legionslagers von Marktbreit gedient haben könnte. Es ist davon auszugehen, dass der mittlere Main (Moenus) schon im frühen 1. Jahrhundert n.Chr. von mittelgroßen römischen Schiffen genutzt wurde, die nachweislich auf dem Rhein und oberen Main verkehrten. Laut den schriftlichen Quellen, wie z.B. Inschriften auf steinernen Bauelementen, wurde am Rhein die römische Schiffsflotte *classis Germanica* um 13 v.Chr. aufgestellt, um einerseits den Rhein als römische Außengrenze zu sichern und um andererseits von dort aus Expeditionen ins Barbaricum östlich des Rheins zu unternehmen (s. Epigraphische Datenbank Heidelberg der Akademie der Wissenschaften online).

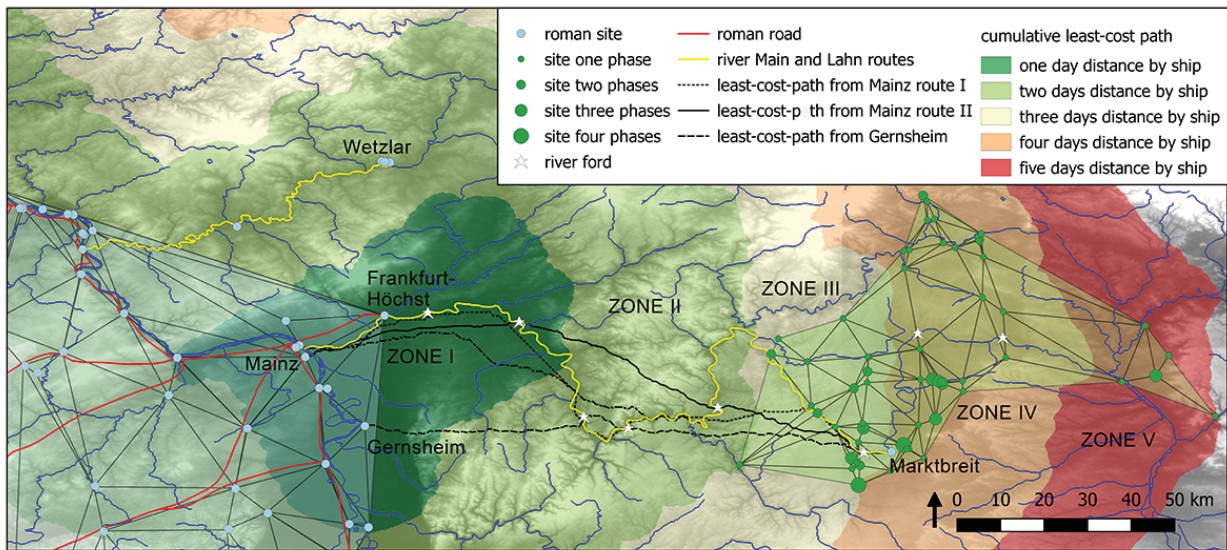


Abb. 4: Roman point of view – Ausgehend vom Auxiliarkastell von Frankfurt-Höchst, das zum Teil mit germanischen Hilfstruppen besetzt war, lag Marktbreit gerade noch in der mittel-gut zu erreichenden Distanzzone III. Die hier in der Abbildung gezeigten fünf Distanzzonen stellen Erreichbarkeitsmodelle dar, die im AIS (GRASS-GIS-Modul) mit kumulativen Least-Cost-Path-Analysen vom Ausgangspunkt Frankfurt-Höchst flächig berechnet und visualisiert wurden (vgl. VERHAGEN et al. 2012). Die Least-Cost-Path-Analysen basieren auf der algorithmischen Auswertung des Digital Surface Modells (DGM1) mit Hangsteigungswerten oder des Netzwerks der größeren Fließgewässer, die ähnlich wie in römischer Zeit existierten, und sie dienen der Berechnung von zusammengefassten Zonen potentiell guter (grün) bis schlechter (rot) Erreichbarkeiten, bzw. geringem (grün) bis hohem (rot) Aufwand in der Erreichbarkeit, wie im gezeigten Fall per Schiff. Sehr interessant ist, dass, bis auf die Ausnahme einer germanischen Fundstelle, alle anderen in den Erreichbarkeitszonen III–V liegen. D.h. aus der Sichtweise Roms waren die nächstgelegenen germanischen Siedlungen am Main potentiell nicht einfach zu erreichen, wobei zielgerichtet das Kastell von Marktbreit unter den Aspekten bestmöglicher Verkehrsanbindung und gleichzeitiger zentraler Lage in der germanischen Siedlungskammer errichtet wurde; Grafik Verfasser

Der Landweg als potenzielle Erschließungsrouten ins Innere des Barbaricums könnte durchaus von großer Bedeutung gewesen sein, da der Main durch geologisch gegebene Furten auch ohne Brücken- oder Floßbau bei temporären Niedrigwasserständen (oder bei Vereisung) leicht zu überqueren war. Interessanterweise haben sich die Hinweise auf natürliche Mainfurten noch in den heute dort existierenden Ortsbezeichnungen erhalten, die – mit Ausnahme von Großkrotzenburg – alle mit dem Suffix „-furt“ enden (s. Abb. 3). Als Ausgangsort für eine römische Expedition ins Barbaricum kommt neben Frankfurt-Höchst auch das bedeutende Legionslager von Mainz (Mogontiacum) mit dem zugehörigen Brückenkopfkastell von Mainz-Kastel (Castellum Mattiacorum = Kastell im Land der Mattiaker) östlich des Rheins in Betracht. Das römische Mogontiacum, das unter Drusus seit 13/12 v. Chr. angelegt worden war, erlangte später Bedeutung als Provinzhauptstadt mit Sitz der Verwaltung und städtischer Infrastruktur. Es wies einige herausragende Monumentalbauten auf und ist auch als ein wichtiger Militärstandort am Rhein zu identifizieren (vgl. ARCHÄOLOGISCHES LANDESMUSEUM BADEN-WÜRTTEMBERG

2005). So bildete Mogontiacum das Tor zum Barbaricum für Schiffstouren auf dem Rhein und Main und ebenso für Landexkursionen.

Eine zweite nördliche Route über Land nach Marktbreit führte an Frankfurt-Höchst im Flachland des niederen Mains vorbei und knüpfte im Westen an eine bereits existierende Römerstraße an. Das dortige Flachland endet östlich erst in der Region um Großkrotzenburg, wo eine natürliche Flussfurt das Überqueren des Mains erleichterte. Erst zwischen 105 und 110 n.Chr. entstand dort ein Kastell, das Garnisonsstandort der *cohors IIII Vindelicorum* war (vgl. Deutsche-Limeskommission online). Spätestens ab 134 n.Chr., zum Ende der Regierungszeit Kaiser Hadrians, wurde bei dem Kastell eine steinerne, permanente Mainbrücke errichtet, wie dendrochronologische Untersuchungen der archäologisch dokumentierten Fundamentpfähle aus dicken Holzstämmen belegen (vgl. Museen-Mainlimes online). Das Kastell von Großkrotzenburg mit angeschlossenem vicus lag auf der östlichen Mainseite genau gegenüber dem kleinen Numerus-Kastell von Hainstadt am westlichen Ufer, das wohl schon im 1. Jahrhundert n.Chr. erbaut worden war (vgl. GREINER 2005).

Östlich des Mains beginnt dort das hügelige Waldland des Spessarts, das durch seine beachtlichen Höhenlagen und das raue Klima bis in das Spätmittelalter hinein nahezu unbewohnt blieb. Im Bereich des Spessarts sind im hochauflösenden digitalen Geländemodell zahlreiche Altwegesysteme erkennbar, die durch die Bewaldung bzw. den nicht vorhandenen Ackerbau und infolgedessen geringere Bodenerosion erhalten blieben (s. Abb. 5). Die Landroute von Mainz nach Marktbreit führt nach dem Höhenabstieg des Spessarts hin zur natürlichen Mainfurt beim heutigen Ort Lengfurt. Östlich davon ist im Anschluss das Relief weit weniger steil und nur noch hügelig, was geologisch durch den dortigen verwitterten Muschelkalk bedingt ist. Der restliche Weg nach Osten bis nach Marktbreit konnte im südlichen Tal des Mains recht angenehm ohne große Steigungen zurückgelegt werden. Über Land beträgt der Weg ausgehend von Mainz bis nach Marktbreit insgesamt 168km, was ca. fünf Tagesmärschen entspricht, da die römische Infanterie oder Reiterei pro Tag bis zu 30–40km problemlos zurücklegen konnte.

Eine alternative Landroute stellt der Weg ausgehend von Gernsheim dar, das am östlichsten Punkt der Grenze des Römischen Reiches in der Region liegt (s. Abb. 4). In Gernsheim bestand am östlichen Rheinufer vermutlich ein Truppen-Standort mit Hafen, der, wie Funde belegen, ab ca. 70 n.Chr. militärisch genutzt wurde (s. HÜSSEN 2000). Der Landweg nach Marktbreit betrug von hier aus nur 148km, die ca. vier Tagesmärschen entsprechen. Bei dieser Route mussten zudem nicht die Höhenzüge des Spessarts überwunden werden. Bei Trennfurt und Kirschfurt konnte der Main in den dortigen natürlichen Furten durchquert werden. Alternativ ließ sich, wenn man auf diese Abkürzung verzichtete und einen Umweg von ca. 5km in Kauf nahm, Marktbreit auch ganz ohne die gefährliche Maindurchquerung erreichen, indem zwischen Trennfurt und Kirschfurt der flache Auenbereich am südlichen Mainufer als natürlich vorgegebene Route genutzt wurde (vgl. Abb. 3). So ist der Landweg von Gernsheim nach Marktbreit eindeutig die bessere und schnellere Route.

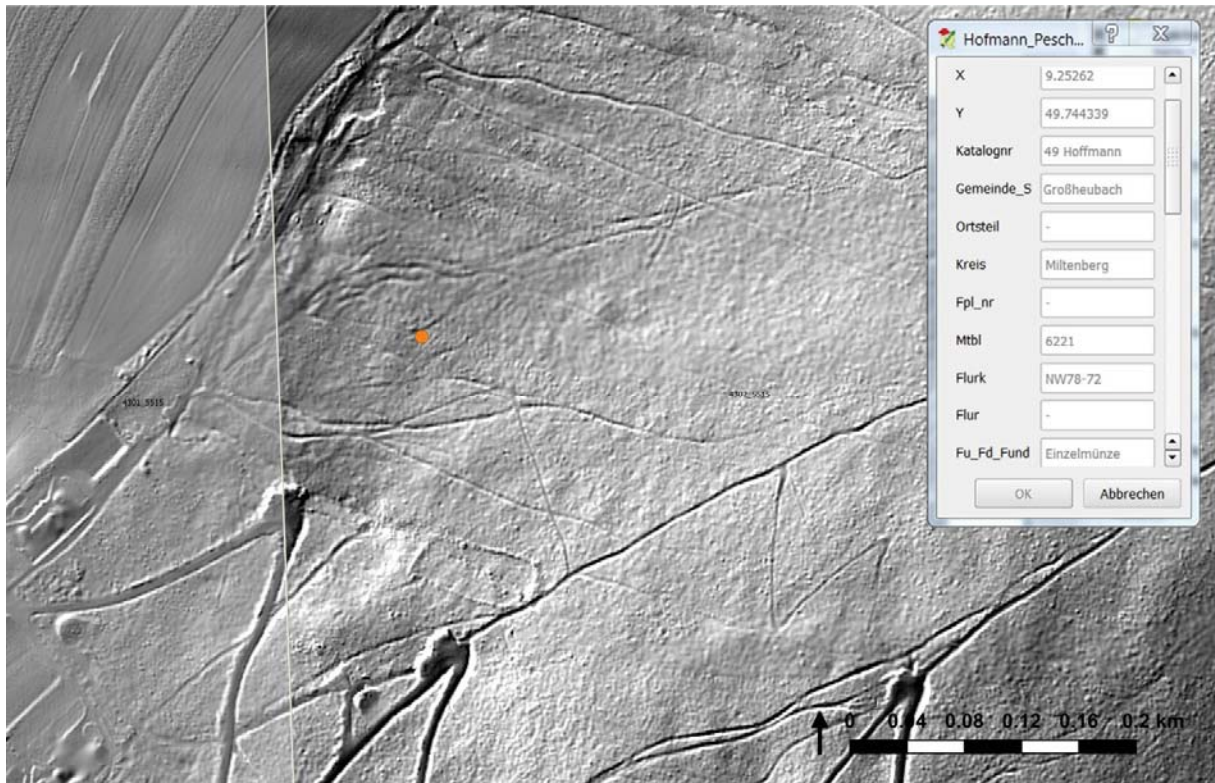


Abb. 5: Hohlwege in einem multiplen Pfadsystem mit temporär genutzten Wegabschnitten, die vom flachen Ackerland (links) zum bewaldeten Plateau (rechts) aufsteigen und deren Datierung allgemein vom Hochmittelalter bis in die Eisenzeit anzusetzen ist. Besonders in Waldgebieten haben sich durch die schützende Baumvegetation, bzw. den nicht vorhanden und durch den Pflug zerstörend wirkenden Ackerbau, neben Altwegesystemen auch vielerlei weitere Bodendenkmale, wie z.B. Hügelgräber, erhalten. Anhand des Fundes einer römischen Münze der späten römischen Kaiserzeit (roter Punkt) erscheint eine entsprechende Feindatierung dieses Wegabschnittes ins späte 3. bis 4. Jahrhundert n.Chr. durchaus realistisch zu sein (ganz rechts oben Fundstelleninformation aus der mit dem GIS verknüpften Datenbank). Folgende geoarchäologische Feldstudien sollen dies verifizieren und die Mehrphasigkeit des multiplen Pfadsystems detailliert untersuchen; Digitales Geländemodell DGM1 des Bayerischen Landesamts für Vermessung und Geoinformation; Grafik Verfasser.

Denkbar ist auch eine Südroute nach Marktbreit, die vom Legionslager Augusta Vindelicorum, der späteren Hauptstadt der römischen Provinz Raetien, im heutigen Augsburg und vom nördlich davon gelegenen Kastell Submuntorium (bei Burghöfe a.d. Donau) ihren Ausgang nahm (s. Abb. 3). Dieses lag unmittelbar südlich der Donau zwar in nur ca. 124 km Entfernung per Luftlinie zu Marktbreit, jedoch musste auf diesem Weg das Nördlinger Ries passiert werden, was mit einer großen Anzahl von zurückzulegenden Höhenmetern einherging. Die Wegstrecke dieser Route Burghöfe-Marktbreit beträgt 142 km (bzw. 157 km bei geringster Steigung) und ist damit interessanterweise ähnlich lang, wie die Route Gernsheim-Marktbreit mit 146 km (bzw. 157 km bei geringster Steigung). Der große Unterschied ist aber ausgehend vom Kastell Burghöfe die zu bewältigende Reliefamplitude und damit einhergehenden, zurückzulegenden Höhenmeter, die durch die Passage über die Ausläufer der Schwäbisch-Fränkischen Alb und Querung des Nördlinger Ries, mit dem steilem Aufstieg aus dem Ries, und weiter nördlich über die

Frankenhöhe bis an den Main bei Marktbreit entstehen. So ist die Summe der zu absolvierenden Höhenmeter wesentlich größer als bei der Route Gernsheim-Marktbreit, die über das Fränkische Schichtstufenland östlich des Odenwalds auf den Höhenzügen parallel im weiteren Umfeld zum Main verläuft. Innerhalb der Strecke von Gernsheim müssen insgesamt 2360 Höhenmeter zurückgelegt werden, wohingegen bei der Strecke von Burghöfe mehr als 3110 Höhenmeter zu absolvieren sind. Daher erscheint die theoretisch mögliche Südroute ausgehend von Burghöfe im römischen Grenzland an der Donau weniger praktikabel gewesen zu sein als die Westroute vom Grenzland im Rhein-Maingebiet.

Neben diesen Landrouten waren aber auch die größeren Flüsse, wie der Main, für die Römer beliebte Verkehrswege: Ausgehend vom kleinen römischen Auxiliarkastell von Frankfurt-Höchst am Main, das bereits um die Zeitenwende datiert, konnte das via Mainroute ca. 275km entfernt liegende Marktbreit per Schiff in ca. drei Tagen erreicht werden (s. Abb. 4). Mit römischen Schiffen konnten mehr als 100km pro Tag zurückgelegt werden, wie zahlreiche auf archäologischen Funden basierende Rekonstruktionen von Römerschiffen mit experimentellen Schifffahrtstouren auf mitteleuropäischen kleineren und größeren Flüssen eindrucksvoll belegen (Römisch-Germanisches Zentral Museum RGZM/ Museum für Antike Schifffahrt Mainz online). Mit hoher Wahrscheinlichkeit ist von Mainz oder vom Kastell Frankfurt-Höchst aus die dreitägige Schiffsreise auf dem Main bevorzugt worden, da von dort aus eine lange, vier bis fünf Tage dauernde und zudem anstrengend steile sowie durch die Mainfurten gefährliche Landroute an sich weniger realistisch erscheint. Jedoch hatte die Schiffsroute den Nachteil, dass ein Schiff auf dem mittleren Main von den steilen Hängen aus schon von weitem gesehen werden konnte und somit ein leichtes Ziel für mögliche Angriffe darstellte. Daher kann die längere Landroute letztlich für exploratorische Militäraktionen der Römer ins Barbaricum durchaus von Vorteil gewesen sein.

3 Fazit & Ausblick

Als Adaption eines GIS wurde ein archäologisches Informationssystem (AIS) konzipiert und umgesetzt, das sehr heterogene Geo- und Archäo-Daten kontextualisiert und auf die Bedürfnisse der fachspezifischen Anwendung und Analyse ausgerichtet ist (vgl. VOLKMANN 2016). Durch die Zusammenbringung der zuvor sehr verstreut vorliegenden Informationen konnte ein ganz neuer Wissenstand zur Besiedlung in der Antike und im Frühmittelalter am mittleren Main generiert werden, der perspektivisch durch die strikt generische Konzeption sowohl zeitlich als auch räumlich erweitert werden kann. Im Rahmen der multiplen Datenanalyse werden Verfahren der Geoinformatik an den aggregierten archäologischen Fachdaten angewandt. Dabei entstehen u.a. neuartige Besiedlungs-Raummodelle, die wiederum die Ausgangslage für weitere Untersuchungen darstellen. In den vorgestellten Analyseschritten dienen diese Raummodelle der Annäherung bzw. der Rekonstruktion von prähistorischen Kulturlandschaftsmodellen mit plausiblen Paläowegverbindungen sowie dem grundlegenden Verständnis der Konzeption des prähistorischen Raums, was nicht nur in der Untersuchungsregion ein seit langem bestehendes Desiderat darstellte. Diese Forschungslücke ist dabei besonders geeignet mit geoinformatischen Methoden neu hinterfragt zu werden, wie es in unserer Untersuchung erfolgt. So können Theorien zur Zentralörtlichkeit (NAKOINZ 2009) und zu Netzwerkstrukturen (KNAPPETT 2013)

durch die Implementierung entsprechender Module direkt im AIS überprüft und analysiert werden, das somit nicht nur der Datenverwaltung und Visualisierung dient, sondern darüber hinaus der Generierung von grundlegend neuem Wissen dienen kann.

Die Studie ist in die Arbeiten der Nachwuchsforschungsgruppe „Digital Humanities and Cultural Heritage with focus on Archaeoinformatics“ am „Cluster of Excellence Asia and Europe in a Global Context“ an der Universität Heidelberg in Form von zwei Dissertation und einer Habilitation eingebunden. Zusammenarbeiten erfolgen dabei mit dem Karlsruher Institut für Technologie (KIT) im Rahmen der Large Scale Data Management Group und mit der Geoinformatik der Universität Heidelberg.

Der Verfasser möchte an dieser Stelle ganz ausdrücklich den Kooperationspartnern: Bayerisches Landesamt für Denkmalpflege, Bayerisches Landesamt für Umwelt, Landesamt für Vermessung und Geoinformation für die großzügige Datenbereitstellung danken, ohne diese die vorliegende Studie nicht möglich gewesen wäre.

4 Literaturverzeichnis

- ARCHÄOLOGISCHES LANDESMUSEUM BADEN-WÜRTTEMBERG (Hrsg.): Imperum Romanum, Roms Provinzen an Neckar, Rhein und Donau. Theis Verlag.
- BRUGHMANS, T., 2013: Thinking through Networks: A Review of Formal. Network Methods in Archaeology, *Journal of Archaeological Method and Theory* 12/2013, **20**(4), 623-662.
- CONOLLY, J., & LAKE, M. 2008: Geographical Information Systems in Archaeology. Cambridge University.
- DEPARTMENT OF HISTORY AND CULTURAL HERITAGE (Hrsg.), 2015: 43rd Computer Applications and Quantitative Methods in Archaeology Annual Conference. Book of Abstracts. University of Siena 2015.
- GREINER, B.-A., 2005: Die Kastellvici am vorderen Limes, ARCHÄOLOGISCHES LANDESMUSEUM BADEN-WÜRTTEMBERG (Hrsg.) Imperum Romanum. Roms Provinzen an Neckar, Rhein und Donau, Theis Verlag, 165-168.
- HERZOG, I., 2012: Least-cost networks, *Archaeology in the Digital Era. Computer Applications and Quantitative Methods in Archaeology 2012*, EARL, G., SLY, T., CHRYSANTHI, A. & MURRIETA-FLORES, P. et al. (Hrsg.), 237-249, <http://dare.uva.nl/aup/nl/record/468863>.
- HOFFMANN, K., 2004: Kleinfunde Der Römischen Kaiserzeit aus Unterfranken. Studien Zur Siedlungsgeschichte und kulturellen Beziehung zwischen Germanen und Römern. *Internationale Archäologie* **80**, Verlag VML.
- HÜSSEN, C.-M., 2000: Sicherung der Reichsgrenze. Die Römer zwischen Alpen und Nordmeer, WAMSER, L. (Hrsg.), Verlag Phillip von Zabern, 58-63.
- KNAPPE, C. (Hrsg.), 2013: Network Analysis in Archaeology. New Approaches to Regional Interaction, Oxford University Press.
- OBST, R., 2012: Die Besiedlungsgeschichte am nordwestlichen Maindreieck vom Neolithikum bis zum Ende des Mittelalters. Würzburger Arbeiten zur Prähistorischen Archäologie **4**, Verlag Marie Leidorf.
- PESCHECK, Chr., 1978: Die germanischen Bodenfunde der römischen Kaiserzeit in Mainfranken, C.H. Beck Verlag 1978.

- KAINZ, W., 2002: Fuzzy Logic and GIS. Department of Geography and Regional Research, Universität Wien 2002.
- KAPPAS, M. 2012: Geographische Informationssysteme. Westermann Verlag.
- DE LANGE, N. 2013: Geoinformatik in Theorie und Praxis. Springer-Verlag.
- NAKOINZ, O., 2009: Zentralortforschung und zentralörtliche Theorie. Archäologisches Korrespondenzblatt **39**, 361-380.
- RATCLIFFE, J. H., 2000: Aoristic analysis: the spatial interpretation of unspecific temporal events, *International Journal Geographical Information Science* **14**(7), 669-679.
- VERHAGEN, P., BRUGHMANS, T., NUNINGER, L. & BERTONCELLO, F., 2012: The Long and Winding Road: Combining Least Cost Paths and Network Analysis Techniques for Settlement Location Analysis and Predictive Modelling, CAA2012 Proceedings of the 40th Conference in Computer Applications and Quantitative Methods in Archaeology, Southampton, United Kingdom, 26-30 March 2012.
- VOLKMANN, A., BOCK, S., SEIBT, D., KÜMMET, S., WEIß, M., DIETZ, E., HUSS, P., HEER, A. & EL HASSAN, N., 2012: Geisteswissenschaft und Geografische Informationssysteme (GIS): Erstellung von Kartierungen mit kommerzieller und Open Source Software im Vergleich, A. Volkmann (Hrsg.) Digital Humanities, Universität Würzburg, 2012, 1-116, DOI: urn:nbn:de:bvb:20-opus-74470
- VOLKMANN, A., PFEIFFER, M. & RAUN, K.-H. M., 2016: "Digital mapping" - Detection and prospection through digital and physical landscapes at Koumasa, Crete. Heidelberg Dokumentenserver, Universität Heidelberg, 1-24, DOI: urn:nbn:de:bsz:16-heidok-219489
- VOLKMANN, A. & RAUN, K.-H. M., 2015: Semi-automated detection of ground monuments in airborne laser scan data (LiDAR). LSDAM (Large Scale Data Management) Borschüre 2015, 1-3.
- VOLKMANN, A., 2014a: Tutorial Geo-referencing of historical maps with QGIS - Georeferenzierung von historischen Karten mit QGIS, Ein Bilderbuch, Wissenschaftsblog archäologiedigitale, ISSN 2197-7453, 2014, 11 Seiten, <http://archdigi.hypotheses.org/420>.
- VOLKMANN, A., 2014b: Beyond administrative boundaries: archaeological networks of research, *Cultural Heritage and new Technologies* **18**, 1-9, <http://www.chnt.at/chnt-18-proceedings/>.
- VOLKMANN, A., 2015a: Climate change, environment and migration, A GIS-based study of the Roman Iron Age until the Early Middle Ages in the river Oder region, *PCA Post-Classical Archaeologies* **5**, 2015, 69-94, <http://www.postclassical.it/vol.5.html>
- VOLKMANN, A., 2015b: Archäologische Fundkartierung und Zeichnungs-Georeferenzierung mit QGIS, Seminar Digital Humanities and Digital Cultural Heritage, Heidelberg Dokumentenserver Heidelberg 2015, 1-47, DOI: urn:nbn:de:bsz:16-heidok-199295
- VOLKMANN, A., 2016: An Archaeological Information System (AIS) for Digital Cultural Heritage management and research: identify settlement dynamics of the Late Roman Iron Age until the Early Middle Ages in the river Oder region, *Sektion für Slawische Archäologie beim 8. Deutschen Archäologie-Kongress 2014 in Berlin, Beiträge zur Ur- und Frühgeschichte Europas*, Verlag Beier & Beran, 2016, 83-99.
- VOLKMANN, A., 2017: Geografische Informationssysteme in den Digital Humanities, JANNIDIS, F., KOHLE, H., REHBEIN, M. (Hrsg.), Digital Humanities, Eine Einführung, Verlag C.B. Metzler, 2017, 300-315.

- VOLKMANN, A., im Druck-1: Perspectives for Network Analyses, Roman roads, ‘Barbarian’ paths and settlement patterns in the borderland at the Limes Germanicus in river Main Region, 21st Annual Meeting of the European Association of Archaeologists (EAA), Session: Uncovering historical routes for a sustainable mobility: methods, tools, Case Studies, Verlag De Gruyter, 19 Seiten.
- VOLKMANN, A., im Druck-2: Methods and perspectives in geoarchaeological site catchment analyses: identify paleoclimate signals from the Iron Age until the Middle Ages in the Oder river region, Digital (Geo-) Archaeology and Digital Cultural Heritage, Natural Science in Archaeology, Verlag Springer, 33 Seiten.

mySentinel – Fernerkundungsdaten für das Bundesland Mecklenburg-Vorpommern

SVEN BALTRUSCH¹ & JULIANE PETERS¹

Zusammenfassung: Das Landesamt für innere Verwaltung Mecklenburg-Vorpommern (LAI V M-V) ist gemäß § 20 GeoVermG M-V für die Sammlung und Registrierung von Fernerkundungsergebnissen zuständig, wenn das Land die Nutzungsrechte besitzt. Im Rahmen des Copernicus-Programms werden durch die European Space Agency (ESA) Sentinel-Daten bereitgestellt, auf die das Bundesland kostenfrei zugreifen und diese nutzen kann. Die Landesluftbildstelle übernimmt für die Aufgaben der Landesfachbehörden die Bündelungsfunktionen, d.h. die ESA-Datensätze (Sentinel-1 und Sentinel-2) werden automationsgestützt periodisch recherchiert, für die Landesfläche aufbereitet und den Fachverwaltungen bereitgestellt. Der Beitrag zeigt erste Erfahrungen und gibt eine Perspektive.

1 Einleitung

Seit 2014 ist der erste Sentinel-Satellit des europäischen Copernicus-Programmes im All und liefert weltweite hochauflösende Fernerkundungsdaten der Erdoberfläche. Die Daten sind für registrierte Nutzer frei verfügbar (EC 2013).

Das Amt für Geoinformation, Vermessungs- und Katasterwesen (AfGVK) des Landesamtes für innere Verwaltung (LAI V) Mecklenburg-Vorpommern ist verantwortlich für die Bereitstellung der amtlichen Geodaten des Landes wie zum Beispiel Luftbilder und Laserscandaten. Laut §20 des Gesetzes über das amtliche Geoinformations- und Vermessungswesen (GeoVermG 2010) ist dabei die Landesluftbildstelle zuständig für die Sammlung von „Luftbildern und anderen Fernerkundungsergebnissen, die für die Landesvermessung und für das Liegenschaftskataster von Bedeutung sind und für die das Land das Nutzungsrecht hat“. Die Daten der Sentinel-Missionen erfüllen beide Voraussetzungen in der Hinsicht, dass sie frei verfügbar sind (siehe oben) und dass sie eine Ergänzung zu den amtlichen Geodaten bilden können. So erfassen sie zum einen Daten in Wellenlängenbereichen, die durch amtliche Luftbilder nicht detektiert werden können (z.B. kurzwelliges Infrarot und Radar). Dies ermöglicht die Ableitung von zusätzlichen Derivaten (z.B. NDWI). Zum anderen eignen sich Satellitendaten für flächenhafte Anwendungen größerer Ausdehnung, da die Abdeckung je Szene wesentlich größer ist. Ein weiterer Vorteil ergibt sich durch die kurzen Wiederholzyklen, die durch die Sentinel-Daten möglich werden.

Ziel des AfGVK ist daher der Aufbau eines Fernerkundungskompetenzzentrums, welches als Datenprovider und Dienstleister fungieren soll. Für Nutzer soll ein landesweites Mosaik aus Sentinel-2-Daten bereitgestellt werden, welches regelmäßig aktualisiert wird (z.B.

¹ Landesamt für innere Verwaltung Mecklenburg-Vorpommern,
Amt für Geoinformation, Vermessungs- und Katasterwesen, Lübecker Straße 289, 19059 Schwerin,
E-Mail: [Sven.Baltrusch, Juliane.Peters]@laiv-mv.de

vierteljährlich). Außerdem sollen Kacheln von Sentinel-1 und -2 langfristig gespeichert werden und bei Bedarf an Nutzer, wie z.B. Umwelt- oder Forstverwaltung, abgegeben werden. Der Landesluftbildstelle kommt dabei die Rolle als Organisator und Qualitätsgarant zu. Dies schließt ein Qualitätsmanagement auf Basis der amtlichen Geobasisdaten ein.

Um diese Aufgaben erfüllen zu können, wird in Kooperation mit Geosystems die Software *mySentinel* entwickelt, welche wesentliche Schritte der Datenspeicherung, -verarbeitung und -ablage automatisiert durchführen soll. Dieser Beitrag beschreibt die Software sowie den bisherigen Stand der Umsetzung. Anhand von Beispielen werden Funktionsweise und Resultate gezeigt. Außerdem werden in einem kurzen Ausblick die weiteren Pläne umrissen.

2 Daten

Das ESA-Programm Copernicus wurde 1998 unter der Bezeichnung GMES (Global Monitoring for Environment and Security) mit dem Ziel ins Leben gerufen, ein europäisches Erdobservationssystem zu schaffen (ESA 2017a). Von den im Rahmen von Copernicus geplanten sechs Sentinel-Missionen mit jeweils zwei aktiven Satelliten sind drei bereits gestartet (ESA 2017b). Für das Land Mecklenburg-Vorpommern sind davon insbesondere optische Daten der Sentinel-2-Mission sowie die Radardaten von Sentinel-1 von Interesse.

Die zwei Sentinel-1-Satelliten sind bereits seit 2014 (Sentinel-1A) bzw. 2016 (Sentinel-1B) im Orbit und ermöglichen damit regelmäßige Radarmessungen im Abstand von maximal sechs Tagen. Das jeweilige SAR-Instrument an Bord der Satelliten operiert im C-Band (ESA 2016a).

Von den geplanten zwei Satelliten der Sentinel-2-Mission ist Sentinel-2A seit 2015 im All, -2B soll 2017 folgen. Momentan wird dadurch eine Wiederholrate von 10 Tagen erreicht; wenn das System vollständig ist, wird eine Rate von 5 Tagen erreicht werden. Der Multispectral Imager an Bord der Satelliten observiert die Erdoberfläche in 13 Kanälen im optischen, nahen und kurzwelligen Infrarotbereich mit einer Bodenauflösung von 10 m (sichtbare und ein Nahinfrarotkanal), 20 m (Nahinfrarot- und Kurzwelleninfrarotkanäle) und 60 m (Kanäle für die Atmosphärenkorrektur) (ESA 2016b). Im Folgenden wird das Level-1C-Produkt der Sentinel-2-Daten verwendet. In diesem Produkt ist bereits eine Wolkenmaske enthalten.

3 mySentinel

Hinter der Funktionsweise der Software *mySentinel* steckt die Idee, dass nur die beschreibenden Parameter einer Anfrage vom Software-Nutzer interaktiv definiert werden sollen. Datensuche, Prozessierung und Speicherung sollen vollautomatisch im Hintergrund erfolgen.

Die Ausführung einer konkreten Anwendung in der *mySentinel*-Software kann in zwei Schritte unterteilt werden. Im ersten Schritt werden interaktiv die Parameter für die Datensuche gesetzt und ein Auftrag (Job) definiert, welcher die Suchparameter sowie Einstellungen zur Prozessierung und Ausgabe der Daten speichert. Der zweite Schritt umfasst die vollständig automatisierte Prozessierung der Daten entsprechend der zuvor festgelegten Einstellungen.

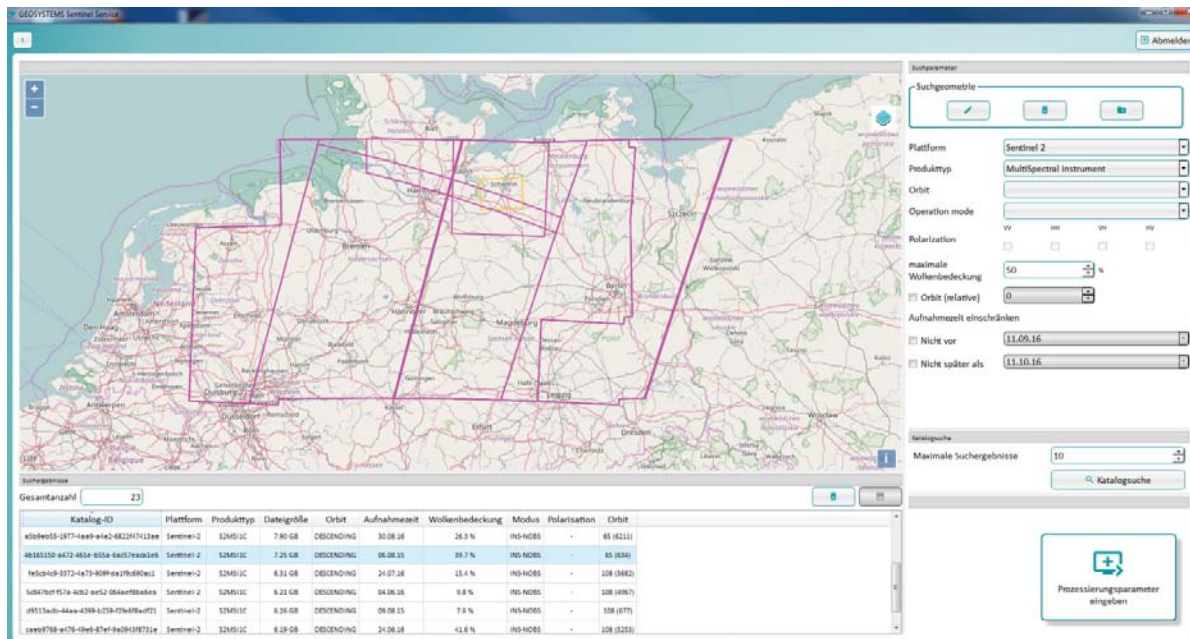


Abb. 1: Datensuche

Bei der Datensuche muss zwingend die Sensorplattform (Sentinel-1 oder Sentinel-2) definiert werden (Abb. 1). Für die Eingrenzung kann eine Suchmaske eingebunden und der Aufnahmezeitraum festgesetzt werden. Im Falle einer Suche nach Sentinel-2-Daten kann zudem die maximale Wolkenbedeckung innerhalb einer Szene definiert werden. Für Sentinel-1-Radardaten kann zwischen den verschiedenen Produkttypen (Single Look Complex, Ground Range Detected, ...) sowie den Operationsmodi (Interferometric Wide Swath, Stripmap, ...) ausgewählt werden. Außerdem besteht die Möglichkeit, einzelne Polarisierungen zu selektieren. Bei der Jobdefinition werden die zuvor gewählten Sucheinstellungen übernommen. Da für Sentinel-1-Daten bisher keine weitere Prozessierung vorgesehen ist, kann nur die Ablage näher bestimmt werden. Für die Daten der Sentinel-2-Satelliten können darüber hinaus die gewünschten Kanäle selektiert werden, wobei hier nur entsprechend der Auflösung unterschieden wird. Das heißt, es kann jeweils gewählt werden, ob alle 10-m-Kanäle und/oder alle 20-m-Kanäle und/oder alle 60-m-Kanäle prozessiert werden sollen. Als Endprodukt wird jeweils ein Layerstack je ausgewählter Auflösungsgruppe gebildet. In jedem Fall kann für die Gruppen zudem definiert werden, in welchem Bildformat die Ergebnisse abgelegt werden, und ob die Einzelbilder mosaikiert werden sollen. Wolken werden momentan durch die Level-1C-Wolkenklassifikation ausmaskiert. Bei der optionalen Mosaikierung werden die entsprechenden Datenlücken durch Pixel aus weiteren (älteren) Kacheln gefüllt, sofern wolkenfreie Daten für den definierten Zeitraum vorhanden sind. Jeder Job kann entweder als einmalig auszuführende Aufgabe oder unter Angabe eines Zeitintervalls für die wiederholte Ausführung angelegt werden.

The screenshot shows the GEOSYSTEMS Sentinel Sentinel interface. At the top, there are navigation icons for 'Job-Ausführung' (Job Execution) and 'Ausführungsdetails' (Execution Details). Below these is a table with the following columns: Katalog-ID, Plattform, Produkttyp, Dateigröße, Orbit, Aufnahmezeit, Wollenbedeckung, Modus, Polarisation, Orbit, Downloadstatus, and Downloadfortschritt. The table contains four rows of data. Below the table, there is a section for 'Weitere Ausführungsinformationen' (Further execution information) showing 'Ausführungsstatus: Download läuft' (Execution status: Download running) and 'Statusinformationen:' (Status information:).

Katalog-ID	Plattform	Produkttyp	Dateigröße	Orbit	Aufnahmezeit	Wollenbedeckung	Modus	Polarisation	Orbit	Downloadstatus	Downloadfortschritt
hf20307-8839-8a08-277621880d13	Sentinel-2	S2MS1C	7.39 GB	DESCENDING	09.09.16	7.1 %	RG-NGBS	-	85 (8554)	Läuft	22%
fc504c9-3372-4a79-909f-da195490a1c1	Sentinel-2	S2MS1C	8.31 GB	DESCENDING	24.07.16	15.4 %	RG-NGBS	-	108 (8582)	In Warteschlange	0%
fb3752c-5558-4758-9aee-b6c7284f536d	Sentinel-2	S2MS1C	1.20 GB	DESCENDING	12.09.16	0.1 %	RG-NGBS	-	108 (8587)	In Warteschlange	0%
6a380d8-1c2f-475e-8fa5-0f436e14f68d	Sentinel-2	S2MS1C	5.25 GB	DESCENDING	12.09.16	0.0 %	RG-NGBS	-	108 (8587)	Läuft	11%
5c8470f-857e-4b32-ba32-064ae8f8a6e6	Sentinel-2	S2MS1C	6.22 GB	DESCENDING	04.06.16	9.8 %	RG-NGBS	-	108 (8587)	In Warteschlange	0%

Abb. 2: Statusanzeige eines Jobs

Die eigentliche Prozessierung der Daten erfolgt automatisch. Der Status eines Jobs kann jederzeit eingesehen werden. In einer Übersicht wird der Downloadfortschritt für die jeweiligen Szenen angezeigt (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Außerdem kann die zugehörige Log-Datei abgerufen werden.

4 Ergebnisse

Für einen ca. 1770 km² großen Bereich im Umkreis der Landeshauptstadt Schwerin soll ein



Abb. 3: Mosaik für Schwerin und Umgebung



Abb. 4: Wolkenmaske (Sentinel-2-Szene vom 09.09.2016)

Mosaik aller Bänder mit 10 m Auflösung (sichtbare Kanäle und Nahinfrarot) erzeugt werden. Als weitere Suchparameter werden eine maximale Wolkenbedeckung von 10% sowie ein Zeitraum vom 21.08.2016 bis zum 21.10.2016 definiert. Die Wolkenmaske soll mit 50 m gepuffert werden, was der Standardeinstellung entspricht. Das Mosaik soll im Bildformat IMG abgelegt werden.

Abb. 3 zeigt das Ergebnis der Prozessierung. Drei Sentinel-2-Szenen mit der entsprechenden Abdeckung existieren für den genannten Zeitraum und gehen somit in die Mosaikierung ein. Die gesamte Prozessierung dauert ca. 50 Minuten. Das Datenvolumen der heruntergeladenen Szenen beträgt insgesamt 14 GB, das finale Mosaik ist 165 MB groß.

Die Klassifikation von Wolken, die in den Level-1C-Daten enthalten ist, scheint Wolken insbesondere in deren Randbereich sowie bei geringer Ausdehnung zu unterschätzen (Abb. 4). Dies kann bei der Mosaikierung in einigen Fällen zu Bildfehlern führen, wenn ausmaskierte Wolkenlücken in einem Bild durch Daten aus einer anderen Szene gefüllt werden (Abb. 5). Im



Abb. 5: Mosaikausschnitt mit gefüllten Wolkenlücken

Beispiel für Schwerin und Umgebung ist ein Füllen von Wolkenlücken innerhalb einer Kachel nicht notwendig, da nahezu der gesamte Suchbereich durch wolkenlose Daten abgedeckt ist.

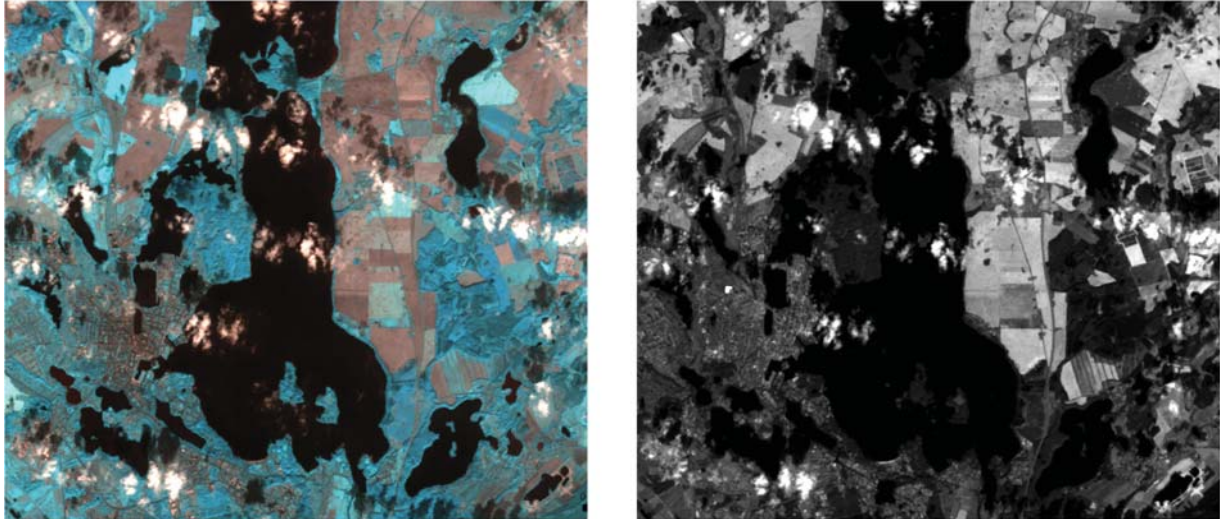


Abb. 6: Kanalkombination der Bänder 5, 6 und 7 (links) und Kanal 12 (rechts) für Schwerin und Umgebung

Für das gleiche Gebiet wie zuvor und den Zeitraum 24.08.2016 bis 24.10.2016 wird ein zweiter Job definiert, der alle Kacheln der Bänder mit 20 m Auflösung speichern soll. Abb. 6 zeigt als Ergebnis die Kombination der Nahinfrarotkanäle (links) sowie Band 12 (kurzwelliges Infrarot; rechts). Die Prozessierung hat wieder rund 50 Minuten in Anspruch genommen. Das Datenvolumen der insgesamt heruntergeladenen Szenen beträgt 14 GB, die Kacheln innerhalb des Suchbereichs ergeben zusammen 1 GB.

5 Ausblick

Aufbauend auf Erfahrungen und in enger Abstimmung mit den Endanwendern in Mecklenburg-Vorpommern erfolgen in Kooperation mit Geosystems weitere Anpassungen der Software, um einerseits die Ergebnisse weiter zu optimieren als auch die Integration der Sentinel-Daten in Fachprozesse zu gewährleisten. Kurzfristig wird die Maskierung von Wolkenbedeckung verbessert. Die bisher genutzte Wolkenmaske des Level-1C-Produktes wird durch eine eigene Wolkenmaske ersetzt. Dabei soll auch der Wolkenanteil innerhalb des definierten Suchbereichs bestimmt werden können. Außerdem wird das Softwaremodul ERDAS ATCOR in die Prozessierung eingebunden, um den Atmosphäreneinfluss auf die Reflektanzen zu korrigieren. Mittelfristig soll mithilfe der Software eine kontinuierliche Datenorganisation erreicht werden, die über die bloße Speicherung hinausgeht. Die Organisation der Daten wird dann durch die Katalogisierung mit ERDAS APOLLO erfolgen. Die Metadaten der Sentinel-Datensätze sollen auf die landesspezifischen Bildanteile angepasst werden, um somit eine eindeutigere Aussage zu erzielen. Zusätzliche Werkzeuge sollen ein Qualitätsmanagement mit den amtlichen Geobasisdaten (z.B. ATKIS-DGM, ATKIS-DOP) als Bezug ermöglichen. Außerdem ist geplant,

auf Basis der Sentinel-2-Daten mithilfe automatisierter *Change Detection* Veränderungshinweise für das Digitale Landschaftsmodell (Basis-DLM) abzuleiten.

Das oben genannte Ziel der flächendeckenden Bereitstellung von Sentinel-2-Daten soll über kontinuierlich aktualisierte WMS-Dienste umgesetzt werden. Neben den üblichen Kombinationen aus optischen Kanälen und ggf. Nahinfrarot sind weitere entsprechend der vorhandenen Sentinel-2-Kanäle denkbar, ebenso wie die Bereitstellung von Radardaten der Sentinel-1-Satelliten. Die Zukunft wird hier zeigen, ob entsprechender Bedarf besteht. Dies gilt auch für das Angebot an verschiedenen Derivaten wie z.B. regelmäßige flächendeckende NDVI-Mosaik, die vom LAiV bereitgestellt werden könnten.

6 Schlussfolgerung

Mit der Software *mySentinel* steht dem LAiV Mecklenburg-Vorpommern ein wirkungsvolles Instrument zur Automatisierung der Workflows rund um die Datenhaltung und die Prozessierung von Sentinel-Daten zur Verfügung. Je nach Bedarf können Einzelkacheln gespeichert werden oder Mosaik mit selbst definierter Flächenabdeckung generiert werden. Wolken können ausmaskiert und die Datenlücken mit Daten aus weiteren (wolkenfreien) Szenen gefüllt werden.

7 Literaturverzeichnis

- EC, 2013: Commission Delegated (EU) Regulation 1159/2013.
- ESA, 2016a: Sentinel-1 – Radar Vision for Copernicus.
http://esamultimedia.esa.int/docs/EarthObservation/Sentinel-1_sheet_March2016.pdf,
 letzter Zugriff 11.01.2017.
- ESA, 2016b: Sentinel-2 – Colour Vision for Copernicus.
http://esamultimedia.esa.int/docs/EarthObservation/SENTINEL-2_sheet_161107.pdf
 letzter Zugriff 11.01.2017.
- ESA, 2017a: Copernicus: Europas Weltraum-Wächter (Special).
http://www.esa.int/ger/ESA_in_your_country/Germany/Copernicus_Europas_Weltraum-Waechter_Special, letzter Zugriff 16.01.2017.
- ESA, 2017b: Copernicus Overview.
http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Overview4, letzter Zugriff 11.01.2017.
- GEOVERMG, 2010: Gesetz über das amtliche Geoinformations- und Vermessungswesen Mecklenburg-Vorpommern. <http://www.landesrecht-mv.de/jportal/portal/page/bsmvprod.psml?showdoccase=1&doc.id=jlr-GeoInfVermGMVrahmen&doc.part=X&st=lr>, letzter Zugriff 20.01.2017.

Fusion of high resolution SAR images with optical and GIS data and its application to change detection in urban areas

CARLOS VILLAMIL LOPEZ¹ & UWE STILLA²

Abstract: SAR images are perfectly suited for change detection, given that they are not affected by different sun illumination conditions and/or clouds. There is potential to improve the SAR change detection results by taking into account prior knowledge of the scene, which can be obtained from other sources of information such as high resolution optical images and data from Geographic Information Systems (GIS). In this paper we will describe how information about the scene geometry and a classification of the scene into different semantic classes can be obtained from the optical and GIS data, and how this information can be transformed to the slant-range coordinate system of SAR images so that it can be easily used in the change detection process. Finally, we will show some initial results that illustrate the benefits of using this information about the scene during the change detection.

1 Introduction

One of the most interesting applications of remote sensing data is the detection of changes due to human activity. This can be applied to the monitoring of urban areas, and can also facilitate urban planning. Ideally, we want to be able to detect changes in certain areas of interest as often as possible. Currently, there are many Earth observation satellites on orbit which have high resolution capabilities and can acquire images of any point on Earth every few days. From the different types of sensors available, Synthetic Aperture Radar (SAR) sensors are better suited for this task, because unlike optical sensors, they are able to operate day and night in all weather conditions. Also, because SAR is an active sensor, the acquired images do not change with the sun illumination conditions, which means that any significant differences between two acquired images will be only due to a change in the imaged scene. This is an important advantage when doing change detection. However, one big disadvantage of SAR sensors is that the interpretation of SAR images of complex scenes, like urban areas, can be a difficult task. This is due to radar specific imaging effects like speckle noise, multiple-bounce propagation, layover and shadowing.

Traditionally, when doing SAR change detection, two or more images of the same scene and acquired at different times are first co-registered (i.e., aligned so that each pixel represents the same location in all images). Then, the amplitude (for incoherent change detection) or the amplitude and phase (for coherent change detection) of the SAR images is compared for every pixel. Typically, the results are shown in a new image that highlights the changes in some color. However, these color images still need to be visualized and interpreted by a human. Modern sensors are capable of acquiring large amounts of data in short periods of time, which means that

¹ Deutsches Zentrum für Luft- und Raumfahrt, Institut für Hochfrequenztechnik und Radarsysteme, Münchener Straße 20, D-82234 Weßling, E-Mail: carlos.villamillopez@dlr.de

² Technische Universität München, Photogrammetrie und Fernerkundung, Arcisstr. 21, D-80333 München, E-Mail: stilla@tum.de

the visual interpretation of all these images is no longer feasible. Therefore, a reliable unsupervised approach that detects changes from a series of SAR images is needed.

Several unsupervised change detection techniques employing very high resolution SAR images can be found in the literature. Most of these are automatic thresholding techniques that work with individual pixels such as the ones presented in (QUIN et al. 2014; SU et Al. 2015). This means that after computing a change detection metric, such as the log-ratio operator (BAZI et al. 2005), a threshold is set automatically and then the pixels below/above this threshold are set as unchanged/changed. In the last few years, a few promising examples of unsupervised change detection approaches that take advantage of some available prior knowledge about the imaged scene have been presented. In BOVOLO et al. (2013), which has a focus on surveillance applications, the authors develop custom feature detectors based on the objects that they expect to find in certain areas of interest of the SAR images. This approach seems to provide very good results, but it needs to be customized for each scene. In TAO & AUER (2016) the authors detect changes on existing buildings and facades by using SAR images together with a Digital Surface Model (DSM). Their approach has the advantage of being capable of working with images acquired with different incidence angles. However, it cannot be used to detect new buildings or any other type of changes. Looking at these results it appears that, if available, prior knowledge about the scene can be used to improve the results of the SAR change detection process.

In this paper we will investigate what kind of information about a scene can be extracted from optical images and GIS data, and how this information can be introduced as prior knowledge into the SAR change detection process to improve the change detection results. Initially, in section 2 we will describe the steps that must be performed in order to achieve this, and some examples using a dataset from the city of Oslo (Norway) will be shown. Finally, the obtained conclusions and the next steps that should be taken to further advance this work will be outlined in section 3.

2 Methodology

In this section, we will describe the different steps for using prior knowledge about a scene together with SAR images, and take it into account when doing SAR change detection. We will obtain this prior knowledge about the scene from optical images and GIS data. The information extracted from these sources must be transformed to match the slant range geometry of SAR images so that it can be used in the SAR change detection process. A simple block diagram showing the different steps of the proposed workflow can be seen in Fig. 1.

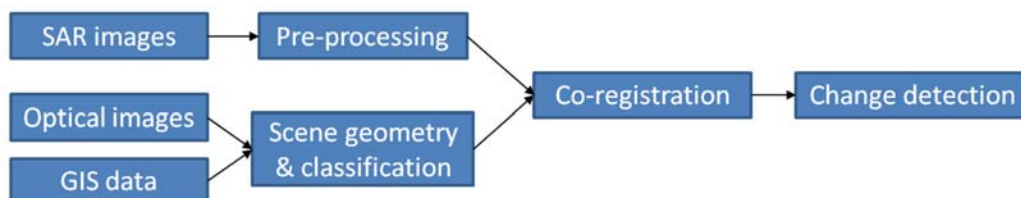


Fig. 1: Block diagram of the proposed workflow for adding information from other sources to the SAR change detection process.

As it can be seen in this block diagram, initially some pre-processing needs to be applied to the SAR images, and information about the scene geometry and a classification of the scene into different semantic classes must be obtained from the GIS and optical data. Then, the information extracted from this data must be co-registered with the SAR images, so that they can be used together during the change detection step. Finally, the changes can be detected by using the SAR images together with the scene knowledge that was obtained from the GIS and optical data. All these steps will be described below.

2.1 Pre-processing of the SAR images

Applying some pre-processing to the SAR images before using them for change detection can help to improve the obtained results. One important task that needs to be performed is the reduction of the speckle noise which is present in all SAR images, as this noise makes it harder to discriminate between the changed and unchanged classes, increasing the number of false alarms. One simple and well established technique for reducing the speckle noise is multi-looking, which significantly improves the radiometric quality of SAR images in exchange for a worse spatial resolution. Depending on the resolution of the SAR images and the size of the smallest changes that must be detected, a worse spatial resolution might not be acceptable. In these cases some other speckle filters can be used, such as the well-known Lee filter (LEE 1986), which attempt to reduce the speckle noise while keeping the spatial resolution of the original SAR image.

After reducing the speckle noise of all the SAR images to be used for change detection, these images must be co-registered. The co-registration of SAR images acquired by the same sensor is not especially difficult, and can be solved using different methods. One of these methods is the geometrical registration presented in (SANSOSTI et al. 2006). Besides being accurate, this method is also very flexible, as it can also be applied to images acquired with different incidence angles and/or different orbits, and not only the typical repeat-pass case.

Once the speckle noise of the SAR images has been reduced and all the images have been co-registered, they are ready to be used for change detection.

2.2 Extracting information about the scene from optical images and GIS data

Although the pre-processed SAR images could already be used for change detection as is, we have already established that there is potential to improve the change detection results if we also use prior information about the scene. This will be especially helpful in urban areas, which are typically very complex scenes in SAR images because of the important role that effects such as layover, shadow and multiple-bounce play in these areas. In order to make our approach fully unsupervised and as general as possible, we will extract this prior information about the scene automatically from other common types of data: optical images and GIS data. More specifically, we will obtain information about the scene geometry in the form of Digital Surface Models (DSM) by photogrammetrically processing stereo optical images, and a classification of the scene into different semantic classes (e.g., buildings, roads, vegetation, etc.) from GIS data (i.e., vector maps).

2.2.1 Extracting information about the scene geometry from stereo optical images

Optical images provide complementary information to the SAR images. Besides being much easier to interpret, the optical images can also be used for obtaining accurate height information in the form of photogrammetric Digital Surface Models (DSM). These DSMs provide us with important information about the scene geometry which can be used to find the layover and shadow areas of the SAR images, as we will describe later in section 2.3.1. The shadow and layover areas of the SAR images play an important role in urban areas, and it has already been shown how their knowledge can be exploited to detect changes in existing buildings (TAO & AUER 2016).

In order to obtain a DSM from stereo optical images, these images need to be photogrammetrically processed. Initially, the orientation obtained from a Global Navigation Satellite System (GNSS) and/or Inertial Measurement Units (IMU) can be refined by using Ground Control Points (GCP) in order to achieve better positional accuracy if needed. Then, a dense stereo matching algorithm, such as the popular Semiglobal Matching algorithm presented in (HIRSCHMULLER 2008), must be applied to obtain the height values. Once the height values are available, the images can be orthogonally reprojected with enforced homogeneity constraints to generate a true orthoimage (VILLAMIL-LOPEZ et al. 2016). Depending on the sensor used for acquiring the images and the extension of the scene to be imaged, several images might need to be processed into a mosaic that covers the whole extent.

For the dataset of the city of Oslo that will be used throughout this paper, a set of UltraCam-X images were processed using the approach described above. The obtained true ortho mosaic and DSM have a ground pixel spacing of 20 cm and can be seen below in Fig. 2, which shows a small area near the Oslo train station.



Fig. 2: True orthoimage (left) and DSM (right) from a small area near the Oslo train station, obtained from a set of optical images by using the approach described above

2.3 Classification of the scene into different semantic classes using GIS data

In addition to the information about the scene geometry that we obtained from the optical images, we can also use data from Geographic Information System (GIS), like vector maps, to obtain additional information about the scene. These maps can provide a classification of the scene in different categories, such as buildings, roads, vegetation, water, etc. This classification can help to improve the results of the change detection (e.g., by allowing us to detect changes using different thresholds for different types of surfaces). Besides, once a change has been detected, knowing in which type of area it took place can help us to distinguish between different types of changes. Vector maps are formed by a series of shapes which have an associated tag/category and with their coordinates specified in a certain map projection. In order to use this information in the change detection these shapes need to be projected into the SAR images, as it will be described later in section 2.3.2.

A great source of free high-quality vector maps is the OpenStreetMap (OSM) project, in which data is added and made freely available by volunteers. The amount of data in OSM and its quality is constantly increasing, although data quality varies worldwide (HAKLAY 2010; ZIELSTRA & HOCHMAIR 2012). The global OSM dataset can be imported and organized into a PostGIS database, which allows to quickly find features of a given type which are located within a certain region. An example of the classification that can be obtained by using OSM data can be seen below in Fig. 3, which shows an optical view of Oslo and the corresponding OSM data for four classes: roads, buildings, vegetation and urban ground.



Fig. 3: Optical view from Oslo (top) and classification obtained using OpenStreetMap vector data (bottom). The classes are: roads (gray), buildings (red), vegetation (green) and urban ground (brown)

2.4 Co-registration of the SAR images with the GIS and optical data

Even though all the SAR images have already been co-registered during the pre-processing step, we still need to co-register them with the GIS and the optical data so that later, in the change detection step, we are able to use the information that we extracted from these sources. This means that we need to transform these different types of data to the same coordinate system. One possibility would be to simply orthorectify the SAR and optical images to the map projection used by the GIS data. However, this is not ideal, as there would be some issues in the layover regions of the SAR images, where multiple points located at different positions are imaged into a single SAR image pixel. To avoid these issues, it is desirable to transform the GIS and optical data into the azimuth and range coordinate system of the SAR images. Initially, we will describe how to transform the optical image (and the DSM we obtained from it) to accurately match the SAR images. Finally, we will briefly mention how to do the same for the GIS data, although this task is much easier.

2.4.1 Transforming optical data to the SAR slant range geometry

The co-registration of SAR and optical images was traditionally a difficult task due to the big radiometric and geometric differences that exist between these two image types. However, the high geolocation accuracy of current state-of-the-art SAR and optical sensors makes this much simpler, by allowing us just to use the geometric information instead of having to use more complex feature or intensity based registration methods. Recently, we presented a method for the co-registration of SAR and optical images using this geometric information, which is capable of handling the ambiguities that appear in the layover areas by using a DSM. We describe this method in detail in (VILLAMIL-LOPEZ et al. 2016), but it will also be summarized below.

For any given pixel in a SAR image, we know the corresponding sensor position and the distance between the sensor and all the points that are imaged in this pixel. By using this information and the range and Doppler equations presented in (CURLANDER 1982), we obtain a circle which defines all the possible locations of these points. By intersecting this circle with the DSM obtained from the optical images, we can find how many points are imaged in this pixel and get their respective locations. With this information, we can easily obtain the shadow and layover areas of the SAR image, and also the pixels in an optical image that correspond to each pixel in the SAR image. The optical image can then be resampled to the coordinate system of the SAR image, obtaining the co-registered optical image.

In order to find all these intersections in an efficient way, we need to be able to perform fast and accurate geocoding (i.e., convert between image and ground coordinates and vice versa). The Rational Polynomial Coefficients (RPC) are a set of polynomials that can be used for fast geocoding, and they can be computed for a SAR image as described in (ZHANG et al. 2011). The loss of accuracy due to the use of RPC instead of to the physical SAR sensor model described in (CURLANDER 1982) is negligible, with errors typically below millimeter level. The optical data has already been processed to a true orthoimage, as we described before in section 2.2.1, which makes the conversion between image and ground coordinates trivial for this type of data.

Once the conversion between image and ground coordinates can be done efficiently for both types of data, we need to find the intersections between the circle that describes the possible locations of a SAR pixel and the DSM. This process is illustrated in Fig. 4, which shows a very

typical case that appears in most SAR images of buildings, where the reflections from the building's roof, façade and the ground are imaged in the same SAR image pixel. In the top-left of this figure, a SAR image of a building with a pixel highlighted in red can be seen. The ground coordinates of this pixel are computed for different height values using the RPC, and the obtained points are projected into the DSM (top-right of Fig. 4). Then, the height value of each of these points can be compared with the corresponding DSM height values. Finally, as it can be seen at the bottom of Fig. 4, the obtained intersections give us the location of the scatterers which can be imaged in the highlighted SAR pixel.

In certain cases something might block the direct line of sight between the SAR sensor and some of these points, which means that the SAR sensor will not be able to see them. For a given SAR pixel, if none of these points can be seen by the SAR sensor, the pixel will correspond to a shadow area; whereas if more than one point can be seen, then the pixel corresponds to a layover area. To check if a point is seen by the SAR sensor, we need to compute the look angle (i.e., the angle between the nadir vector and the line of sight between the sensor and this point) and compare it with the look angle for other points at the same range line (i.e., same azimuth position). A point will be occluded (and therefore not seen by the SAR sensor) if there is another point in the same range line located closer to the sensor and with a higher look angle.

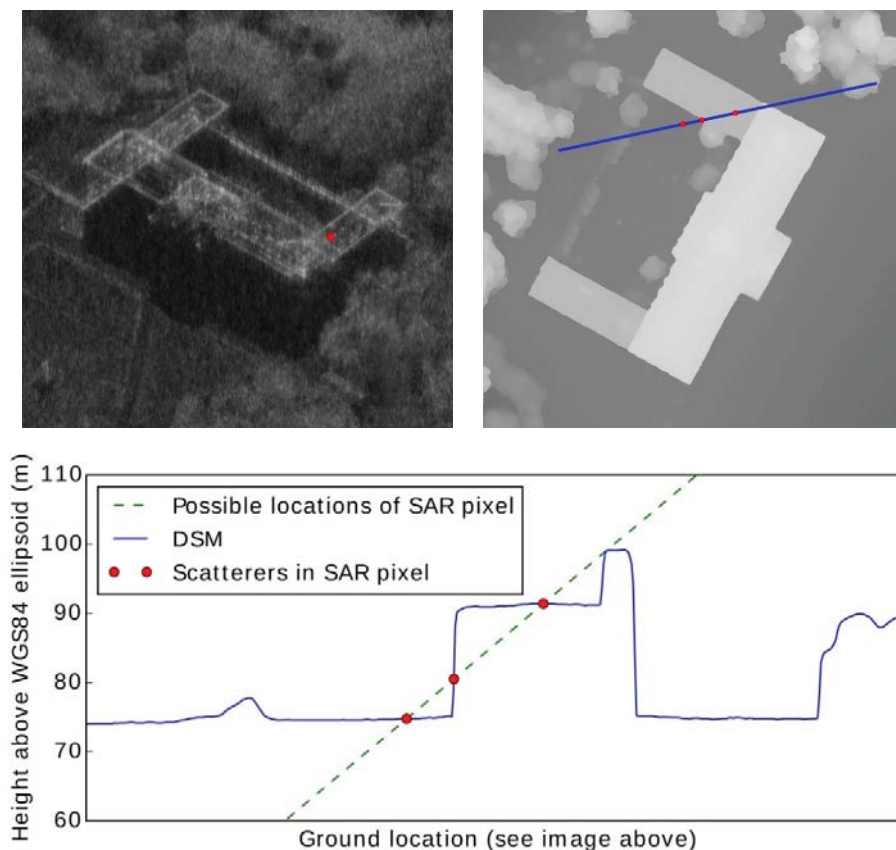


Fig. 4: Finding the location of the scatterers that are imaged in a SAR pixel using a DSM: SAR image with a pixel highlighted in red (top left), DSM with the possible locations of this pixel highlighted in blue (top right), and plot showing the intersections between these locations and the DSM (bottom)

2.4.2 Transforming GIS data to the SAR slant range geometry

The vector shapes from the GIS data need to be projected to the SAR images so that we can use this information later in the SAR change detection step. These shapes are defined by coordinates in a certain map projection (e.g., latitude and longitude or UTM). However, they are typically missing the height coordinate, which we need to project these shapes into the SAR images. Because the shapes in these maps represent objects that are located on the ground surface, we need to obtain those heights from a Digital Elevation Model (DEM) which depicts the terrain height, rather than the DSM that we previously used (which includes the heights of buildings and other structures). Once we know these height values, we can easily project the shapes in the vector map to the SAR images by using the RPCs of the SAR image that we computed before, obtaining a classification of the SAR image in different semantic categories as a result.

2.4.3 Sample co-registration results

To check the accuracy achievable with the co-registration methods described in this section, we applied them to co-register the orthophoto and the OpenStreetMap vector data of Oslo described in the previous sections with a SAR image showing an area with a few buildings. The obtained results are shown below in Fig. 5, where we can see how the co-registered optical image is perfectly aligned with the SAR image even in the layover areas. It is interesting to note that although the optical image used was a true orthophoto (i.e., seen from above), after the co-registration the building façades are now visible because they are also visible in the SAR image, and we transformed the orthophoto to the SAR slant range geometry. Regarding the vector maps from OpenStreetMap, we can see that the building footprints and the roads are accurately projected to the same ground locations they occupy on the SAR image. The vegetation areas with some trees are missing from the OpenStreetMap data, but this has nothing to do with the co-registration method, and more accurate maps from a different source could be used if available.



Fig. 5: Example of co-registration results obtained with the described approach: SAR image of a small area of the city of Oslo (left), co-registered optical image (center), and GIS data showing the location of building footprints in red and roads in dark gray (right)

2.5 Using knowledge about the scene in SAR change detection: initial results

Once we have transformed all the different types of data to the SAR slant range geometry, we can use the information about the scene that we have obtained in the previous steps together with the SAR images for detecting changes caused by human activity. As we already mentioned, the DSM can be exploited for detecting changes in existing buildings by looking for changes in their corresponding shadow and layover areas, whereas the vector maps can be helpful for other types of changes by allowing us to detect the changes in different regions separately. An example of using a DSM for detecting changes in existing buildings has already been shown in (TAO & AUER, 2016). Therefore, to demonstrate the potential of the approach presented in this paper we chose not to focus on buildings, but rather on taking advantage of the vector maps for detecting other types of changes. To illustrate this, we will use two TerraSAR-X images of Oslo (with dates of 12.03.2016 and 23.03.2016) acquired with the very high resolution Staring Spotlight imaging mode, together with the OpenStreetMap data that we described previously in section 2.2.2.

The most prominent changes in our dataset correspond to activity at Oslo the city port, such as moving ships and cargo. The detection of these ships serves as a perfect example for showing how vector maps during the change detection for detecting only specific types of changes. For starters, we know that ships will obviously be located on water, and we know which parts of the image correspond to water thanks to the OpenStreetMap data. Therefore, we only need to look for changes in these image regions, and we can set a threshold for detecting these changes by taking into account the SAR amplitude of just the water regions. The histogram of the log-ratio image (i.e., the logarithm of the ratio between the amplitude of both SAR images) was computed for the complete image and for only the water regions, and the results can be seen in Fig. 6.

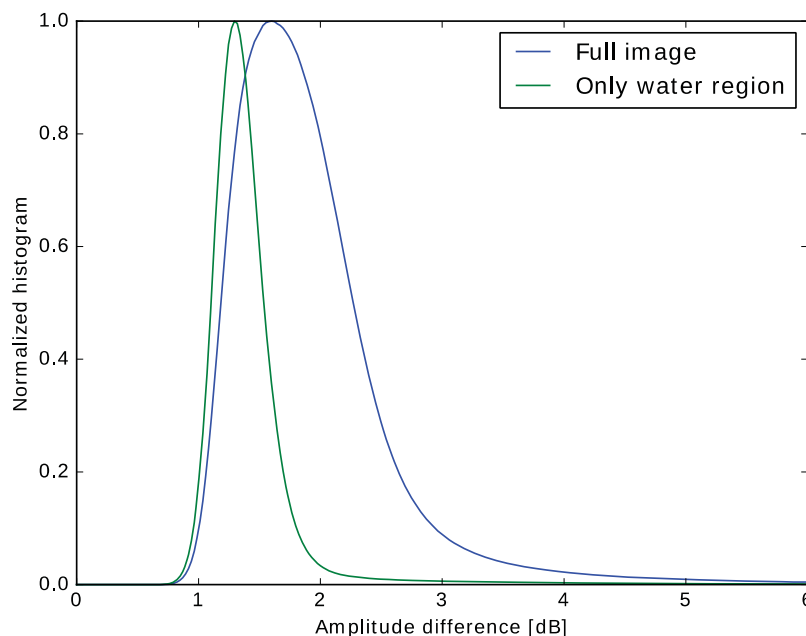


Fig. 6: Plot showing the histogram of the difference between the amplitude of the two SAR images computed for the whole image (blue) and only for the water region (green line)

For finding the threshold to separate the change and no-change classes, an unsupervised thresholding technique such as the ones presented in (BAZI et al 2005; QUIN et. al 2014) can be applied. Looking at the histograms shown in Fig. 6 it becomes clear that computing this threshold for a certain image region (such as water in this case) will provide different results than computing it using the whole image, and it seems obvious that more accurate results will be obtained when computing a specific threshold for each region.

For this simple example of ship detection no automatic thresholding technique was applied, and instead the threshold was set to a fixed 6 dB value. This value was selected by taking into account that ships have bright signatures in SAR images, and the corresponding SAR amplitude is typically at least one order of magnitude (i.e., 10 dB) than that of water (which appears as a dark homogeneous surface). Finally, to remove false alarms due to speckle noise and other detected changes which were not ships, the changes that were too small and therefore could not be ships (which typically have sizes of at least 10 meters) were discarded automatically. This last step was done by analyzing the number of pixels in each of the connected components of the thresholded log-ratio image. The results of this example are shown below in Fig. 7, which shows the bounding boxes of the automatically detected ships in red, painted over a RGB composite image that highlights the changes in orange for easier visualization.



Fig. 7: Results of applying the proposed workflow for detecting ships that arrived/left the port of Oslo. Red rectangles show the ships that were automatically detected. For easier visualization, the rectangles are drawn over a RGB composite of two TerraSAR-X images that highlights the changes in orange

Looking at the results shown in Fig. 7, we can see that all the large ships in this scene were correctly detected. However, one object at the port was erroneously identified as a ship, and three small ships located very close to each other were detected as a single ship. From the results of this simple example, we can establish that the presented method has potential for the unsupervised detection of changes and the classification of these changes (e.g., buildings, ships, etc.). However, these are still initial results which are meant to illustrate the ideas presented in the paper, and more research is needed to be able to robustly detect and identify multiple types of changes, and make it applicable in more scenarios.

3 Conclusions and future work

Remote sensing data can be used for detecting changes due to human activity. Synthetic Aperture Radar (SAR) sensors are better suited for this task than other types of sensors because SAR images are not affected by the sun illumination conditions and/or clouds. However, robustly detecting these changes in an unsupervised manner (i.e., without any human interaction) is still challenging. If available, prior knowledge about the imaged scene can be taken into account during the change detection process to simplify this problem and improve the obtained results. The introduction of this prior knowledge appears especially interesting for urban areas, which are typically very complex scenes in SAR images because of the important role that effects such as layover, shadow and multiple-bounce play in the urban environment. Two interesting sources for this prior knowledge about a scene are GIS and optical data: vector maps provide a classification of the scene into different semantic classes, and stereo optical images can be processed into a DSM that contains accurate information about the scene geometry. The information extracted from these sources needs to be transformed to the slant range geometry of SAR images in order to be used during the change detection step. Another possibility would be to orthorectify the SAR images, but this is not recommended because of the ambiguities that exist in the layover areas, which are very common in urban areas. A complete workflow for extracting this prior knowledge about the scene from GIS and optical data, transforming it to the SAR slant range geometry, and taking it into account during the SAR change detection process has been described in section 2.

Even though we have described the complete workflow, the main focus has been on the initial steps: which kind prior scene knowledge can be useful, where and how to obtain this information, and how to transform it to the geometry of the SAR images so that it can be used together with the SAR images for the change detection. This means that only some initial research has been done on how to exploit this prior scene knowledge for detecting changes with SAR images, and even though we have already shown that there is a big potential, more research is needed to fully take advantage of it.

In addition to this, the potential of other types of GIS data (different than the vector maps from OpenStreetMap which were used in this paper) must also be explored. One very promising source of information for our application is data in CityGML format, which is an open data model and XML-based format for the storage and exchange of virtual 3D city models (Gröger et al. 2012). CityGML is becoming a popular standard, and datasets in this format can be imported and organized into a PostGIS database (which we also used for the OpenStreetMap data) using

the 3DCityDB software (Stadler et al. 2009; Kunde et al. 2013). Therefore, the presented workflow should not change much, and the 3D models of buildings which are available within these datasets might eliminate the need of using a DSM.

4 Acknowledgement

The authors would like to thank Dirk Frommholz and Frank Lehmann from the Institute of Optical Sensor Systems of the German Aerospace Center (DLR) for the processing of the aerial optical image and DSM used in this paper.

The TerraSAR-X staring spotlight images were acquired through the proposal MTH3256 thanks to DLR's TerraSAR-X science service system.

Finally, we would also like to thank OpenStreetMap and its contributors for creating high quality maps and making them available as open data.

5 References

- BAZI, Y., BRUZZONE, L. & MELGANI, F., 2005: An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, **43**(4), 874-886.
- BOVOLO, F., MARIN, C. & BRUZZONE, L., 2013: A hierarchical approach to change detection in very high resolution SAR images for surveillance applications. *IEEE Transactions on Geoscience and Remote Sensing* **51**(4), 2042-2054.
- CURLANDER, J. C., 1982: Location of Spaceborne SAR Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **GE-20**(3), 359-364.
- GRÖGER, G., KOLBE, T. H., NAGEL, C. & HÄFELE, K., 2012: OGC City Geography Markup Language (CityGML) Encoding Standard. Open Geospatial Consortium.
- HAKLAY, M., 2010: How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. *Environment and Planning B: Planning and Design*, **37**(4), 682-703.
- HIRSCHMULLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **30**(2), 328-341.
- KUNDE, F., NAGEL, C., HERRERRUELA, J., ROSS, L. & KOLBE, T., 2013: 3D-Stadtmodelle in PostGIS mit der 3D City. Tagungsband der FOSSGIS-Konferenz.
- LEE, J. S., 1986: Speckle suppression and analysis for synthetic aperture radar images. *Opt. Eng.*, 636-643.
- QUIN, G., PINEL-PUYSSÉGUR, B., NICOLAS, J. M. & LOREAUX, P., 2014: MIMOSA: An automatic change detection method for sar time series. *IEEE Transactions on Geoscience and Remote Sensing* **52**(9), 5349-5363.
- SANSOSTI, E., BERARDINO, P., MANUNTA, M., SERAFINO, F. & FORNARO, G., 2006: Geometrical SAR image registration. *IEEE Transactions on Geoscience and Remote Sensing* **44**(10), 2861-2870.
- STADLER, A., NAGEL, C., KÖNIG, G. & KOLBE, T. H., 2009: Making Interoperability Persistent: A 3D Geo Database Based on CityGML. *3D Geo-Information Sciences*, 175-192.

- SU, X., DELEDALLE, C.-A., TUPIN, F. & SUN, H., 2015: NORCAMA: Change analysis in SAR time series by likelihood ratio change matrix clustering. *ISPRS Journal of Photogrammetry and Remote Sensing* **101**, 247-261.
- TAO, J. & AUER, S., 2016: Simulation-Based Building Change Detection From Multiangle SAR Images and Digital Surface Models. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **9**(8), 3777-3791.
- VILLAMIL-LOPEZ, C., PETERSEN, L., SPECK, R. & FROMMHOLZ, D., 2016: Registration of Very High Resolution SAR and Optical Images. *Proceedings of EUSAR 2016: 11th European Conference on Synthetic Aperture Radar*, 1-6.
- ZHANG, G., LI, Z., PAN, H. B., QIANG, Q. & ZHAI, L., 2011: Orientation of Spaceborne SAR Stereo Pairs Employing the RPC Adjustment Model. *IEEE Transactions on Geoscience and Remote Sensing* **49**(7), 2782-2792.
- ZIELSTRA, D. & HOCHMAIR, H., 2012: Using Free and Proprietary Data to Compare Shortest-Path Lengths for Effective Pedestrian Routing in Street Networks. *Transportation Research Record* **2299**, 41-47.

Investigation of the Potential of Hyperspectral EnMAP Data for Land Cover and Land Use Classification

S. KELLER¹, A. C. BRAUN², S. HINZ¹ & M. WEINMANN¹

Abstract: In this paper, we focus on the classification of hyperspectral data that is expected to be acquired with the Environmental Mapping and Analysis Program (EnMAP) mission, a hyperspectral satellite mission supposed to be launched into space in the near future. More specifically, we consider the dataset presented with the EnMAP Contest, a benchmark that has recently been initiated with the objective of classifying different land cover and land use classes based on EnMAP-like hyperspectral data. We address the EnMAP Contest by presenting a framework that involves (1) standard approaches for dimensionality reduction and feature selection and (2) classifiers relying on different learning principles. The classification results derived for different configurations of our framework clearly reveal the potential of respective techniques and provide the basis for further improvements in different research directions.

1 Introduction

The automated analysis of hyperspectral imagery has become a topic of major interest in remote sensing as such data is valuable for the classification of land cover and land use (PLAZA et al. 2009; CAMPS-VALLS et al. 2014). However, most of the currently available benchmark datasets for evaluating the performance of respective classification approaches consist of (partially) labeled airborne hyperspectral imagery acquired during low-altitude flights. This allows an analysis of areas with a rather limited extent while a large-scale or even nationwide coverage would often be desirable. Recent attempts towards a large-scale coverage with hyperspectral data have resulted in a conceptual transfer towards the use of satellite remote sensing which is addressed by the *Environmental Mapping and Analysis Program (EnMAP)*. EnMAP represents a hyperspectral satellite mission supposed to be launched into space in the near future (KAUFMANN et al. 2008). In order to already obtain first impressions about expected future hyperspectral EnMAP data products before the mission is launched, simulated EnMAP-like hyperspectral data with 244 spectral bands covering the range from 420 nm to 2450 nm has been released and can be used for numerous investigations (GUANTER et al. 2009; SEGL et al. 2010).

In this paper, we focus on the classification of simulated EnMAP-like hyperspectral data (Fig. 1). We present a framework which involves (1) standard approaches for dimensionality reduction and feature selection and (2) several classifiers relying on different learning principles (instance-based learning, probabilistic learning and ensemble learning). To evaluate the performance of this framework, we consider nine different configurations of the framework and present the respective classification results for a recently published benchmark dataset with 20 different land

¹ Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Englerstraße 7, 76131 Karlsruhe, Germany, E-Mail: [sina.keller, stefan.hinz, martin.weinmann]@kit.edu

² Karlsruhe Institute of Technology (KIT), Institute of Regional Science, Reinhard-Baumeister-Platz 1, 76131 Karlsruhe, Germany, E-Mail: andreas.ch.braun@kit.edu

cover and land use classes (five water classes, five forest/meadow classes, five agriculture classes, three urban/industrial classes, and two pasture/fallow classes) in tables and figures. These results correspond to an overall accuracy in the range of about 77-84 % and a kappa value in the range of about 76-83 %, whereby the best classification results are obtained with the classifier relying on ensemble learning. The derived results clearly reveal the potential of EnMAP-like hyperspectral data, and they provide a basis for further investigations in different research directions.

This paper represents an extension of our previous work (KELLER et al. 2016), where we investigated the impact of dimensionality reduction and feature selection on the classification of hyperspectral EnMAP data. In addition to the quantitative classification results obtained for the subset of labeled pixels, we now also focus on the qualitative results obtained for the whole scene. We provide an in-depth analysis of the results obtained for nine different configurations of our framework by considering global evaluation metrics (overall accuracy, Cohen’s kappa coefficient, average completeness, average correctness and average quality) as well as class-wise evaluation metrics (completeness, correctness and quality) and the whole classified scene corresponding to the given simulated EnMAP-like hyperspectral data.

After briefly summarizing related work in Section 2, we explain the single components of our proposed framework for the classification of hyperspectral EnMAP-like data in Section 3. Subsequently, in Section 4, we describe the benchmark dataset used in the scope of our experiments in detail, and we present the respective results derived with our framework and discuss these results with respect to different criteria. Finally, we provide concluding remarks and suggestions for future work in Section 5.

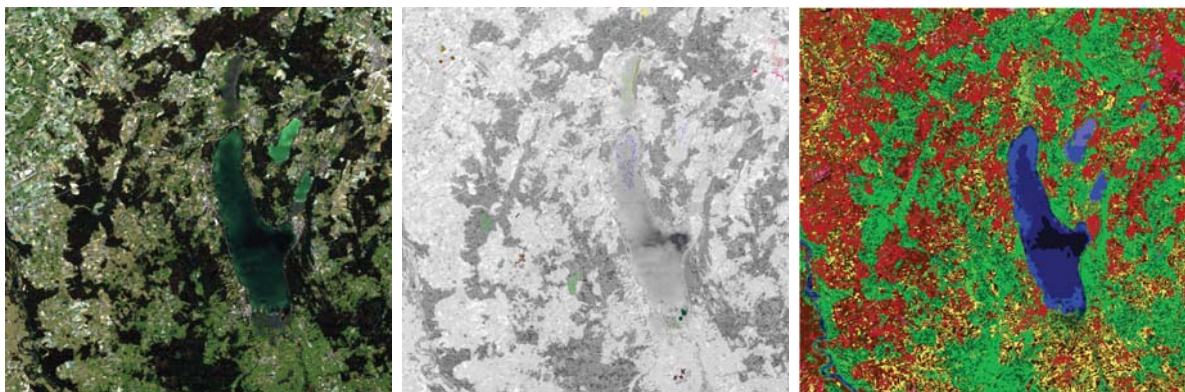


Fig. 1: The *EnMAP Contest* (BRAUN et al. 2015): simulated hyperspectral data covering 244 spectral bands is available for each pixel of the scene (left; image courtesy of Dr. K. Segl). Some pixels within the simulated EnMAP dataset have been labeled (center). The objective is to assign respective class labels to all pixels (right).

2 Related Work

The simplest approach for classifying hyperspectral imagery certainly consists in a per-pixel consideration of the reflectance values of all spectral bands and a subsequent pixel-wise classification based on well-known standard classification approaches which are meanwhile

available in a variety of software packages, e.g. kernel-based methods (CAMPS-VALLS & BRUZZONE 2005), Support Vector Machines (MELGANI & BRUZZONE 2004; CHI et al. 2008), or Random Forests (HAM et al. 2005; JOELSSON et al. 2005). In particular, Support Vector Machines have been widely used for classifying hyperspectral data and related but more sophisticated classification approaches represented by Import Vector Machines and Relevance Vector Machines have also been used for this task (BRAUN et al. 2011; BRAUN et al. 2012).

Despite the classifier, the data representation plays an important role. When classifying hyperspectral data, it should generally be taken into account that the reflectance values of some spectral bands are strongly correlated and that typically not all spectral bands are relevant to the classification task. In particular for high-dimensional data representations as given for hyperspectral data, the Hughes phenomenon (HUGHES 1968) can often be observed. According to this phenomenon, an increase of the number of features over a certain threshold results in a decrease in classification accuracy, given a constant number of training examples (MELGANI & BRUZZONE 2004). Hence, it seems desirable to transfer the given high-dimensional data into a new, more compact representation which encodes almost the same information with respect to the classification task. On the one hand, this can be done via dimensionality reduction (VAN DER MAATEN et al. 2009) which focuses on the mapping of feature vectors from the original feature space onto a specific subspace spanned by meta-features. For this purpose, variants of the well-known Principal Component Analysis or the Independent Component Analysis are commonly applied (LICCIARDI et al. 2012; WANG & CHANG 2006; VILLA et al. 2011). On the other hand, it is possible to apply feature selection (GUYON & ELISSEEFF 2003; SAEYS et al. 2007) which aims at only retaining the relevant features for classification and thus only considers reflectance values of specific spectral bands. Respective methods for instance allow assessing the importance of single spectral bands for land cover classification (LE BRIS et al. 2014) as well as different band selection and band fusion processes (CHEHATA et al. 2014). These studies demonstrate that approaches for both dimensionality reduction as well as feature selection are appropriate to cope with the high degree of redundancy contained in hyperspectral datasets and that the consideration of only a relatively small subspace of the data is still sufficient to derive reasonable classification results.

More recent trends in the classification of hyperspectral imagery mainly address the consideration of spatial context, i.e. the spatial structure within the local image neighborhood of a pixel is considered in addition to the reflectance values of the spectral bands characterizing that pixel. This is commonly referred to as spectral-spatial classification. Respective approaches are for instance given by a probabilistic pixel-wise classification that is followed by either a Markov Random Field regularization (TARABALKA et al. 2010) or a hierarchical optimization (TARABALKA & TILTON 2011). Furthermore, it has been proposed to involve segmentation approaches to support classification, e.g. by using the results of a pixel-wise classification in combination with watershed segments (TARABALKA et al. 2008) which allows a majority voting within watershed segments to account for spatial information. Instead of relying on a segment-based consideration of the results of a pixel-wise classification, spatial context can also be assessed by sampling spectral information within the local surrounding of a pixel, e.g. within adaptive pixel neighborhoods (FAUVEL et al. 2008) or within superpixels (FANG et al. 2015).

3 Methodology

In the scope of this paper, we present a framework for classifying hyperspectral data (Fig. 2). This framework includes three different options for the data representation per pixel (Section 3.1) as well as three different options for supervised classification (Section 3.2).

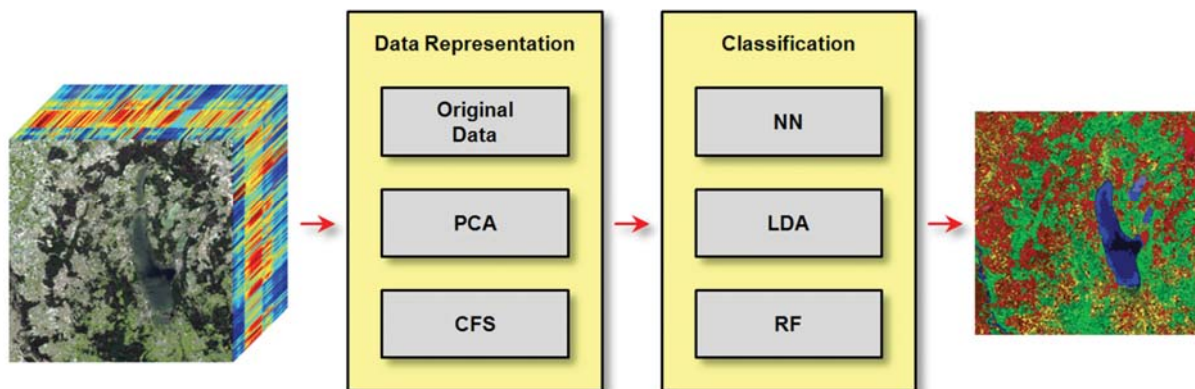


Fig. 2: The proposed framework allowing for three different options for the data representation and three different options for supervised classification.

3.1 Data Representation

In our framework, we focus on different options for deriving feature vectors which serve as input to the subsequent classification procedure. We assume that each pixel is described with a high number of spectral bands and that the spectral information of each spectral band is represented by a reflectance value between 0 % and 100 %. In the following, we focus on three options to derive the feature vectors that serve as input to the subsequent classification procedure.

3.1.1 Original Data Representation

The straightforward approach consists in using the given representation of the data as input to the classification procedure, where we simply concatenate the reflectance values for all spectral bands to obtain the respective feature vectors.

3.1.2 Dimensionality Reduction: Principal Component Analysis

For dimensionality reduction, we focus on using a standard *Principal Component Analysis (PCA)* in the scope of this paper and assume that this is appropriate, even though the PCA is not necessarily the best approach for dimensionality reduction (CHERIYADAT & BRUCE 2003). The main idea of the PCA consists in transforming a set of feature vectors from a high-dimensional feature space (which is spanned by possibly correlated features) to a new feature space which is spanned by linearly uncorrelated meta-features, the principal components, which account for the data variability along the respective dimension. Ranking the meta-features with respect to the respectively covered variability allows us to consider the few meta-features that are most

relevant with respect to data variability, while other meta-features are almost irrelevant with respect to data variability. In the scope of our work, we use the most relevant meta-features that cover 99.9 % of the variability of the given training data to define the new feature vectors, and we assume that we do not have a significant loss of relevant information with respect to the classification task when discarding all other, less relevant meta-features.

3.1.3 Feature Selection: Correlation-based Feature Selection

For feature selection, we focus on using *Correlation-based Feature Selection (CFS)* (HALL 1999) which takes into account (1) the correlation between features and classes to identify relevant features and (2) the correlation among features to identify and discard redundant features. More specifically, CFS exploits the average correlation $\bar{\rho}_{FC}$ between features and classes as well as the average correlation $\bar{\rho}_{FF}$ among classes to evaluate the relevance R of a feature subset comprising n features according to

$$R = \frac{n\bar{\rho}_{FC}}{\sqrt{n + n(n-1)\bar{\rho}_{FF}}}$$

where the correlation metric is given by the symmetrical uncertainty (HALL 1999). Deriving a suitable feature subset thus corresponds to maximizing the relevance R over the feature subset space. Beginning with an initial feature subset either a feature is added to the feature subset or a feature is removed from the feature subset until the relevance R converges to a stable maximum. The new feature vectors thus correspond to a selection of a few dimensions of the original data representation.

3.2 Supervised Classification

For classification, we integrate three classifiers relying on different learning principles into our framework. This allows us to draw more general conclusions about the impact of dimensionality reduction and feature selection on the classification of hyperspectral EnMAP data.

3.2.1 Nearest Neighbor Classifier

The *Nearest Neighbor (NN) classifier* relies on the principle of instance-based learning. For each feature vector belonging to a sample in the test set, this classifier performs a comparison to the feature vectors belonging to the samples in the training data and selects the class label associated with the most similar feature vector in the training data. Accordingly, the induction process is delayed to the prediction stage, i.e. there is no training stage and the prediction stage might correspond to a higher computational effort due to the required comparisons of vectors (here with respect to the Euclidean distance).

3.2.2 Linear Discriminant Analysis Classifier

The *Linear Discriminant Analysis (LDA) classifier* relies on the principle of probabilistic learning. In the training stage, it is assumed that the instances of different classes exhibit a Gaussian distribution in the feature space. Accordingly, the training of the LDA classifier consists in fitting a multivariate Gaussian distribution to the given training data, whereby the parameters of a Gaussian distribution have to be estimated for each class. Due to a lack of

knowledge about the behavior of single classes in the feature space, the same covariance matrix is assumed for each class so that only the means vary for the different classes. In the prediction stage, the class probabilities are evaluated for each feature vector belonging to a sample in the test set and the label of the class with the maximum probability is selected.

3.2.3 Random Forest Classifier

The *Random Forest (RF) classifier* (BREIMAN 2001) relies on the principle of ensemble learning. In the training stage, an ensemble of randomly trained decision trees is generated via bagging (BREIMAN 1996). Thereby, subsets of the training data are randomly drawn with replacement and an individual decision tree is trained for each subset. In the prediction stage, for each feature vector belonging to a sample in the test set, each decision tree casts a vote for one of the defined classes and the majority vote is selected to obtain the most probable class label.

4 Experimental Results

In the following, we first introduce the used benchmark dataset (Section 4.1). Subsequently, we present the derived results (Section 4.2) and we also provide a short discussion with respect to different aspects (Section 4.3).

4.1 Dataset

To evaluate the performance of our framework, we use the dataset presented in the scope of the recent EnMAP contest (BRAUN et al. 2015). This dataset is based on the simulated *EnMAP Alpine Foreland dataset* (GUANTER et al. 2009; SEGL et al. 2010), but it additionally contains a subset of pixels that are labeled and thus allow a benchmarking of different approaches for the classification of hyperspectral data.

Like the simulated EnMAP Alpine Foreland dataset, the considered dataset covers an area of about $30 \text{ km} \times 30 \text{ km}$ around the Ammersee in Bavaria, Germany. It is represented by an image of 1000×1000 pixels (i.e. the ground sampling distance is 30 m), where each pixel is associated with values on 244 simulated spectral bands covering the spectral range of 420-2450 nm with a varying spectral sampling of 6.5-10 nm. A near natural color visualization of the considered image is provided in Fig. 1 (left) and reveals a diversity of water, vegetation, agricultural, and urban/industrial classes.

The additional labeling has been performed for a subset of pixels as shown in Fig. 1 (center) and with respect to 20 different land cover and land use classes, whereby the manual annotation has been done by considering visual differences in the image (considering several band combinations) as well as the individual spectra of selected pixels. More specifically, the labeling takes into account five water classes, five forest/meadow classes, five agriculture classes, three urban/industrial classes, and two pasture/fallow classes. The average spectra of these classes are shown in Fig. 3, and it can be observed that some of the classes reveal a spectrally very similar behavior, which makes classification more challenging. With the annotation, a split of the labeled pixels into training data and test data is already provided, whereby the training data comprises hyperspectral data for 2,617 labeled pixels and the test data contains hyperspectral data for 1,124 labeled pixels.

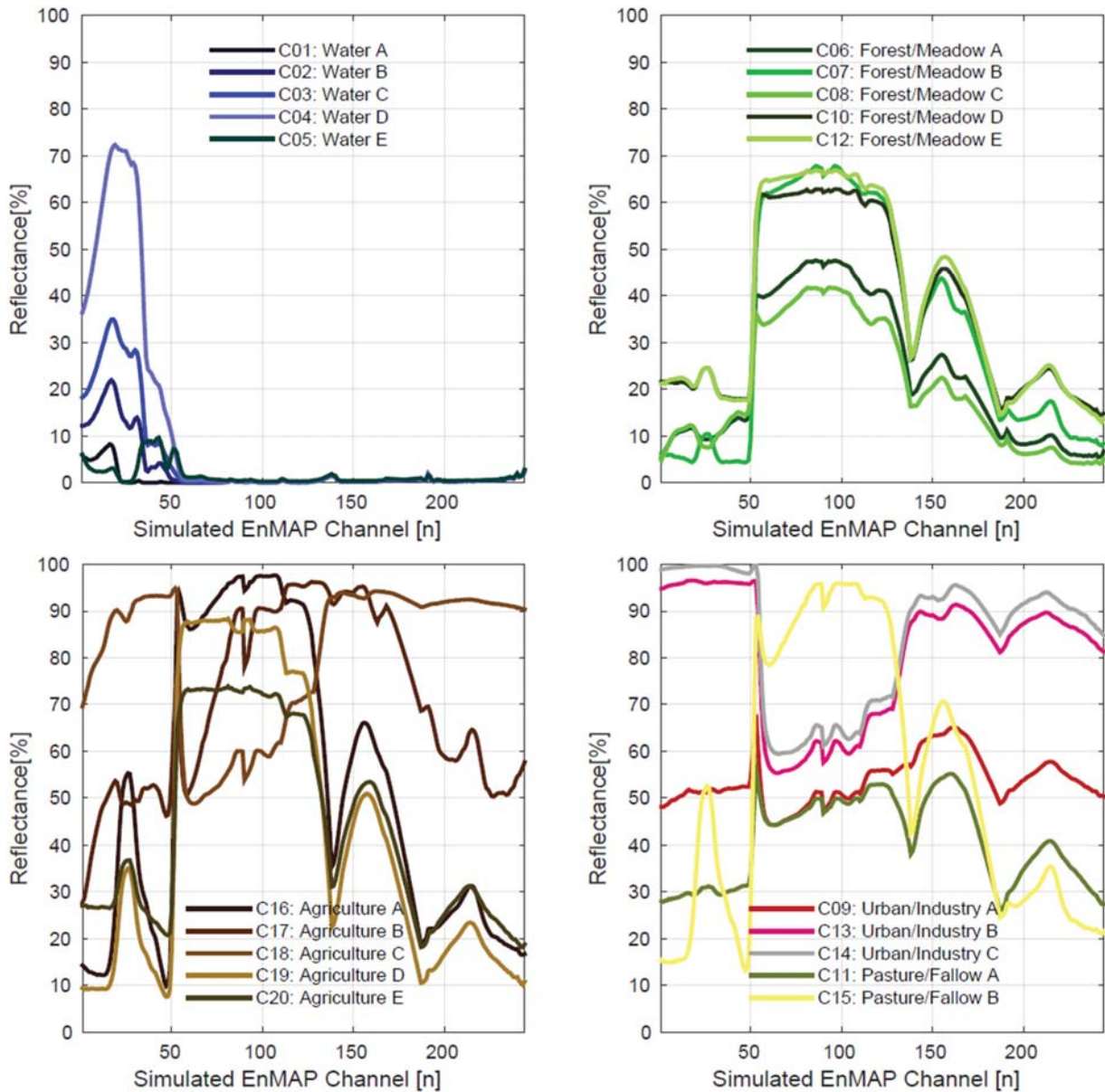


Fig. 3: Average spectra for the 20 land cover and land use classes considered in the scope of the EnMAP contest (BRAUN et al. 2015): water classes (top left), forest/meadow classes (top right), agriculture classes (bottom left) and other classes (bottom right)

4.2 Results

To judge about the performance of different configurations of our framework for classifying hyperspectral EnMAP data, we first consider the general evaluation metrics represented by overall accuracy OA , Cohen's kappa coefficient κ , average completeness \overline{CMP} , average correctness \overline{COR} and average quality \overline{Q} .

In total, we consider nine different configurations of our framework, resulting from all conceivable combinations that are possible when considering three options for classification (1:

NN classifier; 2: LDA classifier; 3: RF classifier) and three options for the data provided to the respective classifier (1: use of the values of all 244 spectral bands per pixel; 2: use of meta-features derived via PCA-based dimensionality reduction; 3: use of the values of those spectral bands that are selected via CFS). The derived classification results are visualized in Fig. 4.

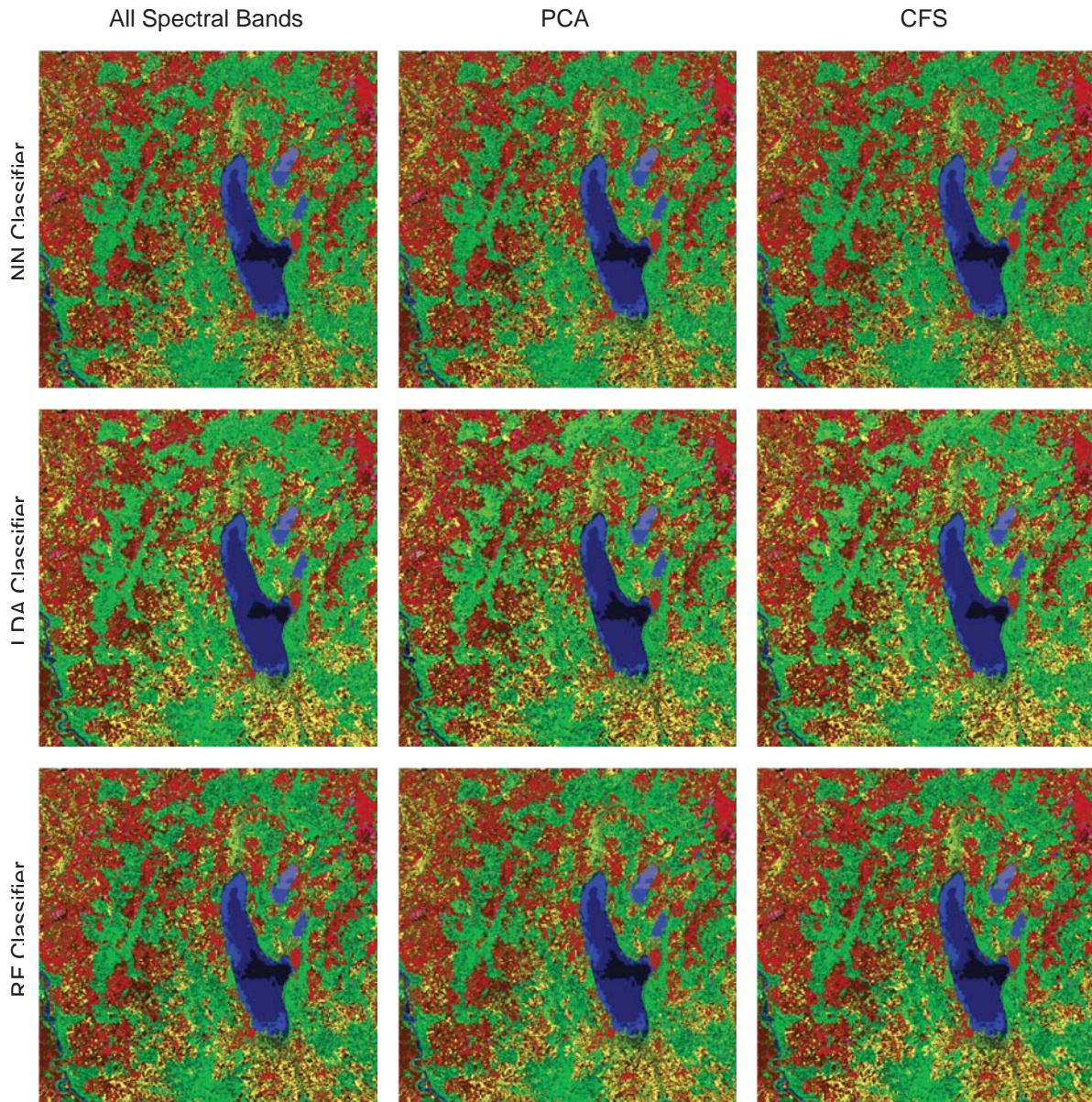


Fig. 4: Classification results obtained for the NN classifier (top row), the LDA classifier (center row) and the RF classifier (bottom row) when using all spectral bands (left column), PCA-based dimensionality reduction (center column) and feature selection via CFS (right column): the color encoding of the different classes is explained in Fig. 3.

The corresponding values for the general evaluation metrics are provided in Tab. 1 and reveal an overall accuracy in the range of about 77-84 % for the nine different configurations of our framework. The kappa value is in the range of about 76-83 % for the considered configurations. The table also contains the average values of completeness, correctness and quality across all 20 classes.

Tab. 1: Classification results obtained for nine different configurations of our framework

Classifier	DR/FS	OA [%]	κ [%]	$\overline{\text{CMP}}$ [%]	$\overline{\text{COR}}$ [%]	\overline{Q} [%]
NN	-	80.07	78.98	80.68	80.04	69.94
NN	PCA	80.96	79.92	81.52	80.99	71.08
NN	CFS	79.89	78.80	80.29	79.95	69.81
LDA	-	78.65	77.48	79.21	78.37	68.18
LDA	PCA	77.40	76.17	77.47	77.11	65.98
LDA	CFS	79.45	78.34	80.32	78.33	68.18
RF	-	83.19	82.27	83.69	83.19	73.70
RF	PCA	82.74	81.80	83.31	82.72	73.30
RF	CFS	82.65	81.71	83.25	82.55	72.94

Whereas the general evaluation metrics allow statements about the complete classification results, the class-wise evaluation metrics of completeness, correctness and quality can be used to assess statistics about the behavior of single classes. With a focus on distinguishing between 20 different land cover and land use classes, we provide the minimum and maximum values of the three class-wise evaluation metrics for all nine configurations of our framework in Tab. 2. It can be observed that the maximum is always reached at 100 % which indicates that some classes are not too hard to distinguish, even when using different configurations of our framework.

Tab. 2: Classification results obtained for nine different configurations of our framework

Classifier	DR/FS	CMP_{\min} [%]	CMP_{\max} [%]	COR_{\min} [%]	COR_{\max} [%]	Q_{\min} [%]	Q_{\max} [%]
NN	-	40.00	100.00	48.00	100.00	27.91	100.00
NN	PCA	45.00	100.00	56.25	100.00	33.33	100.00
NN	CFS	45.00	100.00	45.76	100.00	29.35	100.00
LDA	-	30.00	100.00	42.86	100.00	21.43	100.00
LDA	PCA	16.67	100.00	43.48	100.00	13.70	100.00
LDA	CFS	20.00	100.00	50.00	100.00	16.67	100.00
RF	-	45.00	100.00	58.18	100.00	38.03	100.00
RF	PCA	48.33	100.00	60.00	100.00	36.71	100.00
RF	CFS	48.33	100.00	54.72	100.00	34.52	100.00

4.3 Discussion

As expected, the RF classifier provides the highest quality of the derived classification results. This might be due to the fact that it provides a better generalization capability in case of strongly correlated features. When transforming the original feature vectors via PCA and keeping only the most relevant meta-features that cover 99.9 % of the variability of the given training data, it can be observed that only eight meta-features and therefore a significantly more compact data representation are used, while the quality of the derived classification results is similar with a slight decrease in most cases, but also a slight increase for one case. When using CFS for feature selection, the reflectance values of particular spectral bands are selected. A closer look at the selected bands reveals that reflectance values of 33 spectral bands are considered for classification instead of reflectance values of all 244 spectral bands. Thereby, referring to Fig. 3, 15 spectral bands are considered within the interval [1,50], ten within the interval [51,100], four within the interval [101,150], one within the interval [151,200] and three within the interval [201,244]. Again, the quality of the derived classification results is similar with a slight decrease in most cases and a slight increase for one case.

5 Conclusions

In this paper, we have investigated the potential of hyperspectral EnMAP data for land cover and land use classification. We have presented a framework that involves (1) standard approaches for dimensionality reduction and feature selection and (2) classifiers relying on different learning principles. For nine different configurations of our framework, we have presented the results of a performance evaluation on the benchmark dataset presented in the scope of the recent EnMAP contest, a simulated hyperspectral dataset with 20 different land cover and land use classes. The derived results clearly reveal that a variety of land cover and land use classes can be distinguished based on hyperspectral EnMAP data. Furthermore, we have demonstrated that the transfer of the original, high-dimensional data representation corresponding to 244 spectral bands to a significantly more compact data representation allows to derive classification results of similar and partially even better quality.

6 References

- BRAUN, A. C., WEIDNER, U. & HINZ, S., 2011: Support vector machines, import vector machines and relevance vector machines for hyperspectral classification. Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-4.
- BRAUN, A. C., WEIDNER, U. & HINZ, S., 2012: Classification in high-dimensional feature spaces – Assessment using SVM, IVM and RVM with focus on simulated EnMAP data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5(2), 436-443.
- BRAUN, A. C., WEINMANN, M., KELLER, S., MÜLLER, R., REINARTZ, P. & HINZ, S., 2015: The EnMAP contest: developing and comparing classification approaches for the Environmental Mapping and Analysis Programme – Dataset and first results. The

- International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **XL** (3/W3), 169-175.
- BREIMAN, L., 1996: Bagging predictors. *Machine Learning* **24**(2), 123-140.
- BREIMAN, L., 2001: Random forests. *Machine Learning* **45**(1), 5-32.
- CAMPS-VALLS, G. & BRUZZONE, L., 2005: Kernel-based methods for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing* **43**(6), 1351-1362.
- CAMPS-VALLS, G., TUIA, D., BRUZZONE, L. & BENEDIKTSSON, J. A., 2014: Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. *IEEE Signal Processing Magazine* **31**(1), 45-54.
- CHEHATA, N., LE BRIS, A. & NAJJAR, S., 2014: Contribution of band selection and fusion for hyperspectral classification. *Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 1-4.
- CHERIYADAT, A. & BRUCE, L. M., 2003: Why principal component analysis is not an appropriate feature extraction method for hyperspectral data. *Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS)* **6**, 3420-3422.
- CHI, M., FENG, R. & BRUZZONE, L., 2008: Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem. *Advances in Space Research* **41**(11), 1793-1799.
- FANG, L., LI, S., KANG, X. & BENEDIKTSSON, J. A., 2015: Spectral-spatial classification of hyperspectral images with a superpixel-based discriminative sparse model. *IEEE Transactions on Geoscience and Remote Sensing* **53**(8), 4186-4201.
- FAUVEL, M., CHANUSSOT, J. & BENEDIKTSSON, J. A., 2008: Adaptive pixel neighborhood definition for the classification of hyperspectral images with support vector machines and composite kernel. *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 1884-1887.
- GUANTER, L., SEGL, K. & KAUFMANN, H., 2009: Simulation of optical remote-sensing scenes with application to the EnMAP hyperspectral mission. *IEEE Transactions on Geoscience and Remote Sensing* **47**(7), 2340-2351.
- GUYON, I. & ELISSEEFF, A., 2003: An introduction to variable and feature selection. *Journal of Machine Learning Research* **3**, 1157-1182.
- HALL, M. A., 1999: Correlation-based feature subset selection for machine learning. PhD thesis, Department of Computer Science, University of Waikato, New Zealand.
- HAM, J., CHENG, Y., CRAWFORD, M. M. & GHOSH, J., 2005: Investigation of the random forest framework for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing* **43**(3), 492-501.
- HUGHES, G. F., 1968: On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory* **14**(1), 55-63.
- JOELSSON, S. R., BENEDIKTSSON, J. A. & SVEINSSON, J. R., 2005: Random forest classifiers for hyperspectral data. *Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 160-163.
- KAUFMANN, H., SEGL, K., GUANTER, L., HOFER, S., FOERSTER, K.-P., STUFFLER, T., MUELLER, A., RICHTER, R., BACH, H., HOSTERT, P. & CHLEBEK, C., 2008: *Environmental Mapping*

- and Analysis Program (EnMAP) – Recent advances and status. *Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS)* **4**, 109-112.
- 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 Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 1-6.
- LE BRIS, A., CHEHATA, N., BRIOTTET, X. & PAPANIKOLAOS, N., 2014: Use intermediate results of wrapper band selection methods: a first step toward the optimization of spectral configuration for land cover classifications. *Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 1-4.
- LICCIARDI, G., MARPU, P. R., CHANUSSOT, J. & BENEDIKTSSON, J. A., 2012: Linear versus nonlinear PCA for the classification of hyperspectral data based on the extended morphological profiles. *IEEE Geoscience and Remote Sensing Letters* **9**(3), 447-451.
- MELGANI, F. & BRUZZONE, L., 2004: Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing* **42**(8), 1778-1790.
- PLAZA, A., BENEDIKTSSON, J. A., BOARDMAN, J. W., BRAZILE, J., BRUZZONE, L., CAMPS-VALLS, G., CHANUSSOT, J., FAUVEL, M., GAMBA, P., GUALTIERI, A., MARCONCINI, M., TILTON, J. C. & TRIANNI, G., 2009: Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment* **113**, S110-S122.
- SAEYS, Y., INZA, I. & LARRANAGA, P., 2007: A review of feature selection techniques in bioinformatics. *Bioinformatics* **23**(19), 2507-2517.
- SEGL, K., GUANTER, L., KAUFMANN, H., SCHUBERT, J., KAISER, S., SANG, B. & HOFER, S. 2010: Simulation of spatial sensor characteristics in the context of the EnMAP hyperspectral mission. *IEEE Transactions on Geoscience and Remote Sensing* **48**(7), 3046-3054.
- TARABALKA, Y., CHANUSSOT, J., BENEDIKTSSON, J. A., ANGULO, J. & FAUVEL, M., 2008: Segmentation and classification of hyperspectral data using watershed. *Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS)* **3**, 652-655.
- TARABALKA, Y., FAUVEL, M., CHANUSSOT, J. & BENEDIKTSSON, J. A., 2010: SVM- and MRF-based method for accurate classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters* **7**(4), 736-740.
- TARABALKA, Y. & TILTON, J. C., 2011: Spectral-spatial classification of hyperspectral images using hierarchical optimization. *Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 1-4.
- VAN DER MAATEN, L. J. P., POSTMA, E. O. & VAN DEN HERIK, H. J., 2009: Dimensionality reduction: a comparative review. Technical Report TiCC-TR 2009005, Tilburg University, Tilburg, The Netherlands.
- VILLA, A., BENEDIKTSSON, J. A., CHANUSSOT, J. & JUTTEN, C., 2011: Hyperspectral image classification with independent component discriminant analysis. *IEEE Transactions on Geoscience and Remote Sensing* **49**(12), 4865-4876.
- WANG, J. & CHANG, C.-I., 2006: Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing* **44**(6), 1586-1600.

Bilder vom Boden und vom UAV: Orientierung, dichte 3D-Rekonstruktion und funktionale Modellierung

HELMUT MAYER¹, HAI HUANG¹, ANDREAS KUHN¹, MARIO MICHELINI¹, WILLIAM NGUATEM¹, LUKAS ROTH¹, MATTHIAS SCHMITZ¹ & MARTIN DRAUSCHKE²

Zusammenfassung: Für eine detaillierte Modellierung von Gebäuden und ihrer Umgebung bietet sich die gemeinsame Auswertung von Bildern vom Boden und von kleinen Unmanned Aerial Vehicles (UAVs) an. Die Automatisierung der Auswertung führt zu einer Vielzahl von technischen und wissenschaftlichen Fragestellungen, von denen eine Reihe in diesem Beitrag angesprochen und für die an der Universität der Bundeswehr München und am Institut für Robotik und Mechatronik des Deutschen Zentrums für Luft- und Raumfahrt (DLR) entwickelte Lösungen dargestellt werden. Hierbei wird davon ausgegangen, dass nur Bilder mit ungefähre Kalibrierung, aber keine Navigationsinformation zur Verfügung steht. Die Leistungsfähigkeit der vorgestellten Ansätze für Orientierung, dichte 3D-Rekonstruktion und funktionale Modellierung werden anhand eines aussagekräftigen Beispiels aufgezeigt.

1 Einleitung

Für die detaillierte dreidimensionale (3D) Modellierung von Gebäuden und ihrer Umgebung eignet sich insbesondere eine Kombination von Bildern vom Boden und von kleinen (< 5 kg) Unmanned Aerial Vehicles (UAVs). Diese erlaubt eine Abbildung der Gebäude von allen Seiten (siehe Abb. 1) auch aus unterschiedlichen Höhen sowie von oben und von schräg oben. Sie führt damit zu hoch aufgelöster Bildinformation aus frontaler Perspektive sowohl für Wände als auch für Dächer. Aus diesen sind extrem detaillierte 3D-Modelle ableitbar, die für Stadtplanung, Architektur, Navigation, Tourismus, Filme und auch für Einsatzplanung im urbanen Raum geeignet sind.

Der vorgestellte Ansatz ist sehr allgemein gehalten und geht für die Bilder nur davon aus, dass eine ungefähre Kalibrierung, wie sie aus den EXIF Daten für fast alle Kameras verfügbar ist, aber keine Navigationsinformation zur Verfügung steht. Damit kann für die Orientierung der 5-Punkt-Algorithmus von NISTÉR (2004) verwendet werden. In Kombination mit RANdOm SAMple Consensus – RANSAC (FISCHLER & BOLLES 1981), kleinster Quadrate Bildzuordnung und robuster Bündelausgleichung erhält man einen Algorithmus, der hoch genaue Orientierungen auch für große Basen bestimmen kann. Die hierfür notwendige Information über die Verknüpfung der Bilder in Form von Bildtriplets wird mittels des Ansatzes von (MICHELINI & MAYER 2016) bestimmt.

Auf Grundlage der hoch genauen Orientierung wird eine dichte 3D-Rekonstruktion mit dem Ansatz von KUHN et al. (2016a) durchgeführt. Aus den mittels SemiGlobalem Matching – SGM (HIRSCHMÜLLER 2008) berechneten Tiefenkarten wird die Genauigkeit der Zuordnung für Paare abgeschätzt und in den 3D-Raum projiziert. Die Information für verschiedene Paare wird fusioniert und daraus hoch detaillierte 3D-Oberflächen abgeleitet. Da die Fusion lokal begrenzt ist, kann sie

¹ Universität der Bundeswehr München, Institut für Angewandte Informatik, Werner-Heisenberg-Weg 39, D-85577 Neubiberg, E-Mail: [Helmut.Mayer, Hai.Huang, Andreas.Kuhn, Mario.Michelini, Lukas.Roth, Matthias.Schmitz]@unibw.de

² Deutsches Zentrum für Luft- und Raumfahrt, Institut für Robotik und Mechatronik, Münchener Straße 20 D-82234 Oberpfaffenhofen-Weßling, E-Mail: Martin.Drauschke@dlr.de

parallel durchgeführt werden, was die Skalierbarkeit der 3D-Rekonstruktion auf große Mengen von hoch aufgelösten Bildern ermöglicht. Durch die Abschätzung der Zuordnungsgenauigkeit wird trotz der lokalen Begrenzung eine hohe Qualität erzielt.

Die bei der 3D-Rekonstruktion entstehenden 3D-Modelle sind extrem detailliert, damit allerdings auch bezüglich des Speicherplatzes sehr umfangreich und besitzen keine Semantik. Daher wird in der funktionalen Modellierung nach Objekten und deren Teilen, hier Gebäude mit Fassaden, Dächern, Fenstern und Türen gesucht. Grundlage ist eine effiziente Klassifikation, der bei der dichten 3D-Rekonstruktion entstehenden farbigen Punktwolke (KUHNER et al. 2016b) sowie die Berechnung zusammenhängender Flächenelemente (NGUATEM & MAYER 2016), die zu Dach-, Fassaden- und Bodenflächen führen. Die Dachflächen werden mit statistischen Verfahren klassifiziert (Giebeldach, Walmdach, Mansarddach etc.). Türen und Fenster werden als Strukturen hinter der Fassade bestimmt (NGUATEM et al. 2016). Parallel dazu werden in entzerrten Bildern der Fassade mittels eines Convolutional Networks Fenster und Türen klassifiziert und segmentiert (SCHMITZ & MAYER 2016).

Im Weiteren werden die einzelnen Ansätze für Orientierung, dichte 3D-Rekonstruktion und funktionale Modellierung im Detail vorgestellt.

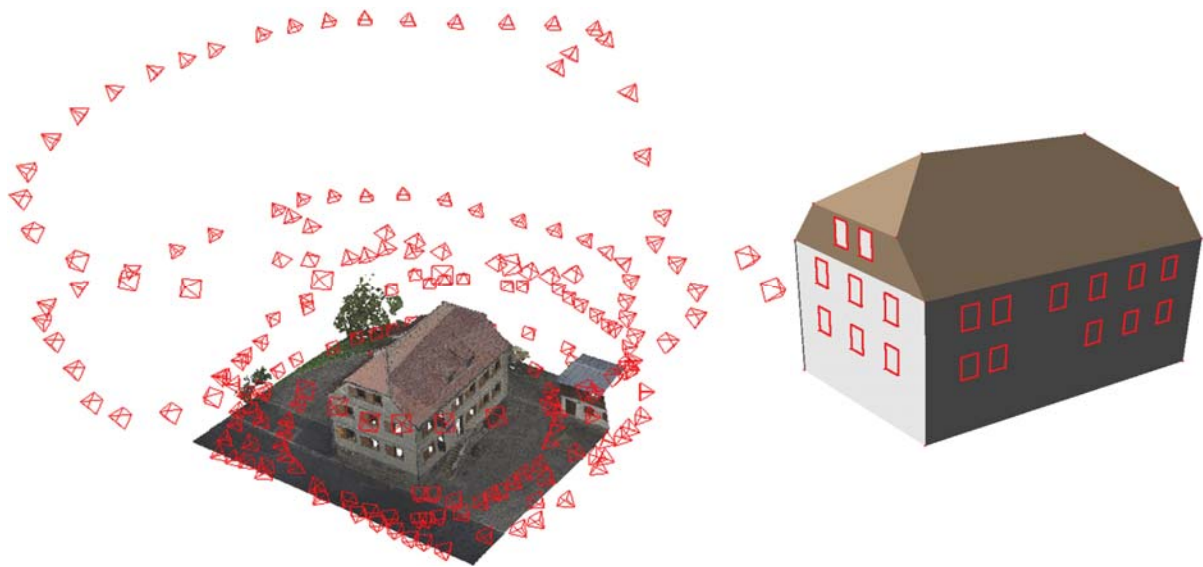


Abb. 1: Dichte 3D-Rekonstruktion eines Gebäudes inkl. der als Pyramiden dargestellten Orientierungen der Kameras (links) und funktionale Modellierung des Gebäudes mit Wänden (grau), Dach (braun) und Fenstern (rot umrandet) (rechts)

2 Orientierung

Die (relative) Orientierung der genähert kalibrierten Bilder erfolgt auf Grundlage des 5-Punkt-Algorithmus sowie RANSAC. Für große Basen ergeben sich wenig Inlier für RANSAC. Insbesondere sind die mittels 5-Punkt-Algorithmus bestimmten Lösungen auch für fünf Inlier oft nur lokal, d.h. für Teile der maximal möglichen Inlier repräsentativ. Deswegen wird RANSAC mit robuster Bündelausgleichung in Form eines an Expectation Maximization (EM) angelehnten Ansatzes (MAYER et al. 2012) kombiniert, bei dem lokale Lösungen schrittweise erweitert werden

und womit eine stabile Lösung auch für große Basen möglich wird. Diese Vorgehensweise wird für Paare und für Triplets verwendet. Triplets werden hierarchisch zu Bildmengen vereinigt (MAYER 2014).

Die effiziente automatisierte Bestimmung der Verknüpfung für größere Bildblöcke erfolgt auf Grundlage von Merkmalsextraktion auf der Graphikkarte. Das zentrale Problem für größere Mengen von (ungeordneten) Bildern ist der exponentielle Anstieg der Paare. Um mit diesem umzugehen, findet eine Reduktion des bei der Bestimmung der Verknüpfung entstehenden Graphen statt. Da eine Übertragung der Skalierung der Orientierung nur für Triplets möglich ist und auch nur diese für große Basen eine stabile Orientierung ermöglichen, wurden zudem Baumstrukturen (Steinerbäume) für Triplets verwendet, die eine (näherungsweise) minimale Verbindung garantieren. Nach Verifikation von Paaren und Triplets mittels des Ansatzes von (MAYER et al. 2012) und einer evtl. notwendigen Neuberechnung der Verknüpfung erfolgt zuletzt eine gezielte Verdichtung der Netzstruktur zur Stabilisierung der Geometrie (MICHELINI & MAYER 2016).

Die in (MICHELINI & MAYER 2016) beschriebene Vorgehensweise eignet sich für große Basen. Die Zuordnung von Paaren mit extrem großen Basen ist wünschenswert, im Fall unbekannter Verknüpfung aber unrealistisch, da erst nach der rechenaufwändigsten Variante der Zuordnung klar wird, ob ein Paar eine extreme Basis besitzt oder einfach nicht denselben Ausschnitt der Welt zeigt. Je größer die Bildmenge ist, desto größer wird die Wahrscheinlichkeit für Letzteres. Für kleine Mengen (z.B. < 20) oder bei manueller Vorgabe ist allerdings eine Zuordnung von Bildern mit extremer Basis sinnvoll. Hierzu wird in (ROTH et al. 2017) ein Ansatz vorgestellt, bei dem projektiv verzerrte Bilder simuliert werden. In den verzerrten Bildern werden verschieden aufwändige Merkmalsdetektoren angewandt, die Merkmale werden zugeordnet und zuletzt die Zuordnungen mittels 5-Punkt-Algorithmus überprüft. Abb. 2 demonstriert, dass mit der vorgestellten Vorgehensweise auch perspektiv extrem verzerrte Bilder mit kleinem Überlappungsbereich erfolgreich zugeordnet und orientiert werden können.

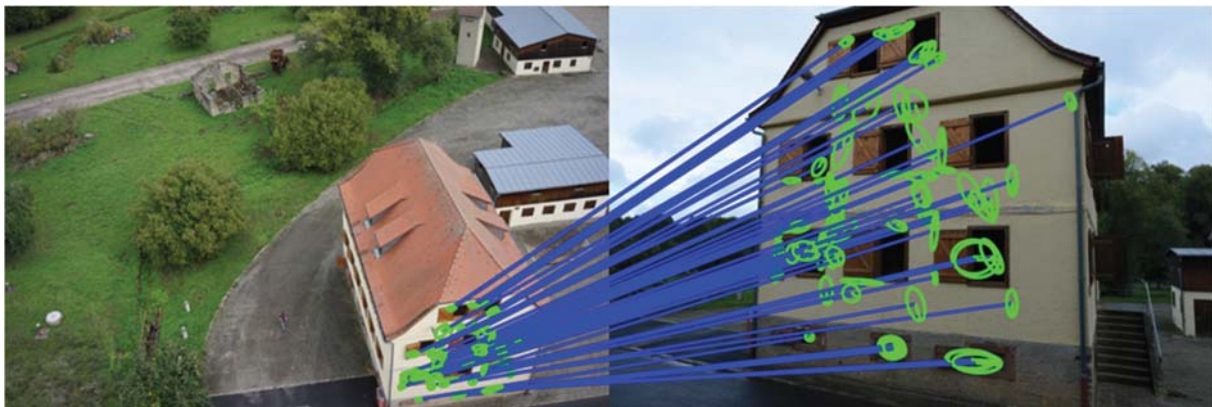


Abb. 2: Zuordnung von Punkten zwischen Bildern mit extrem großer Basis. Grün: Punkte mit Größe und Ausrichtung; blau: Zuordnung.

3 Dichte 3D-Rekonstruktion

Die hoch genaue Orientierung inkl. Genauigkeitsinformation aus der Bündelausgleichung ist die Grundlage für die dichte 3D-Rekonstruktion (KUHN et al. 2016a). Hierfür werden mit SGM für Bildpaare Disparitätskarten berechnet. Für die Disparität wird die Unsicherheit auf Grundlage eines auf Total Variation (TV) basierenden Maßes geschätzt. Die Zuordnungsfunktion wurde aus Ground Truth Daten mittels eines Lernansatzes unter Nutzung von EM bestimmt. Die Unsicherheit wird in den 3D-Raum propagiert und führt zu Wahrscheinlichkeiten für Voxel, dass diese vor bzw. hinter einer Oberfläche liegen. Die Ergebnisse verschiedener Paare werden mittels binärer Bayes Fusion kombiniert und die Oberfläche als der wahrscheinlichste Ort zwischen den Voxeln vor und hinter der Oberfläche bestimmt.

Der Ansatz wird skalierbar, indem der 3D-Raum mittels eines Octrees dynamisch unterteilt wird (KUHN & MAYER 2015). Da die Fusion lokal begrenzt durchgeführt wird, reicht eine vorgegebene Überlappung aus, um ein global konsistentes Ergebnis zu erzielen. Weil die Fusion in den unterschiedlichen Zweigen eines Octrees unabhängig durchgeführt werden kann, ist der Ansatz parallelisierbar. Nach Filterung bezüglich Sichtbarkeit wird die Oberfläche zuletzt trianguliert.

Der Ansatz ist durch die Parallelisierbarkeit gut für Rechner mit vielen Kernen und für Rechnercluster geeignet. Durch Nutzung der Genauigkeitsinformation der Zuordnung führt er zudem zu Ergebnissen mit hoher Qualität. Dies gilt im Gegensatz zu vielen anderen Ansätzen auch für große Mengen von Bildern (hunderte oder tausende) mit hoher Auflösung (50 Megapixel und mehr).

Abb. 3 stellt zwei schattierte Ansichten des Gebäudes aus Abb. 1 dar. Insbesondere zeigt die rechte Ansicht eine Treppe, bei dem die dünnen Stangen des Geländers 3D rekonstruiert wurden. Dies zeigt die Detaillierung der 3D-Rekonstruktion.

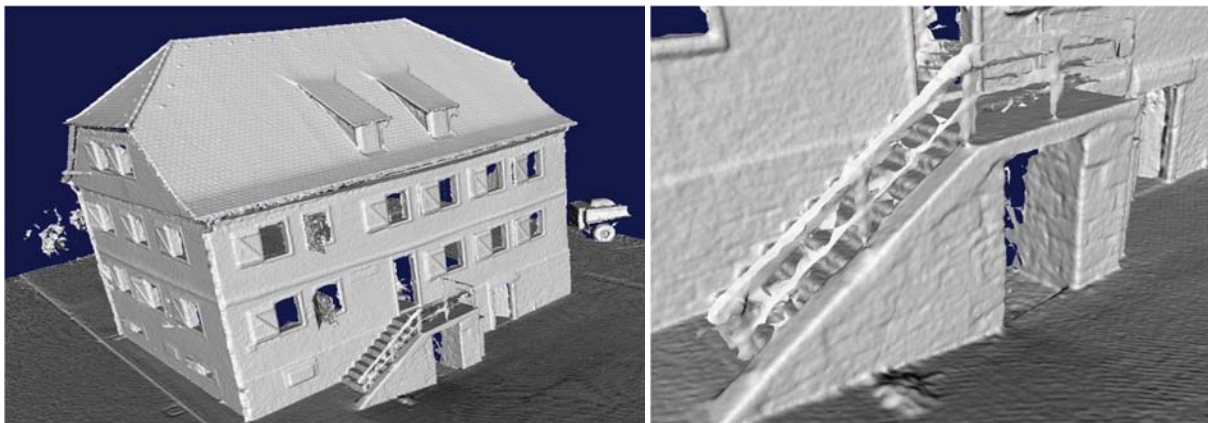


Abb. 3: Schattierte Ansichten der 3D-Rekonstruktion des Gebäudes aus Abb. 1. Der Ausschnitt rechts zeigt die 3D-Rekonstruktion der dünnen Stangen des Geländers.

4 Funktionale Modellierung

Die Modelle der 3D-Rekonstruktion zeigen kleinste Details, benötigen aber auch große Mengen an Speicherplatz (für das in Abb. 1 und 3 dargestellte Modell sind es für 23,6 Millionen Dreiecke bei binärer Speicherung ein halbes Gigabyte). Darüber hinaus besitzen die Dreiecke keine Semantik. In der funktionalen Modellierung werden daher Objekte und ihre Teile, hier Gebäude mit Fassaden, Dächern, Fenstern und Türen bestimmt.

Der erste Schritt besteht in einer Klassifikation der bei der dichten 3D-Rekonstruktion entstehenden farbigen Punktwolke in die Klassen Boden, Gebäude, Gras und Bäume (KUHNS et al. 2016b). Der Ansatz ist effizient, weil er ausnutzt, dass für die Klassifikation keine hoch aufgelöste 3D-Oberfläche notwendig ist. Mittels Occupancy Grids wird eine sehr schnelle Fusion von 3D-Punkten inkl. Wahrscheinlichkeiten und Farbwerten möglich. Eine weitere Beschleunigung wird durch die Nutzung von (erweiterten) Superpixeln erzielt, die mittels Segmentierung bezüglich Farbe, Höhe, und Neigung abgeleitet werden (HUANG & MAYER 2015). Durch die Nutzung relativer Merkmale ist der Ansatz nicht nur robust bezüglich Änderungen in den Umgebungsbedingungen, sondern auch parallelisierbar.

Die Klassifikation führt zu Bereichen mit einer hohen Wahrscheinlichkeit für Gebäude. In solchen Bereichen werden zusammenhängende Flächenelemente mittels eines Clustering Verfahrens berechnet (NGUATEM & MAYER 2016). Die Schätzung der Flächennormalen für Voxel auf Grundlage von RANSAC macht den Ansatz robust und effizient. Eine Dirichlet-Prozess Mixtur (DPM) Modell von Gauß-Funktionen vermeidet die Vorgabe der Zahl der Flächen. Zuletzt werden zusammenhängende Flächen bestimmt.

Aus den zusammenhängenden Flächen werden mittels statistischer Verfahren Dach, Fassaden und Boden abgeleitet. Für die Dachflächen führen Sequentielles Monte Carlo Sampling und Modellselektion zur Unterscheidung von Giebeldach, Walmdach und Mansarddach (NGUATEM et al. 2013). Das Dach wird um vertikale Fassaden ergänzt und Fenster sowie Türen werden als Strukturen hinter den Fassaden bestimmt. Hierzu werden verschiedene Hypothesen für die Fassadenebene gebildet, die zu unterschiedlichen Segmentierungen der 3D-Punkte führen. Auf deren Grundlage erfolgt eine probabilistische Musterzuordnung inkl. Modellselektion (NGUATEM et al. 2014). Einen Überblick über alle Schritte der funktionalen Modellierung von Gebäuden gibt (NGUATEM et al. 2016).

In einem parallel zu den statistischen Verfahren für 3D-Punkte entwickelten Ansatz (SCHMITZ & MAYER 2016) werden in entzerrten Bildern der Fassade mittels eines Convolutional Networks Fenster und Türen klassifiziert und segmentiert. Für Letzteres werden neben „Convolutional“ auch „Deconvolutional“ Verknüpfungen verwendet, die die Auflösung nach der Klassifikation vergrößern. Die Nutzung der Bildinformation macht den Ansatz unabhängig von Löchern und unscharfen Kanten in der Punktwolke.



Abb. 4: 3D-Projektion der Ergebnisse der Interpretation von entzerrten Fassadenbildern mittels eines Convolutional Networks mit Fenstern (rot) und Türen (blau) für das Gebäude aus Abb. 1. Die Fassadenebenen wurden mit Hilfe des Ansatzes von NGUATEM & MAYER (2016) bestimmt.

5 Fazit & Ausblick

Der Beitrag gibt einen Überblick über verschiedene an der Universität der Bundeswehr München und dem Institut für Robotik und Mechatronik des Deutschen Zentrums für Luft- und Raumfahrt (DLR) entwickelte Ansätze, die von der Orientierung über die dichte 3D-Rekonstruktion bis hin zur funktionalen Modellierung von Gebäuden reichen.

In der Zukunft sollen die Ansätze für die funktionale Modellierung in Richtung komplexerer Gebäude sowie weiterer Objektklassen erweitert werden. Für eine Verbesserung der Zuverlässigkeit sollen die Ergebnisse für die statistischen Verfahren mit denen der Convolutional Networks kombiniert werden.

6 Literaturverzeichnis

- FISCHLER, M. & BOLLES, R., 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.
- HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30**(2), 328-341.
- HUANG, H. & MAYER, H., 2015: Robust and Efficient Urban Scene Classification Using Relative Features. *23rd SIGSPATIAL International Conference on Advances in Geographic Information System*, 81:1-81:4.
- KUHN, A., HIRSCHMÜLLER, H. & MAYER, H., 2013: Multi-Resolution Range Data Fusion for Multi-View Stereo Reconstruction. *German Conference on Pattern Recognition*, 41-50.
- KUHN, A., HIRSCHMÜLLER, H., SCHARSTEIN, D. & MAYER, H. 2016a: A TV Prior for High-Quality Scalable Multi-View Stereo Reconstruction. *International Journal of Computer Vision*, Online First.
- KUHN, A., HUANG, H., DRAUSCHKE, M. & MAYER, H., 2016b: Fast Probabilistic Fusion of 3D Point Clouds via Occupancy Grids for Scene Classification. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences (III-3)*, 325-332.
- KUHN, A. & MAYER, H., 2015: Incremental Division of Very Large Point Clouds for Scalable 3D Surface Reconstruction. *International Conference on Computer Vision Workshop (ICCVW)*, 4321-4329.
- KUHN, A., MAYER, H., HIRSCHMÜLLER, H. & SCHARSTEIN, D., 2014: A TV Prior for High Quality Local Multi-View Stereo Reconstruction. *2nd International Conference on 3D Vision (3DV)*, 65-72.
- MAYER, H., 2014: Efficient Hierarchical Triplet Merging for Camera Pose Estimation. *German Conference on Pattern Recognition*, 399-409.
- MAYER, H., BARTELTSEN, J., HIRSCHMÜLLER, H. & KUHN, A., 2012: Dense 3D Reconstruction from Wide Baseline Image Sets. *Outdoor and Large-Scale Real-World Scene Analysis – 15th International Workshop on Theoretical Foundations of Computer Vision, Lecture Notes in Computer Science 7474*, Springer, Berlin, 285-304.
- MAYER, H. & MICHELINI, M., 2016: Orientierung großer Bildverbände. *Handbuch der Geodäsie, Band Photogrammetrie und Fernerkundung*, 32 Seiten.

- MICHELINI, M. & MAYER, H., 2016: Efficient Wide Baseline Structure from Motion. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences (III-3)*, 99-106.
- NGUATEM, W., DRAUSCHKE, M. & MAYER, H., 2013: Roof Reconstruction from Point Clouds Using Importance Sampling, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences II (3/W3)*, 73-78.
- NGUATEM, W., DRAUSCHKE, M. & MAYER, H., 2014: Localization of Windows and Doors in 3D Point Clouds of Facades. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences II (3)*, 87-94.
- NGUATEM, W., DRAUSCHKE, M. & MAYER, H., 2016: Automatic Generation of Building Models with Levels of Detail 1-3. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (XLI-B3)*, 649-654.
- NGUATEM, W. & MAYER, H., 2016: Contiguous Patch Segmentation in Pointclouds. *German Conference on Pattern Recognition*, 131-142.
- NISTÉR, D., 2004: An Efficient Solution to the Five-Point Relative Pose Problem, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(6), 756-770.
- ROTH, L., KUHN, A. & MAYER, H., 2017: Wide-Baseline Image Matching with Projective View Synthesis and Calibrated Geometric Verification. *Photogrammetrie – Fernerkundung – Geoinformation*, angenommen.
- SCHMITZ, M. & MAYER, H., 2016: A Convolutional Network for Semantic Facade Segmentation and Interpretation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (XLI-B3)*, 649-654.

Innenraumrekonstruktion aus semantisch angereicherten 3D Punkten und Linien

DOROTA IWASZCZUK¹, TOBIAS KOCH¹ & UWE STILLA¹

Zusammenfassung: In diesem Artikel präsentieren wir ein Konzept zur semantikbasierten Innenraumrekonstruktion aus überlappenden Bildserien. Herkömmliche Verfahren zur Erzeugung von bildbasierten 3D Punkten und 3D Linien erschweren besonders in schwach texturierten Innenräumen die Detektion von Begrenzungsflächen und resultieren meist in unvollständige Gebäudemodelle. Durch eine semantische Interpretation dieser 3D Strukturen können relevante 3D Objekte identifiziert werden um ein vollständiges 3D Raummodell zu erzeugen. Unsere Methode basiert auf einer pixelweisen semantischen Segmentierung der Bilder deren Klasseninformationen auf 3D Strukturen, wie 3D Punkte und 3D Linien, übertragen werden. Dabei werden den 3D Punkten und 3D Linien die Signaturen aus der Bildklassifikation zugewiesen. Aus diesen signierten Merkmalen werden die Innenräume eines Gebäudes rekonstruiert. Dabei werden die Hypothesen für die zu rekonstruierende Ebenen aus 3D Linien einer Klasse gebildet und mit den 3D Punkten derselben Klasse unterstützt. Die optimale Konfiguration aus den Ebenen wird unter Berücksichtigung der Orthogonalität und Parallelität iterativ gesucht.

1 Einleitung

Aufgrund der Urbanisierung und der damit einhergehenden stetigen Zunahme der Einwohneranzahl in städtischen Bereichen steigt die Nachfrage städtische Gebiete mit hohem Detaillierungsgrad zu modellieren. Von besonderem Interesse sind dabei Gebäude, da sich Menschen überwiegend in Gebäuden aufhalten und deren Aktivitäten dort konzentrieren.

In den letzten Dekaden sind zahlreiche Methoden entstanden die eine automatische 3D Gebäuderekonstruktion ermöglichen. Die Detaillierungsgrade erstrecken dabei von einfachen Klötzchenmodellen (LOD 1), über erweiterte Modelle mit zusätzlicher Modellierung von Dachstrukturen und einfachen Texturierungen (LOD 2), bis hin zu komplexen 3D Modellen mit vollständigen und texturierten Fassadenstrukturen (LOD 3). Ein aktueller Trend besteht in der Gebäudeinnenraumrekonstruktion, die zu einer Generierung von LOD 4 Gebäudemodellen dienen können. Besonders in den letzten Jahren wird diesem Bereich größere Aufmerksamkeit in der Forschung und Praxis gewidmet (COHEN et al. 2016).

Moderne Methoden aus dem Bereich der Computer Vision (SNAVELY et al. 2007; ENGEL et al. 2014; ENGEL et al. 2016) verfügen über robuste Verfahren um 3D Geometrien als fotorealistische 3D Punktwolken zu rekonstruieren, und die Objekte effektiv in 3D darzustellen. Bisher steht allerdings meist die visuelle Betrachtung von solchen Datensätzen im Vordergrund, während automatische Analysen dieser Daten oft vernachlässigt werden. Erschwerend für die Rekonstruktion von Innenräumen ist die oftmals fehlende oder schwach ausgeprägte Textur

¹ Technische Universität München, Ingenieurfacultät Bau Geo Umwelt, Arcisstr. 21, D-80333 München, E-Mail: [Dorota.Iwaszczuk, Tobias.Koch, Stilla]@tum.de

besonders bei Wänden und Decken. Dies resultiert meist in unvollständige Innenraummodelle, die eine geometrische als auch semantische Interpretation erschwert.

Innenraumrekonstruktion ist ein großer Bestandteil der Forschung mit mobilen RGBD-Kameras (WANG et al. 2016; NEVEROVA et al. 2013; CHOI et al. 2015). Die Verwendung dieser aktiven Sensoren ermöglicht auch bei texturlosen Oberflächen das Erzeugen von sehr dichten 3D Punktwolken, wodurch die Schätzung der Raumbegrenzungsflächen einfacher ist. In diesem Artikel beschränken wir uns jedoch auf die Verwendung von herkömmlichen RGB Kameras, die einen größeren Einsatzbereich unserer Methode abdecken.

Ein Grund für die meist unzufrieden stellenden Ergebnisse einer Innenraumrekonstruktion auf Basis von RGB Bildern ist, dass meist keine semantische Information verwendet wird und die als Punktwolken gespeicherte Geometrien keine Zusammenhänge darstellen. Um solche Analysen zu ermöglichen, sollen die im Innenraumbereich erzeugten Geometrien mit existierenden 3D Gebäudemodellen kombiniert werden. Dafür wird typischerweise eine Repräsentation als Polygone benötigt. Semantische Informationen, die aus den Bildern gewonnen werden, sollen nun zur besseren und vollständigeren Repräsentation von Innenraummodellen verwendet werden.

2 Konzept

In diesem Beitrag präsentieren wir ein Konzept zur 3D Polygonrekonstruktion von Gebäudeelementen aus Bilderserien von Innenräumen. Eine Übersicht über den Workflow der Methode ist in Abb. 1 dargestellt. Aus einer Vielzahl von Bildern des Innenraumes werden zum einen 3D Strukturen berechnet und zum anderen eine pixelweise semantische Segmentierung der Bilder vorgenommen. Die 2D Klassifikationsergebnisse werden anschließend auf die 3D Strukturen übertragen. Von besonderem Interesse sind hierbei Elemente, wie Wände, Decken, Böden, Türen und Fenster, während Mobiliar, Personen und andere Störobjekte vom Rekonstruktionsprozess ausgeschlossen werden. Aus den verbleibenden 3D Strukturen werden mehrere Ebenenhypothesen geschätzt und mithilfe zusätzlicher Bedingungen eine sinnvolle Konfiguration dieser Ebenen bestimmt.

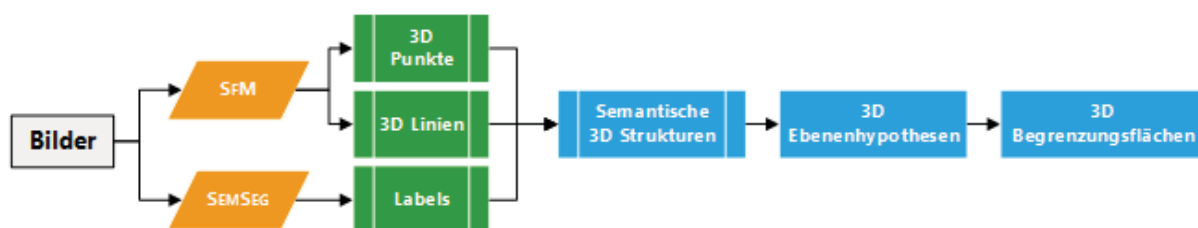


Abb. 1: Workflow zur semantischen Innenraumrekonstruktion. Aus einem Set von überlappenden Bildern werden Kamerapositionen und anschließend 3D Punkte und 3D Linien berechnet. Die Klassifikationsergebnisse einer semantischen Segmentierung aller Bilder in Bezug auf Raumbegrenzungsflächen (Wand, Boden, Decke, Tür, Fenster) werden auf die 3D Strukturen übertragen. Nach der Eliminierung irrelevanter Strukturen wird eine Vielzahl von 3D Ebenenhypothesen aufgestellt. Eine ML-Methode unter Berücksichtigung geometrischer Bedingungen wird zur Bestimmung einer sinnvollen Ebenenkonfiguration angewendet.

2.1 2D semantische Segmentierung

Um eine semantische Segmentierung der Bilder durchzuführen, stehen heutzutage zahlreiche Methoden zur Verfügung (CIREŞAN et al. 2011; GERKE & XIAO 2013; ARMENI et al. 2016). Großes Potential birgt hierfür ein überwachter Klassifikator in Form eines konvolutionalen neuronalen Netzes, um eine pixelweise Klassifikation aller Eingabebilder mit den gewünschten Klassen zu ermöglichen. Als Trainingsdaten für unsere Problemstellung stehen öffentlich zugängliche Datensätze zur Verfügung, wie z.B. in (LIU et al. 2015) mit über 1500 annotierten Bildern der Klassen Fenster und Türen von unterschiedlichen Raum- und Bebauungstypen. Da insgesamt jedoch nur eine geringe Anzahl von Datensätze für die Innenraumklassifikationen zur Verfügung steht, müssen eigens annotierte Bilder zur Verbesserung der Klassifikationsergebnisse hinzugefügt werden.

Eine vielversprechende Methode zur pixelweisen semantischen Segmentierung stellt *SegNet* dar (KENDALL et al. 2015). Neben einer pixelweisen Klassifikation von unbekanntem Eingabebildern werden durch die Verwendung eines Softmax-Klassifikators in der letzten Schicht des konvolutionalen neuronalen Netzwerkes ebenfalls Wahrscheinlichkeitsverteilungen aller Klassen pro Pixel ausgegeben. Dies ermöglicht die Ausgabe von Unsicherheiten für alle klassifizierten Pixel und führt im nächsten Schritt zu robusteren Klassifikationsergebnissen der 3D Strukturen. Für unsere Anwendung müsste dieses Netz nochmals mit einer Vielzahl neuer Daten und Klassen antrainiert werden. Abb. 2 stellt beispielhaft die Ergebnisse der semantischen Bildsegmentierung für zwei Bilder dar.



Abb. 2: Beispielergebnisse der 2D semantischen Segmentierung. Jedes Bild wird mit einem vorher antrainierten Klassifikator pixelweise segmentiert. Die Klassen Wand (grün), Boden (grau), Decke (schwarz), Tür (rot) und Fenster (cyan) werden für die Bestimmung der Raumbegrenzungsflächen verwendet während die letzte Klasse Objekt (blau) zur Eliminierung unbedeutender Störobjekte dient.

2.2 Initiale 3D Modellierung der Innenräume

Für die 3D Rekonstruktion von Innenräumen werden zunächst herkömmliche Verfahren aus den Bereichen Photogrammetrie und Computer Vision verwendet. Aus einer Reihe von überlappenden Bildern des zu modellierenden Raumes werden zunächst mithilfe von *Structure-from-Motion* (SfM) Verfahren interne und externe Orientierungen der einzelnen Bilder berechnet die im nächsten Schritt zur Generierung von 3D Strukturen in Form von 3D Punkten und 3D Linien dienen. Aus einer Vielzahl bestehender SfM Implementierungen wird hier *VisualSfM* (WU 2011) verwendet.

2.2.1 3D Punkte

Die 3D Rekonstruktion erfolgt zunächst durch eine dichte 3D Punktwolke. Hierbei steht mit *Semi-Global-Matching* (SGM) weiterhin eine aktuelle Methode für das Erzeugen einer dichten 3D Punktwolke zur Verfügung. Hierfür werden mithilfe der geschätzten internen und externen Kameraparametern aus dem SfM Schritt und den Eingabebildern pixelweise Korrespondenzen in Bildpaaren gesucht, die anschließend trianguliert werden um eine dichte Punktwolke zu erzeugen. Dies setzt jedoch gut texturierte Oberflächen in den Bildern voraus, welches für Innenräume durch oftmals weiße Wände nicht immer gegeben ist. Als Folge entstehen oft große Lücken in den 3D Punktwolken an den Wänden und Decken des Raumes, welches das Bestimmen von 3D Begrenzungsflächen der Räume erschwert. Eine dichte Punktwolke eines Büros ist in Abbildung 3 dargestellt. Für die Generierung der dichten Punktwolke wird die Methode von (ROTHERMEL et al. 2012) verwendet.

2.2.2 3D Linien

Zur Unterstützung der 3D Strukturen werden neben einer dichten 3D Punktwolke ebenfalls 3D Linien rekonstruiert. Auch hierfür werden die Kameraparameter und Ursprungsbilder verwendet um korrespondierende 2D Linien in Bildpaaren zu identifizieren und anschließend zu triangulieren. Obwohl auch hier an schwach texturierten Flächen nur wenige Linienkorrespondenzen gefunden werden können, ermöglichen wenige 3D Linien eine robustere Schätzung von Begrenzungsflächen wie Wände oder Decken als die verlässlichen 3D Punkte (Abb. 3). Für die Erstellung der Linien wird die Methode von (HOFER et al. 2016) verwendet.

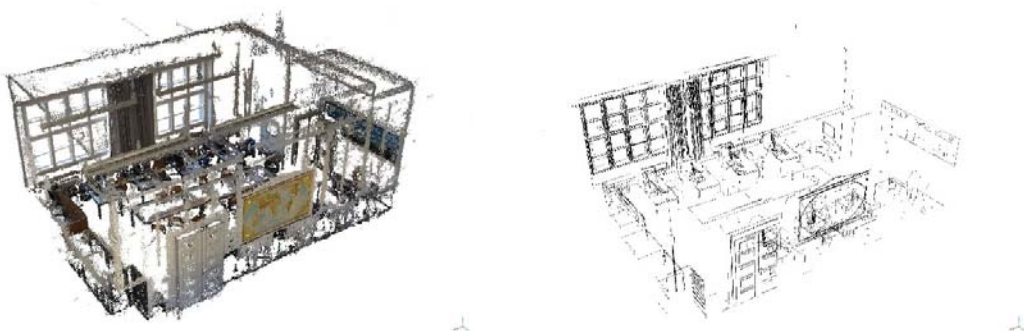


Abb. 3 Erzeugte 3D Strukturen eines Büros. Links: dichte 3D Punktwolke und rechts: 3D Linien

2.3 Verwendung der semantischen Signaturen in 3D Strukturen

Die im vorherigen Schritt erzeugten 3D Punkte und Linien enthalten neben den für die Ermittlung von Begrenzungsflächen relevanten Strukturen (wie Wände, Türen oder Fenster) vorwiegend Strukturen von zunächst unbedeutenden Objekten wie Mobiliar oder anderen Gegenständen. Da diese für die Bestimmung der Raumstruktur hinderlich sind kann die semantische Information aus dem ersten Schritt herangezogen werden.

Da sowohl die 3D Punkte als auch 3D Linien aus korrespondierenden 2D Punkten bzw. 2D Linien in Bildpaaren generiert wurden und für jedes Bild eine pixelweise Klassifikation vorliegt, können

den 3D Strukturen diese semantischen Informationen aus den Segmentierungsergebnissen zugewiesen und als Label gespeichert werden.

Für jeden 3D Punkt wird zunächst die jeweilige Klasse und Klassifikationsgenauigkeit des korrespondierenden 2D Punktes aus allen Bildern in denen dieser 3D Punkt gesehen worden ist gespeichert. Schließlich wird aus der Liste aller Klassifikationsergebnissen inklusive Genauigkeiten eine finales Label geschätzt. Für die Klassenzuordnung der 3D Linien wird ähnlich verfahren, indem jede 3D Linie in diejenigen Bilder, in denen diese 3D Linie sichtbar ist, projiziert wird. Diese 2D Linie gilt oftmals als Übergang zweier Klassen und somit werden für jede 3D Linie zwei Klassen gespeichert, und zwar diejenigen, die in der lokalen Nachbarschaft auf beiden Seiten der 2D Linie vorkommt.

Ein Beispiel dieser Klassenzuordnung ist in Abbildung 4 dargestellt.

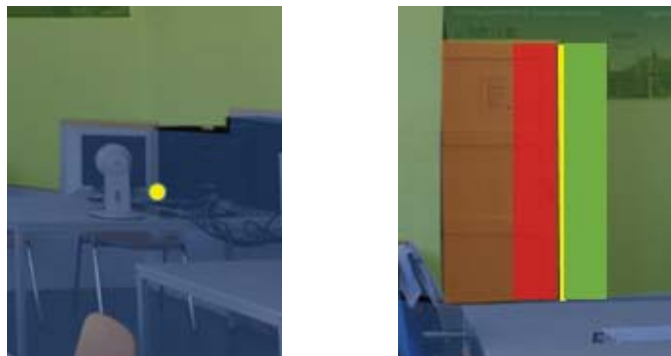


Abb. 4: Projektion (gelb) eines 3D Punktes (links) und einer 3D Linie (rechts) in ein Bild. Links: Klassenzuordnung des 3D Punktes entsprechend des Klassifikationsergebnisses an dieser Bildstelle (hier: blau für Objekt). Rechts: Zwei Zuordnungen entsprechend beider Seiten 2D Linie (hier: rot für Tür und grün für Wand).

Da für die Bestimmung der Begrenzungsflächen von Innenräumen lediglich 3D Strukturen mit den Klassen Wand, Boden, Decke, Fenster oder Tür von Bedeutung sind, werden zunächst alle 3D Strukturen, die keine dieser Klassen beinhalten, entfernt.

2.4 Rekonstruktion der Innenräume

Aus der dichten 3D Punktwolke und der 3D Linien sollen nun Polygone geschätzt werden, die einen geschlossenen Raum bilden. Es wird angenommen, dass die Wände, Decken und Böden orthogonal und parallel zueinander stehen und sich in den Raumecken schneiden. Der Vorteil der gemeinsamen Verwendung von 3D Punkten und 3D Linien ist, dass die Punkte eine dichtere Rekonstruktion der Flächen liefern, die Linien dagegen häufig in Raumecken extrahiert und rekonstruiert werden. Dies erlaubt es die Begrenzungsebenen robuster und akkurater zu schätzen. Zusätzlich kann in Innenräumen häufig von einer Manhattan-World Annahme ausgegangen werden in der die Begrenzungsflächen entlang der drei Hauptrichtungen verlaufen. Diese Annahme wird auch hier für die Schätzung der Raumbegrenzungsflächen verwendet.

2.4.1 Aufstellen von Ebenenhypothesen

Um die Ebenenhypothesen zu ermitteln werden zunächst nur diejenigen 3D Punkte und 3D Linien betrachtet, die auf Basis der Signaturen in der Wand-, Decke- und Bodenebene zu liegen (3D Strukturen mit den Signaturen "Wand", "Decke", "Boden", „Fenster“ oder „Tür“). Anschließend werden die drei Hauptrichtungen der 3D Linien ermittelt um die zu erwartete Ausrichtung der rekonstruierten Ebenen zu bestimmen. Hierfür werden zunächst die Richtungsvektoren aller 3D Linien berechnet und mit einem k-Means Verfahren in drei Gruppen aufgeteilt. Mithilfe eines RANSAC Verfahrens können die drei Cluster der Hauptrichtungen robust geschätzt werden. Die Mittelwerte der Richtungsvektoren der drei Cluster ergeben deren Hauptrichtungen. Abb. 5 und Abb. 6 verdeutlichen dieses Verfahren.

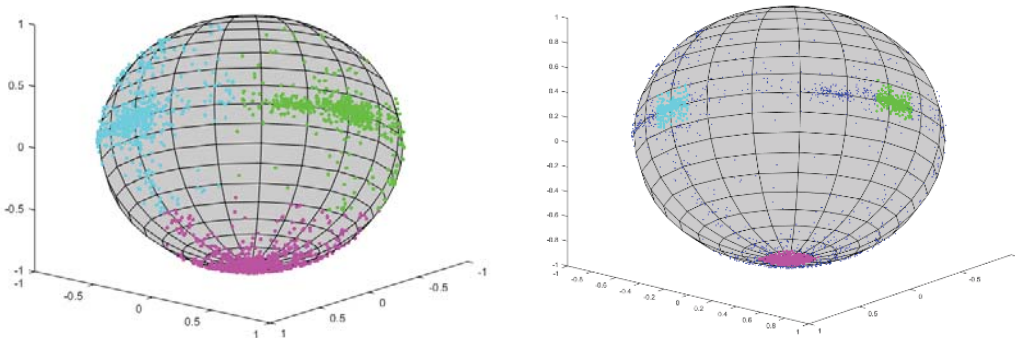


Abb. 5: Bestimmung der Hauptrichtungsvektoren (cyan, magenta, grün) aus den 3D Linien. Hauptrichtungen dargestellt auf einer Sphäre nach (links) k-Means Clustering und (Mitte) RANSAC Verfahren.

Anschließend werden die 3D Linien jeder Hauptrichtung in Ebenen gruppiert. Für diese Gruppierung werden die Mittelpunkte der 3D Linien und die Richtungsvektoren verwendet und ebenfalls ein k-Means Verfahren zur Gruppierung verwendet. Das Ergebnis der Liniengruppierung zeigt Abb. 7. Für jedes gruppierten Liniencuster werden 3D Punkte in derselben Ebene gesucht. Aus den Linien und Punkten eines Clusters werden anschließend die Parameter einer 3D Ebene geschätzt, die als Ebenenhypothese weiter bearbeitet werden.

2.4.2 Festlegung der Raumbegrenzungsflächen

In diesem Schritt wird jede Ebenenhypothese mit allen anderen Hypothesen auf Identität geprüft und bei Bedarf zusammengefügt. Die Prüfung erfolgt mit einem Schwellwertverfahren auf Basis des Winkels zwischen beiden Ebenennormalen und dem Unterschied im vierten Ebenenparameter. Für jede Hypothese wird die Anzahl der zugehörigen 3D Punkte und 3D Linien mitgespeichert, sowie auf Übereinstimmung mit den Hauptrichtungen geprüft. Aus diesen drei Größen wird pro Ebenenhypothese ein Faktor berechnet, der die Qualität dieser Hypothese bestimmt. Die Suche nach der besten Konfiguration aus allen Ebenen erfolgt iterativ. Für jede Ebenenhypothese werden weitere orthogonale und parallele Ebenen gesucht und der Qualitätsfaktor aller diesen Ebenen summiert. Als Endergebnis wird eine Konfiguration gewählt, die die den höchsten Qualitätsfaktor aufweist.

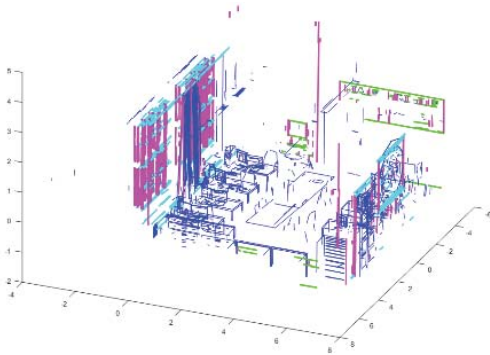


Abb. 6: 3D Linien mit Farbkodierung der drei Hauptrichtungen. In blau werden Linien dargestellt, die zu keiner der drei Richtungen gehören

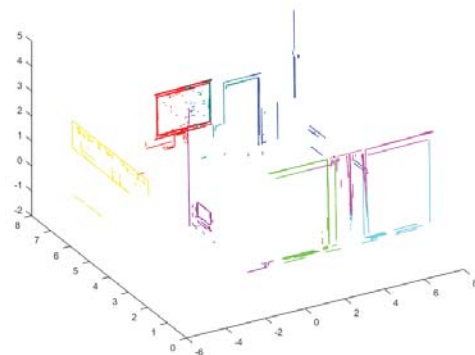


Abb. 7: Gruppierte 3D Linien für die drei Hauptrichtungen

3 Erste Ergebnisse

Um die vorgestellten Methoden zu evaluieren, wurden Bilder einer kalibrierten RGB-Kamera von einem Innenbereich eines Bürogebäudes verwendet (Abb. 8).

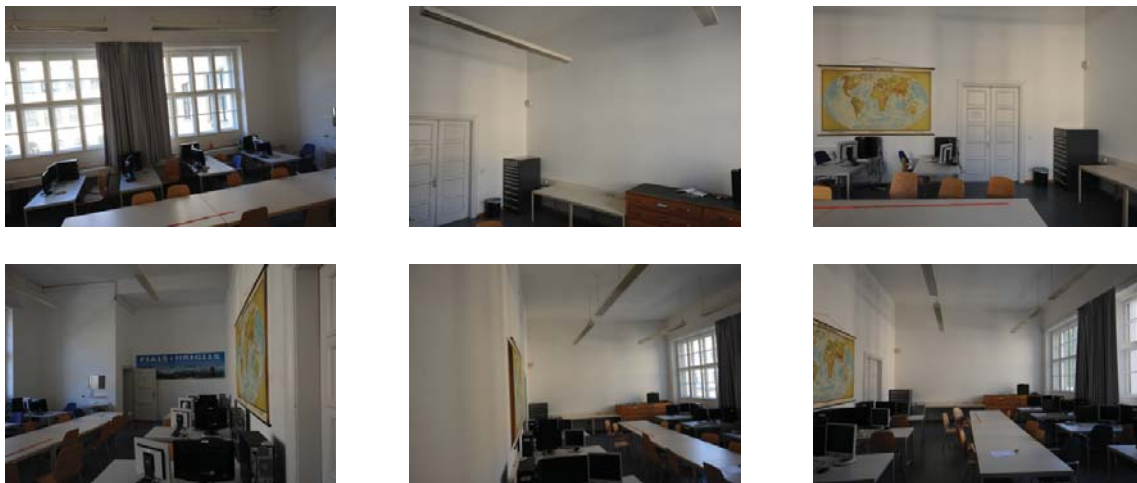


Abb. 8: Beispielbilder des Innenraumes

Aus den in Abbildung 8 präsentierten Bildern wurden 3D Punkte und 3D Linien rekonstruiert und bereits in Abbildung 3 vorgestellt. Die Klassifikation erfolgte zunächst manuell. Abbildung 9 stellt die klassifizierte 3D Punktwolke (links) und die klassifizierten 3D Linien (rechts) dar. Danach wurden die Hauptrichtungen mittels k-Means-Clustering und RANSAC Verfahren bestimmt. Die Ergebnisse dieser Suche wurden bereits in Abb. 5 und Abb. 6 gezeigt. Im Anschluss wurden alle 3D Linien, die den Hauptrichtungen entsprechen, in Cluster gruppiert und 3D Punkte gesucht, die in diesen Ebenen liegen. Der Schwellwert bei dem ein Punkt zu einer Ebene gehört wurde auf 10 cm festgelegt.

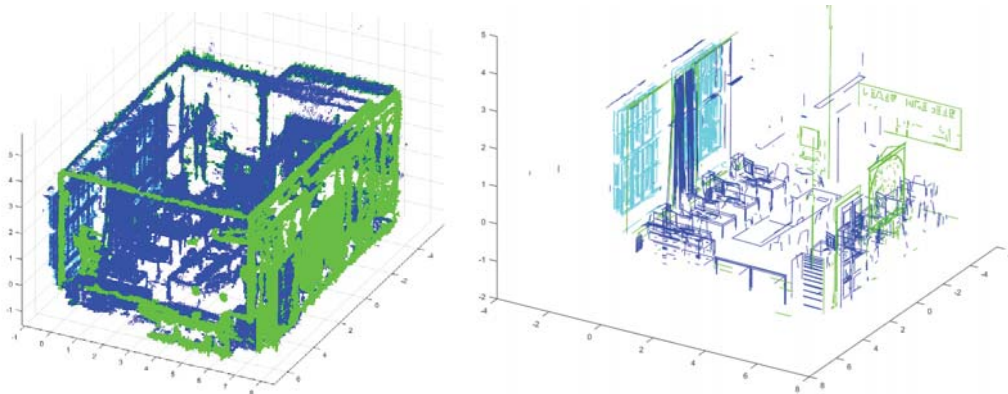


Abb. 9: Klassifizierte 3D Merkmale. Links: 3D Punkte; rechts: 3D Linien. Farbkodierung: grün – „Wand“, cyan – „Fenster“, blau – „Objekt“

Aus den Ebenenhypothesen wurden Ebenenkonfigurationen gebildet und die beste Konfiguration ausgewählt. Das Endergebnis ist in der Abbildung 10 zu sehen.

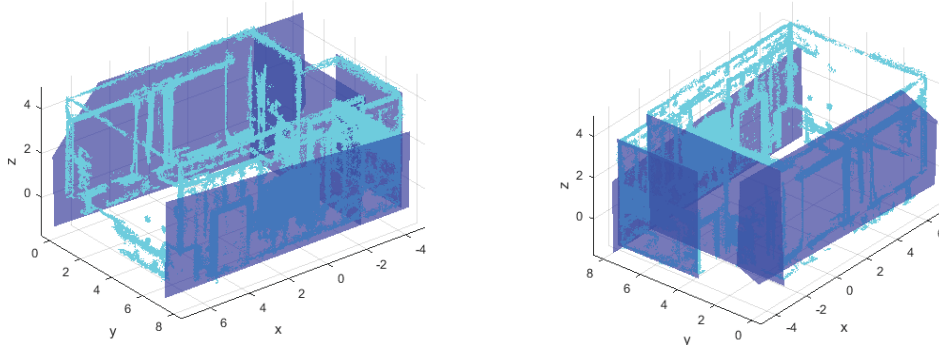


Abb. 10: Rekonstruierte Raumbegrenzungsflächen (blau) mit abgedichtete Punktwolke (cyan).

4 Fazit & Ausblick

In diesem Artikel präsentierten wir ein Konzept für eine Methode zur Innenraumrekonstruktion auf Basis von 3D Strukturen, die aus Bilderserien generiert wurden. Mithilfe einer semantischer Segmentierung der Bilder können die relevanten 3D Strukturen aus den rein geometrischen 3D Primitiven extrahiert werden und zu einer robusteren Bestimmung von Raumbegrenzungsflächen verwendet werden. In Zukunft soll neben der Realisierung der semantischen Segmentierung mittels eines konvolutionalen neuronalen Netzes die geometrischen Unsicherheiten der 3D Linien und 3D Punkte mitberücksichtigt werden, wodurch die verwendeten Schwellwerte vermieden werden können. Eine Erweiterung dieser Methode um eine nachträgliche Suche nach fehlenden Ebenen wird angestrebt und kann über eine Bedingung eines geschlossenen Raumes erreicht werden. Zusätzlich können auch Linien verwendet werden, um die Suche nach Raumecken zu unterstützen. Um das Verfahren zu evaluieren sollen in Zukunft die Eckpunkte des Raumes geodätisch aufgenommen und zur Erstellung eines Referenzmodells verwendet werden. Die Parameter der rekonstruierten Wandebenen werden dann mit den Referenzebenen verglichen und die Genauigkeit der Rekonstruktion bestimmt.

5 Literaturverzeichnis

- ARMENI, I., SENER O., ZAMIR, A.R., JIANG, H., BRILAKIS, I., FISCHER. M. & SAVARESE, S., 2016: 3D Semantic Parsing of Large-Scale Indoor Spaces. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1534-1543.
- CIREŞAN, D.C., MEIER, U., MASCI, J., GAMBARDELLA, L.M. & SCHMIDHUBER, J., 2011: Flexible, High Performance Convolutional Neural Networks for Image Classification. Proceedings of International Joint Conference on Artificial Intelligence **22**(1), 1237.
- COHEN, A., SCHÖNBERGER, J.L., SPECIALE, P., SATTTLER, T., FRAHM, J.M. & POLLEFEYS, M., 2016: Indoor-Outdoor 3D Reconstruction Alignment. In: European Conference on Computer Vision 2016, Springer International Publishing, 285-300.
- ENGEL J., SCHÖPS, T. & CREMERS D., 2014: LSD-SLAM: Large-Scale Direct Monocular SLAM. European Conference on Computer Vision, Springer International Publishing, 834-849.
- ENGEL, J., KOLTUN, V. & CREMERS, D., 2016: Direct Sparse Odometry. arXiv:1607.02565
- GERKE, M. & XIAO, J., 2013: Supervised and unsupervised MRF based 3d scene classification in multiple view airborne oblique images. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **II-3/W3**, 25-30.
- HOFER, M., MAURER, M. & BISCHOF, H., 2016: Efficient 3D Scene Abstraction Using Line Segments. Computer Vision and Image Understanding (CVIU), Elsevier B.V., <http://dx.doi.org/10.1016/j.cviu.2016.03.017>
- SNAVELY, N., SEITZ, S.M. & SZELISKI, R., 2007: Modeling the World from Internet Photo Collections. International Journal of Computer Vision **80**(2), 189-210.
- KENDALL, A., BADRINARAYANAN, V. & CIPOLLA, R., 2015: Bayesian SegNet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding. arXiv preprint arXiv:1511.00561
- LIU, C., SCHWING, A.G., KUNDU, K., URTASUN, R. & FIDLER, S., 2015: Rent3D: Floor-Plan Priors for Monocular Layout Estimation. Conference on Computer Vision and Pattern Recognition, 3413-3421.
- ROTHERMEL, M., WENZEL, K., FRITSCH, D. & HAALA, N., 2012: Sure: Photogrammetric surface reconstruction from imagery. Proceedings LC3D Workshop, Berlin, http://www.ifp.uni-stuttgart.de/publications/2012/Rothermel_etal_lc3d.pdf.
- WU, C. 2011: VisualSFM: A visual structure from motion system. <http://www.cs.washington.edu/homes/ccwu/vsfm>.
- WANG, H., WANG, J. & LIANG, W., 2016: Online reconstruction of indoor scenes from rgb-d streams. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3271-3279.
- NEVEROVA, N., MUSELET, D. & TRÉMEAU, A., 2013: 21/2 d scene reconstruction of indoor scenes from single rgb-d images. Computational Color Imaging, Springer Berlin Heidelberg, 281-295.
- CHOI, S., ZHOU, Q. Y. & KOLTUN, V., 2015: Robust reconstruction of indoor scenes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 5556-5565.

Gleichwertige Ziele in dynamischen Navigationskarten

JOHANNES OEHRLEIN¹, THOMAS C. VAN DIJK² & JAN-HENRIK HAUNERT¹

Zusammenfassung: Die Generierung übersichtlicher Karten erfordert Verfahren der automatischen Generalisierung. Zur Darstellung auf Navigationskarten werden beispielsweise Objekte eines Verkehrsnetzes anhand ihrer Bedeutung für das Netz ausgewählt. Beschränkt man sich auf die Navigation von einem festen Startpunkt aus, verlieren viele Objekte für die Karte an Bedeutung. Durch Verzicht auf deren Darstellung wird die Karte übersichtlicher. Diesen Umstand nutzen van Dijk et al. (2016) für einen Algorithmus zur standortbasierten Generalisierung von Straßennetzen. Dieser trifft – abhängig von einem fest gewählten Standort – durch die Zusammenfassung von als gleichwertig erkannten Zielen eine Auswahl. Hebt man die Fixierung des Standorts auf, ergeben sich neue Möglichkeiten für mobile Geräte. Dieser Beitrag beschäftigt sich mit den Problemen, die mit der Dynamik Einzug in diesen Algorithmus halten, und bietet erste Lösungsansätze.

1 Einleitung

Bei der Erzeugung von Karten bedarf es einer Abwägung zwischen Übersichtlichkeit und Informationsfülle. Dabei müssen Objekte ausgewählt und vereinfacht werden, um die Lesbarkeit einer Karte zu gewährleisten. Soll dieser Prozess automatisch vonstattengehen, benötigt man Algorithmen, die die einzelnen Schritte der kartographischen Generalisierung beherrschen (HAKE et al. 2002)

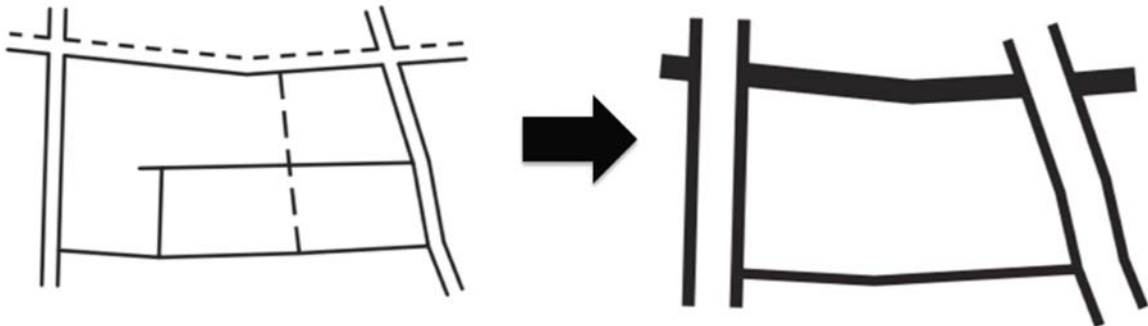


Abb. 1: Generalisierung eines Straßennetzes (nach HAKE et al. 2002)

Einer dieser Schritte betrifft die Generalisierung von Straßennetzen (Abb. 1). Hierbei müssen zum einen Straßenzüge vereinfacht werden. Zum anderen müssen Straßen anhand ihrer Bedeutung für den auf der Karte dargestellten Inhalt ausgewählt werden. Mit diesem Problem beschäftigen sich VAN DIJK et al. (2016). Dabei konzentrieren sie sich auf die Generalisierung von Straßennetzen,

¹ Universität Bonn, Institut für Geodäsie und Geoinformation, Lehrstuhl Geoinformation,
E-Mail: [haunert, oehrlein]@igg.uni-bonn.de

² Universität Würzburg, Institut für Informatik, Lehrstuhl für Informatik I,
E-Mail: thomas.van.dijk@uni-wuerzburg.de

bei der die Bedeutung einzelner Straßen hinsichtlich eines festen Punkts berücksichtigt wird. Dieser Punkt kann beispielsweise als aktueller Standort oder auch als Ziel einer Navigation aufgefasst werden.

Bei der standortbasierten Generalisierung von Straßennetzen wird die Auswahl einer Straße anhand ihrer Bedeutung für den gegebenen Standort getroffen. Dabei kommt einer Straße im Allgemeinen eine größere Bedeutung zu, je näher sie sich zum Standort befindet (HAMPE et al. 2004; KOPF et al. 2010).

Unser Beitrag ist nun die Untersuchung der Eignung eines so generalisierten Straßennetzes für die Erstellung von dynamischen Karten, beispielsweise zur Navigation. Im Rahmen der Zielführung ist dabei insbesondere von Interesse, wie sich eine solche Karte verhält, wenn der Standort, auf den sich die Generalisierung bezieht, verändert. Insbesondere für mobile Geräte bietet es sich an, hier den aktuellen Standort des Nutzers der Navigation zu verwenden.

In Abschnitt 2 wird zunächst die Arbeit von van Dijk et al. vorgestellt und ihre Eignung zur Zielführung diskutiert. Anschließend werden in Abschnitt 3 die Probleme analysiert, die durch Veränderung des Standortes auftreten. Außerdem wird erörtert, wie sich diese Probleme sinnvoll beheben lassen. Schließlich bietet Abschnitt 4 Raum für ein Fazit.

2 Standortbasierte Generalisierung von Straßennetzen

Bei der Reise von A nach B mithilfe eines Navigationsgeräts ist die genaue Adresse des Ziels für den Reisenden erst einmal uninteressant, sobald sie in das Gerät eingegeben wurde. Van Dijk et al. (2016) ziehen als Beispiel eine Reise von Berlin nach Groningen heran: Zu Beginn der Reise interessiert es den Nutzer nicht, wo sich die Zieladresse genau in Groningen befindet. Für den ersten Abschnitt seiner Reise ist dies unerheblich. Wie man Berlin in Richtung der Niederlande verlässt, hängt nicht von der Straße oder des Stadtviertels des Zielorts in Groningen ab. Auf dieser Idee basiert der Algorithmus.

2.1 Ausgangssituation

Im zu generalisierenden Straßennetz betrachtet man mögliche Zielorte von einem Standort aus. Zielorte, zu denen die Routen vom gegenwärtigen Standort aus zu großen Teilen übereinstimmen, sollen zu abstrakten Zielen zusammengefasst werden. Hierzu wird das Straßennetz als **Graph** $G = (V, E)$ aufgefasst: Die Kanten stellen Straßenabschnitte dar, die sich an den Knoten treffen. G enthält dabei den Standort $s \in V$. Die Kanten von G sind gewichtet. Diese **Gewichtung** $w: E \rightarrow \mathbb{R}^{\geq 0}$ bestimmt das Kriterium, nach dem die Route geplant werden soll; für die kürzeste Wegstrecke verwendet man beispielsweise als Gewicht einer Kante ihre euklidische Länge. Betrachtet man einen Weg, so lässt sich dieser als eine endliche Folge von Kanten formalisieren. Dabei entspricht das Gewicht des Weges der Summe der Gewichte der einzelnen in ihm enthaltenen Kanten. Unabhängig vom gewählten Kriterium ist man bei der Navigation an einem optimalen, also gewichtsminimalen Weg interessiert. Ermittelt man nun von s den Weg minimalen Gewichts (bezüglich w) zu jedem möglichen Ziel, so ergibt eine Baumstruktur. Im Folgenden beziehen wir uns auf den Fall der euklidischen Abstände. Dann sind diese gewichtsminimalen Wege kürzeste Wege und wir

sprechen vom **Kürzeste-Wege-Baum T** . Auf diesem operiert der Algorithmus. Der Weg von einem beliebigen Knoten $u \in V$ zu einem Knoten $v \in V$ in diesem Baum ist eindeutig und wird im Folgenden mit P_{uv} bezeichnet.

2.2 Gleichwertige Zielorte

Ziel ist es, mögliche Bestimmungsorte zusammenzufassen, die über größtenteils übereinstimmende Wege erreicht werden können. Hierzu wird ein Kriterium eingeführt, mit dem entschieden werden kann, ob zwei Knoten $u, v \in V$ **kompatibel**, also zusammenfassbar, sind. Hierzu wird jeweils der kürzeste Weg P_{su} bzw. P_{sv} vom Standort aus betrachtet. Diese beiden Wege stimmen von s aus bis zu einem gewissen Knoten $x \in V$ überein. u und v werden als kompatibel bezeichnet, wenn für einen bestimmten Schwellenwert $\alpha \in [0,1]$ folgende Ungleichungen erfüllt sind:

$$\frac{w(P_{sx})}{w(P_{su})} \geq \alpha \quad \text{und} \quad \frac{w(P_{sx})}{w(P_{sv})} \geq \alpha.$$

Das heißt: u und v sind kompatibel, wenn der gemeinsame Weg mindestens den Anteil α am Gesamtweg sowohl nach u als auch nach v einnimmt (siehe Abb. 2).

Der Parameter α ist hierbei die einzige nicht-geometrische Information, die dem Algorithmus übergeben werden muss. Dieser Wert entscheidet über die Dichte des dargestellten Straßennetzes. Kleine Werte (0,5 und kleiner) sorgen für eine starke Ausdünnung des kürzeste-Wege-Baums.

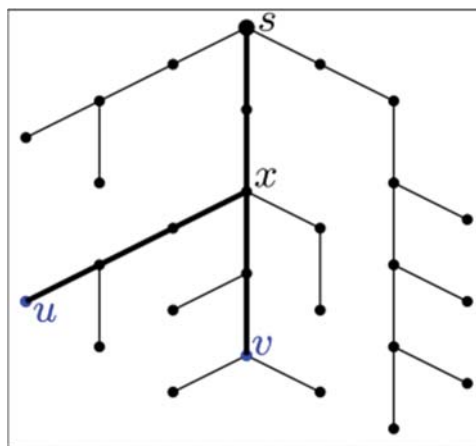


Abb. 2: u und v mit gemeinsamen Vorgänger x . Nimmt man gleiche Kantengewichte an, so beträgt der gemeinsame Weg $2/5$ des Weges nach u und $1/2$ des Weges nach v . (Abb.: VAN DIJK et al., 2016)

Große Werte (0,9 und größer) erhalten viele Details und geben auch in größerer Entfernung zum gewählten Standort noch viele Informationen über den Eingabegraphen wieder. Das Ziel, dass das Gebiet nahe dem Standort detaillierter dargestellt wird als die weitere Umgebung, wird allerdings unabhängig von der Wahl von α erreicht.

2.3 Algorithmus

Der Algorithmus fasst nun ausgehend von den Blättern des Baumes möglichst viele kompatible Knoten zusammen. Dabei wird jeder Knoten genau einer **Zelle** zugewiesen, die nur Knoten enthält, die paarweise kompatibel sind (siehe Abb. 3, links). Das Resultat ist ein von T induzierter Baum, der für jede Zelle einen Repräsentanten enthält und diese untereinander gemäß den in T enthaltenen Kanten verknüpft. Ein sinnvoller Repräsentant ist der (eindeutige) Knoten, der dem Standort s am nächsten ist (siehe Abb. 3, rechts).

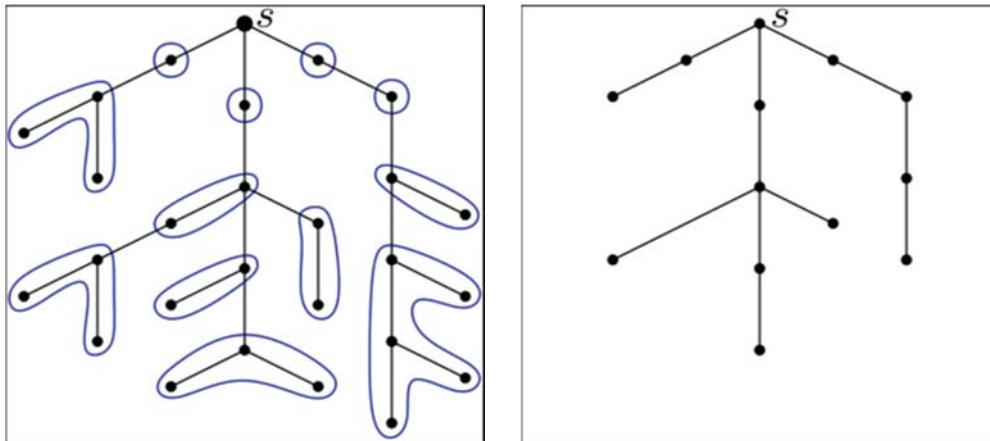


Abb. 3: Ergebnis des Algorithmus für $\alpha = 2/3$. Links: vom Algorithmus ermittelte Zellen. Rechts: Ausgabe nach Zusammenfassung. (Abb. nach VAN DIJK et al., 2016)

Van Dijk et al. (2016) zeigen, dass ihr **Greedy-Algorithmus** ein optimales Ergebnis zurückliefert in dem Sinne, dass es keine Möglichkeit gibt, die kompatible Knoten zu einer geringeren Anzahl an Repräsentanten zusammenzufassen. Dazu benötigt der Algorithmus asymptotisch Linearzeit (bezüglich der Anzahl $|V|$ an Knoten in T).

2.4 Querverbindungen

Bei der Reduktion eines Graphen (Eingabe-Sträßennetz) auf einen Baum (Kürzeste-Wege-Baum) gehen viele Informationen über den Zusammenhang im Graphen verloren. Insbesondere gibt es Orte im Sträßennetz, die über zwei völlig verschiedene, gleichlange Wege vom gegenwärtigen Standort aus erreichbar sind. In einem gewöhnlichen kürzeste-Wege-Baum ist dieser Umstand nicht mehr erkennbar. Daher wird die folgende Modifikation am Kürzeste-Wege-Baum vorgenommen.

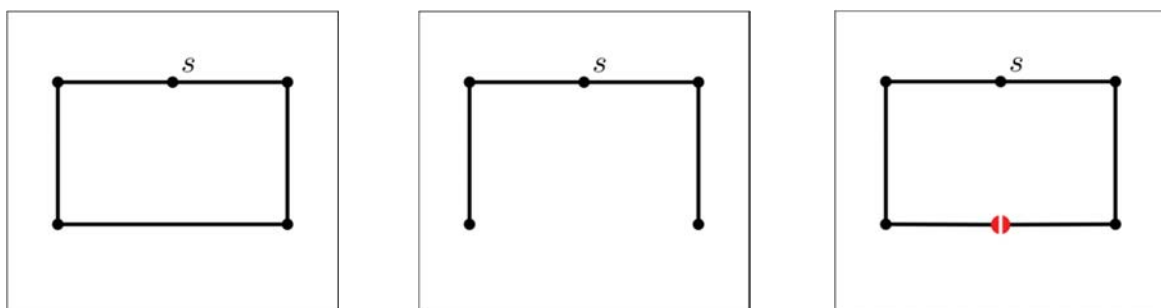


Abb. 4: Vom Graph (links) über den kürzeste-Wege-Baum (Mitte) zum modifizierten kürzeste-Wege-Baum (rechts) mit virtuellen Knoten (rot)

Es gibt Kanten $e = (u, w)$ in G , die nicht in T enthalten sind. Die Ursache hierfür ist, dass der kürzeste Weg von s nach w nicht über u führt und umgekehrt. Unter der Annahme, dass das ge-

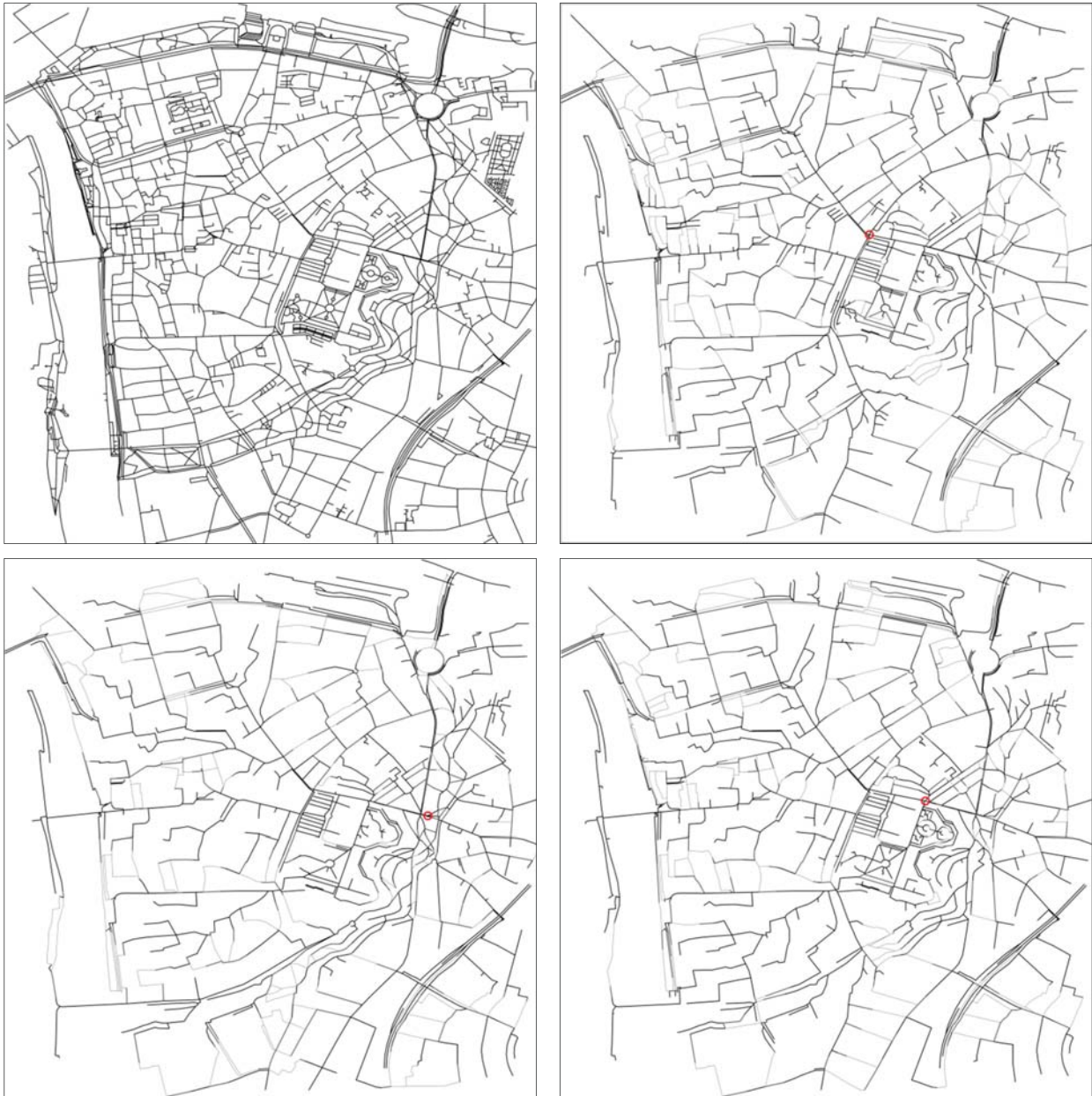


Abb. 5: Beispiel an einem Ausschnitt des Würzburger Straßennetzes: Die Eingabedaten werden oben links dargestellt. Alle weiteren Abbildungen zeigen das Ergebnis des Algorithmus für verschiedene Standorte (●) mit demselben Wert $\alpha = 0,9$. Die schwarzen Kanten bilden den erzeugten Baum, die grauen Kanten die Querverbindungen (siehe Abschnitt 2.4). Verschiedene Standorte erzeugen verschiedene Ausgaben. Betrachtet man die Bilder von oben rechts nach unten links im Uhrzeigersinn (oder umgekehrt), so kann man die Auswirkungen einer Standortveränderung entlang eines Straßenzugs nachvollziehen.

wählte Kantengewicht sich **stetig entlang der Kante verteilt**, findet sich auf diesen Kanten jeweils ein Punkt (kein Knoten!), der auf zwei verschiedenen Wegen gleichen Gewichts erreicht werden kann. Im Fall euklidischer Abstände ist diese Annahme erfüllt, das Gewicht einer Kante ist sogar gleichmäßig verteilt. Für solche Punkte werden jeweils zwei **virtuelle Knoten** eingefügt, die entsprechend der ursprünglichen, in T fehlenden Kante über zwei eingefügte Kanten mit T verbunden werden. So entsteht der **modifizierte kürzeste-Wege-Baum T'** (Abb. 4).

Um derlei Information im Ergebnis wiederzugeben, wird für jedes Paar Repräsentanten, das von mindestens einem Paar virtueller Knoten jeweils einen Knoten repräsentiert, eine **Querverbindung** eingeführt. Diese verbindet die Repräsentanten über das Paar virtueller Knoten, welches dem aktuellen Standort s am nächsten ist.

2.5 Diskussion

Der vorgestellte Algorithmus liefert generalisierte Informationen über ein Straßennetz, die in der Nähe des Standorts detailreicher sind als in der weitläufigen Umgebung. Dieses Problem wird nach einem exakt definierten Kriterium optimal und effizient gelöst. Damit bildet der Algorithmus eine gute Grundlage für standortbasierte Karten (Hampe et al. 2004). Da für die Generalisierung von Straßennetzen allerdings weitere Kriterien berücksichtigt werden müssen, bedarf der Algorithmus in jedem Fall einer Ausarbeitung für die Anwendung. Beispiel hierfür ist die Erhaltung von Straßenzügen, die der Algorithmus in der vorgestellten Form nicht gewährleistet (THOMSON & RICHARDSON 1999).

3 Auswirkungen eines beweglichen Standorts

Neben statischen Anwendungen wie beispielsweise Anfahrtsbeschreibungen bieten sich standortbasierte Karten insbesondere für **dynamische Karten** an. Hier wird die Karte regelmäßig an einen aktualisierten Standort angepasst. Dieser ergibt sich durch Simulation oder durch die Verwendung eines mobilen Geräts. Dabei gilt, dass verschiedene Standorte im Allgemeinen verschiedene Ausgaben erzeugen. Eine Möglichkeit, daraus eine dynamische Karte zu erzeugen, besteht darin, bei einer Standortaktualisierung eine neue Ausgabe zu erzeugen und anzuzeigen. In Abb. 5 wird eine Folge solcher Darstellungen präsentiert. Um die Ansicht für einen Standort in die Ansicht für den nächsten Standort sinnvoll zu überführen, müssen die Auswirkungen einer Standortveränderung auf das generalisierte Straßennetz untersucht werden. Diese Auswirkungen werden im Folgenden erläutert.

3.1 Veränderungen im modifizierten kürzeste-Wege-Baum

Durch die Variation des Standorts muss der Kürzeste-Wege-Baum aktualisiert werden. Dabei ist es nicht ausreichend, die Abstände der Knoten zum (veränderten) Standort zu aktualisieren. Auch die Struktur des Baums kann sich dadurch verändern.

Untersucht man die Auswirkungen einer Standortveränderung auf den kürzeste-Wege-Baum, so stellt man fest, dass es vom Standort s abhängt, welche Kanten des Graphen G in T nicht enthalten sind. Diese Veränderung lässt sich am modifizierten Kürzeste-Wege-Baum T' nachvollziehen. Hierzu betrachten wir ein Paar virtueller Knoten (v_1, v_2) auf einer Kante (u, w) . Der Weg P_{sv_1} erreicht v_1 über den Knoten u , der Weg P_{sv_2} führt von s nach v_2 über w . Die Wege P_{sv_1} und P_{sv_2}

haben hierbei dieselbe Länge. Nun betrachtet man die Veränderung des Standorts als stetigen Prozess. Verschiebt man s , so müssen auch die virtuellen Punkte v_1 und v_2 hinsichtlich Lage und Abstand zu s aktualisiert werden.

Da T' ein Baum ist, existiert ein eindeutiger Weg von v_1 nach v_2 . In G beschreibt dieser Weg einen Kreis K , da das „virtuelle Hindernis“ (v_1, v_2) in G nicht existiert. Unterteilt man die Verschiebung von s in infinitesimal kleine Etappen, so muss für diese zwischen folgenden Fällen unterschieden werden (siehe Abb. 6):

- a) $s \in K$. Liegt der Standort auf diesem Kreis, so bewegen sich s und das Punktepaar (v_1, v_2) gleichförmig auf diesem Kreis. Der Abstand von s bleibt unverändert.
- b) $s \notin K$. Liegt der Standort nicht auf diesem Kreis, so hat eine Variation von s keinen Einfluss auf die Position von (v_1, v_2) . Der Abstand von s muss aktualisiert werden.

Somit ist die Position der virtuellen Punkte stetig abhängig vom Standort (Abb. 6).

Alle übrigen (also nicht-virtuellen) Knoten in T' besitzen eine feste Position. Bei einer Variation des Standorts, verändert sich bei diesen Knoten lediglich der Abstand zum Standort.

3.2 Auswirkungen auf den Baum der Ausgabe und Stabilisierungsversuche

Durch Veränderung des Eingabe-Baums liefert der Algorithmus mit großer Wahrscheinlichkeit ein anderes Ergebnis. Hierzu betrachten wir eine Zelle im Ausgabe-Baum. Insbesondere gilt unser Interesse dem Knoten x dieser Zelle, der dem Standort s am nächsten ist.

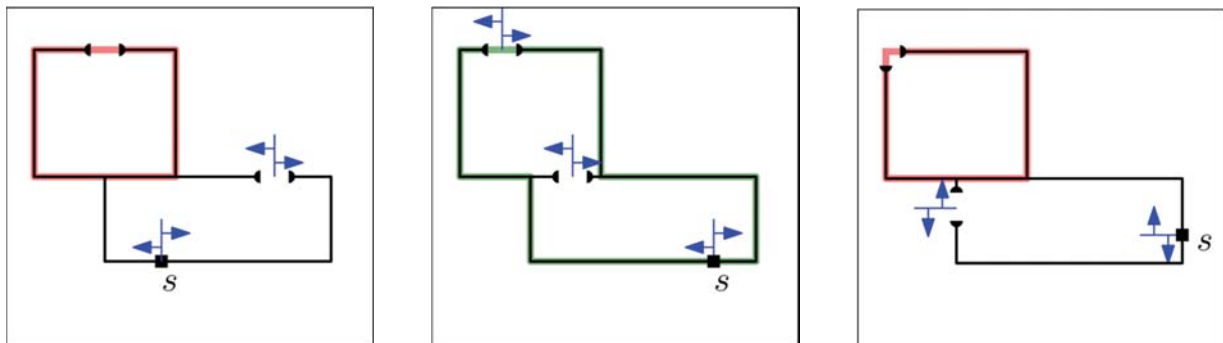


Abb. 6: Für das obere virtuelle Punktepaare ist K farbig markiert, Lageveränderungen sind in blau dargestellt. Eine Variation des Standorts s sorgt nur dann für eine Veränderung der Lage eines virtuellen Punktepaars in T' , falls s sich auf dem oben beschriebenen Kreis befindet.

Für die Bildung einer solchen Zelle ist der Parameter α von großer Bedeutung. Nur die Nachfahren v eines Knotens x werden der Zelle zugewiesen, für die

$$\frac{w(P_{sx})}{w(P_{sv})} \geq \alpha$$

gilt. Im allgemeinen Fall besitzt x Nachfahren in T' und ist somit kein virtueller Knoten. Somit ist x statisch und $w(P_{sx})$ verändert sich um den Betrag der Variation von s . Bei v kann es sich um einen virtuellen Knoten handeln. Somit besteht nach Fallunterscheidung aus Abschnitt 3.1 die Möglichkeit, dass sich die Länge des Weges P_{sv} nicht ändert. Das heißt, dass für eine Standortänderung um einen kleinen Wert $\varepsilon > 0$ folgende Szenarien möglich sind:

$$\frac{w(P_{sx}) \pm \varepsilon}{w(P_{sv}) \pm \varepsilon} \geq \alpha \quad \text{oder} \quad \frac{w(P_{sx}) \pm \varepsilon}{w(P_{sv})} \geq \alpha.$$

Unabhängig vom Nenner der linken Seite wird der Wert des Bruchs also größer, wenn sich der Standort von der untersuchten Zelle entfernt, und kleiner, wenn er sich auf sie zu bewegt. Insbesondere bleibt die Zuweisung eines Knotens v zu einer Zelle mit Repräsentant x legitim, solange sich s von der Zelle wegbewegt. Diese Beobachtung ist im Einklang mit der Grundidee der standortbasierten Generalisierung, dass ein Gebiet detailreicher dargestellt wird, je näher es sich zum Standort befindet.

Durch die Auffassung als Graph wird das Straßennetz diskretisiert. Diese Diskretisierung ist im Resultat verstärkt sichtbar. Da lediglich Knoten im modifizierten Kürzeste-Wege-Baum auf Kompatibilität überprüft werden, wird die mögliche Größe einer gebildeten Zelle im Allgemeinen nicht voll ausgenutzt. Die folgende Rechnung verdeutlicht diesen Umstand, wobei ausgenutzt wird, dass man ausgehend vom Standort s einen Knoten v in der Zelle von x in T' nur über x erreichen kann.

$$\frac{w(P_{sx})}{w(P_{sv})} = \frac{w(P_{sx})}{w(P_{sx}) + w(P_{xv})} \geq \alpha \quad \Leftrightarrow \quad \frac{1 - \alpha}{\alpha} w(P_{sx}) \geq w(P_{xv})$$

Da der Algorithmus jeweils diskrete Punktmengen zu solchen Zellen zusammenfasst, herrscht hier im Allgemeinen keine Gleichheit. Somit besteht für die einzelnen Zellen ein gewisser Rahmen, in dem sich s bewegen darf, ohne dass dies Konsequenzen für die Stabilität der einzelnen Zellen hat. Eine Möglichkeit, die Stabilität der Ausgabe zu gewährleisten, besteht darin, die Toleranz für die Kompatibilität der Knoten zu erhöhen. Das heißt, die Zellen werden grundsätzlich für einen gegebenen Wert α gebildet. Verändert sich der Standort, so wird die Kompatibilität der Zellen jedoch für einen Wert $\alpha^* < \alpha$ überprüft. Erst wenn Knoten einer Zelle bezüglich α^* inkompatibel sind, wird die Zelle neu berechnet.

Eine andere Möglichkeit, die Ausgabe zu stabilisieren, besteht darin, die dynamische Karte möglichst in Echtzeit zu aktualisieren. Die stetige Abhängigkeit der virtuellen Punkte des modifizierten Kürzeste-Wege-Baums verleitet zu diesem Ansatz. Eine Echtzeit-Realisierung ist nur möglich, wenn die Prozessierung nicht zu aufwändig wird. Diesbezüglich ist ein Blick auf die mögliche Anzahl der virtuellen Knoten interessant. Da je Kante maximal zwei virtuelle Knoten hinzukommen, liegt deren Anzahl in $\mathcal{O}(|E|) \subseteq \mathcal{O}(|V|^2)$. Nehmen wir an, dass an einer Kreuzung stets höchstens $k \in \mathbb{N}$ Straßen beteiligt sind. Dann ergibt sich die Abschätzung $|E| \leq k \cdot |V|$. Die Annahme ist sinnvoll, da in einem realen Straßennetz im Allgemeinen höchstens eine geringe zweistellige Anzahl an Straßen eine Kreuzung bildet. Für planare Eingabegraphen lässt sich diese Abschätzung zu $|E| < 3 \cdot |V|$ verschärfen. Allerdings ist die Planarität eines Straßennetzes durch Über- bzw. Unterführungen verletzt. Gemessen an der Anzahl Über- bzw. Unterführungen ist diese Verletzung jedoch geringfügig.

Für eine kontinuierliche Dynamik in der zu erzeugenden Karte müssen die Auswirkungen einer Standortveränderung auf die gebildeten Zellen näher untersucht werden. Der stetige Zusammenhang zwischen der Lage des Standort und der Lage der virtuellen Knoten des modifizierten Kürzeste-Wege-Baums ist nicht hinreichend.

4 Zusammenfassung und Ausblick

Der diskrete Ansatz klingt auf den ersten Blick vielversprechend. Durch eine gewisse Fehlertoleranz bezüglich der Kompatibilität soll die Stabilität der Ausgabe erhöht werden. Allerdings führt

eine stete Bewegung des Standorts in eine Richtung zum Aus- und Überreizen dieser Toleranz. Zu einem bestimmten Zeitpunkt muss die Ausgabe aktualisiert werden. Durch die erhöhte Fehlertoleranz ist dann der Unterschied zur korrekten Lösung allerdings vermutlich auffälliger. Es muss überprüft werden, ob diese selteneren, womöglich heftigeren Veränderungen der Ausgabe die Informationsvermittlung in stärkerem Maße beeinflussen, als dies eine sofortige Behebung auftretender Kompatibilitätsprobleme nach sich zöge.

Die Echtzeit-Lösung stimmt hier zuversichtlicher, insbesondere da der effiziente Algorithmus und eine Topologie nahe der Planarität diesen Ansatz begünstigen. In Abschnitt 3.2 wurde bereits festgestellt, dass die Auswirkung einer Standortveränderung noch stärker untersucht werden muss. Davon abgesehen ist das Verhalten der Querverbindungen bei Veränderung des Standorts von großem Interesse. Diese sind unverzichtbar, wenn dem Leser der Karte Informationen zum Zusammenhang des Graphen nicht gänzlich verwehrt werden sollen. Da die einzelnen Querverbindungen entweder dargestellt oder nicht dargestellt werden, liegt kein stetiger Zusammenhang mit der Lage des Standorts vor. Daher muss untersucht werden, wie sprunghaft diese Auswahl ist und wie diese möglichst stabilisiert werden kann.

5 Literaturverzeichnis

- VAN DIJK, T. C., HAUNERT, J.-H. & OEHRLEIN, J., 2016: Location-dependent generalization of road networks based on equivalent destinations. *Computer Graphics Forum* **35**(3), 451-460.
- HAKÉ, G., GRÜNREICH, D. & MENG, L., 2002: *Kartographie: Visualisierung raum-zeitlicher Informationen*. 8th ed., Berlin: De Gruyter Lehrbuch.
- HAMPE, M., SESTER, M. & HARRIE, L., 2004: Multiple Representation Databases to Support Visualization on Mobile Devices. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **35**(B4), 135-140.
- KOPF, J., AGRAWALA, M., BARGERON, D., SALESIN, D. & COHEN, M., 2010: Automatic generation of destination maps. *ACM Transactions on Graphics* **29**(6), 158.
- THOMSON, R. C. & RICHARDSON, D. E., 1999: The “Good Continuation” Principle of Perceptual Organization applied to the Generalization of Road Networks. *Proceedings of the ICA 19th International Cartographic Conference*, 14-21.

Indoor Point Cloud Segmentation for Automatic Object Interpretation

LAVINIA S. RUNCEANU¹, SUSANNE BECKER¹, NORBERT HAALA¹ & DIETER FRITSCH¹

Summary: The paper presents an algorithm for the automatic segmentation of point clouds from low cost sensors for object interpretation in indoor environments. This algorithm is considering the possible noisy character of the 3D point clouds and is using an iterative RANSAC approach for the segmentation task. For evaluating the robustness, it is applied on two indoor datasets, acquired with the Google Tango tablet and with the NavVis M3 trolley. The realized evaluation reveals the potential of the two systems for delivering data suitable for automatically interpreting indoor structures.

1 Introduction

A large variety of systems and applications are offering mapping, localization and navigation services for outdoor environment. However, people spend most of their time indoor, where there is a lack in digital maps and where conventional GPS services do not work. Despite this need in indoor navigation applications, the developments in the field of augmented and virtual reality, including also game industry, and Building Information Modeling (BIM), are also requiring intelligent indoor models. In order to obtain the needed models, innovative equipment is required to replace the traditional systems, which are mostly expensive and sometimes inconvenient to use. Recently, a variety of systems, designed for this purpose, were made available on the market. In order to increase the mapping efficiency and in the same time to reduce the mapping costs, these systems integrate laser scanners, cameras and sometimes inertial measurement units. Also, they adapt their design to the indoor space, being built as a trolley or as a backpack (e.g. NavVis M3 Trolley, Leica Pegasus). Another step further is made by the availability of devices integrating depth cameras, at a consumer-level. These platforms are using low cost sensors which made them affordable to the general public (e.g. Phab 2 Pro Phone, Google Tango Tablet, Microsoft Kinect). All this progress enables unexperienced users to contribute to the indoor mapping request, but in the same time this rises new challenges which need to be overcome. The acquired 3D data needs to be automatically interpreted in order to obtain the models for the aforementioned applications. For the automatic 3D data interpretation, segmentation is needed. Point cloud segmentation is a subject of research for many years, however point clouds coming from low cost sensors, rise new challenges for the segmentation and interpretation process, which are addressed by this work.

The contributions of the paper are:

- a robust algorithm for automatic extraction of wall structures in indoor scenarios;
- analysis of the potential of different sensors to provide interpreted indoor structures with a data-based algorithm.

¹ University of Stuttgart, Institute for Photogrammetry (ifp), Geschwister-Scholl-Straße 24D, D-70174 Stuttgart, E-Mail: [Forename.Lastname]@ifp.uni-stuttgart.de

The paper is structured as follows: Section 2 provides an overview of the related work focused on the available indoor mapping systems and on the segmentation and reconstruction methods using point clouds. The used systems and the obtained indoor data are presented in section 3. The definition of the proposed interpretation algorithm is given in section 4. Section 5 gives an accuracy analysis of the input data, as well as of the automatically interpreted structures. Conclusions are given in section 6.

2 Related Work

The related work is presented separately considering the available indoor mapping systems and the point clouds reconstruction methods.

2.1 Indoor mapping systems

Being well suited for mapping textureless surfaces, often present in indoor environments, active systems are the most used for satisfying the indoor mapping request. Improvements in sensor design and technology as well as in the used algorithms have contributed to the large variety of systems used for 3D data collection today. A lot of research work is addressing the available data acquisition solutions for indoor modeling. However, in this paper, the focus is on the recent ones using Indoor Mobile Mapping Systems (IMMS) commercially available and low cost consumer-grade range cameras will be mentioned here.

Initially introduced as *Zebedee* by BOSSE et al. (2012), the ZEB1 from GeoSLAM (GEOSLAM 2017) is composed of a 2D laser scanner, a low-cost IMU and two robotic systems springs. THOMSON et al. (2013) are comparing it and the iMS 3D from Viامتريس (VIAMETRIS IMS 3D 2017) with Terrestrial laser scanner (TLS) in terms of accuracy. Both IMMS systems delivered centimeter-accuracy, but the iMS 3D, composed of three 2D laser scanners and a Point Grey Ladybug spherical camera, proved to generate higher quality point clouds. BASSIER et al. (2015) investigated data acquisition techniques and workflows considering the transitory tendency from TLS to IMMS. Despite the previous mentioned systems, they considered also the M3 trolley from NavVis in their comparison. Another solution offered by Leica Geosystems is the integration of the needed mapping sensors in the Pegasus Backpack. KURIAN & MORIN (2016) used the 3D point cloud generated by it for developing a method of minimizing the computation cost and data storage for real-time mapping applications. Although range cameras were available for several years in the game industry, the research interest grew with the release of Microsoft Kinect (MSDN KINECT 2017) in 2010. Being firstly based on structured light and later on time-of-flight (ToF) principle, the Microsoft Kinect became one of the most affordable device for 3D mapping. BÖHM (2014) includes it in his investigations, where he checked the performance of some structured light sensors in regard to accuracy and repeatability. Continuously improvement in the hardware and software made available low-cost range camera devices, aiming among others at indoor data acquisition, like it is the case of DPI-8 (DOTPRODUCT DPI-8 2017) and Google Tango tablet (WIKIPEDIA TANGO 2017). The capabilities of the last mentioned one, Tango tablet, have been subject of research for indoor scanning by DIAKITÉ & ZLATANOVA (2016).

Due to the indoor scene complexity and to sensors limitations (like battery life), normally the acquired data, is composed of multiple sessions (sometimes from different viewpoints), which

have to be aligned and geo-referenced, before other processing is performed. While there is the option to manually define the point correspondences and to apply an algorithm like Iterative Closest Point (ICP) (BESL & MCKAY 1992), this problem is mostly solved automatically with Simultaneous Localization and Mapping (SLAM) (LEONARD & DURRANT-WHYTE 1991) approaches. Each of the before mentioned systems are using a type of SLAM implementation for partially or totally solving the registration problem. Being commercial products, the used algorithm is not always good documented. However, some research works are addressing the SLAM problem specifically to the device, like: BOSSE et al. (2012) for *Zebedee*, NEWCOMBE et al. (2011) for Kinect, LASKAR et al. (2016) for Tango tablet.

2.2 Segmentation and reconstruction methods using point clouds

The resulted point clouds after the registration task are normally not used as an end product. They are usually modelled as surfaces and volumes with the use of a variety of reconstruction methods. A lot of research work was realized in this regard. However, most of it is making use of the Manhattan World constraints and it is dealing with high accurate data, coming from TLS. JEKE et al. (2009) are making use of the Manhattan World constraints and a graph structure for fitting cuboids to a point cloud, which are further merged to rooms and corridors. BUDRONI & BÖHM (2009) are performing a plane sweeping algorithm for identifying walls in a Manhattan World scenario. The approach used by VALERO et al. (2012) for wall segmentation is based on the work done by OKORN et al. (2010) and is finally delivering a boundary representation model. PREVITALI et al. (2014) are presenting an automatic reconstruction algorithm from TLS point clouds to semantically enriched models using terrestrial laser scanner data. BECKER et al. (2015) are proposing a grammar-based approach for automatic reconstructing 3D interiors from laser scanner point clouds. In his PhD thesis, KHOSRAVANI (2016) proposed an approach for obtaining topological correct indoor model from Kinect measurements. However, this solution is limited to small scale data (a room, a hallway, etc.), while we aim at reconstructing an entire floor. Also, part of the point clouds used within our investigations are less accurate than the point clouds provided by Kinect. Therefore, the algorithm from section 4 is proposed.

3 Indoor Mapping Systems and Delivered Data

In this work, two different systems, Google Tango tablet and NavVis M3 trolley, were used for the indoor mapping task. Their characteristics are presented in the followings.

3.1 Google Tango tablet

Being firstly introduced by Google in 2014, the Tango tablet (WIKIPEDIA TANGO 2017) is a development kit with 3D motion detection and depth measurement capabilities. Among sensors normally available for mobile devices, like accelerometer, gyroscope, GPS, etc., it also integrates a 3D depth sensor, composed by an infrared (IR) projector and IR sensor, and a wide-angle motion tracking camera (fisheye camera) (Fig. 1 (a)).

The depth is perceived with the help of infrared structured light (FOFI et al. 2004), giving the device the possibility of measuring the distance to the surrounding objects in a range of 0.5 – 4 m, with approximately 1% accuracy. The resulted depth data can be obtained in form of a point cloud or a

textured mesh, its quality being directly influenced by the light source lighting the objects and by their reflectivity (GOOGLE TANGO 2016a).

The device can track its own position and orientation in space (its pose) by making use of visual-inertial odometry (LI & MOURIKIS 2012). Hence, the device's pose combines the change in position, determined by tracking the features in the images from the motion tracking camera, and the rotation and acceleration changes, coming from the inertial motion sensors. However, this concept has limitations, over time the device poses being affected by drift (GOOGLE TANGO 2016b). In order to reduce the aforementioned drift and to estimate the device's position within the already measured areas, the device has area learning capabilities (GOOGLE TANGO 2017) which performs Simultaneous Localization and Mapping (THURN & LEONARD 2008).

As a mobile device used for indoor scanning, the Tango tablet is very easy to handle and it gives to the user the flexibility of moving through the space, which a fixed scanner cannot offer. Nevertheless, it has limited resources which influence also the space extend which can be measured in a session.

3.2 NavVis M3 trolley

The M3 trolley is a 3D mapping solution, being released in 2014 by the company NavVis (NAVVIS M3 TROLLEY 2017). It integrates three Hokuyo UTM-30LX laser scanners, a HDR panoramic camera head composed of six cameras, an inertial measurement unit (IMU), WiFi sensors, a magnetometer and an on-board computer (Fig. 1 (b)).

Despite the 3D point clouds, acquired in a range between 0.1 - 30 m, with an accuracy of approximately 3 cm, the M3 trolley delivers 360° panorama images with the possibility of virtually navigate through them afterwards. Also, these panorama images are used, during post-processing, to generate coloured point clouds, with different level of details up to 5 mm resolution (NAVVIS DEMO DATA 2017).

During the data acquisition, quality maps, as 2D floor plans, are generated for each measurement session, by using a graph-SLAM algorithm (NAVVIS M3 TROLLEY 2017). This algorithm is based not only on the 2D laser input, but also on the panoramic images, IMU, WiFi and magnetic field measurements (BASSIER et al. 2015), and also enables the M3 trolley to centimeter-accurate estimate its position within the 2D map. This feature helps the user to be aware of the space still unmapped and to improve the data quality in terms of completeness.

Even if the M3 trolley allows to maximum measure 45 minutes in a session, its 1.98 m height together with the normally low-height of the door openings is forcing the operator to pass it from room to room in compact form, which implies the need of a new calibration for the IMU sensors and so the start of a new session. Target points, previously set up, before the measurement took place, are called anchor points and are considered to be constrains for the SLAM back-end approach (GRISSETTI et al 2010), enabling, in post-processing, the 2D maps optimization and the automatic registration of the point clouds. For each measurement session at least 2 anchor points have been used. Being designed especially for indoor scanning, the NavVis M3 trolley, together with the embedded algorithms, succeeded to exceed classical fixed scanner solutions in terms of efficiency and flexibility. However, it may not be such a suitable solution for staircases, very narrow spaces, or with very low ceiling, due to its construction.

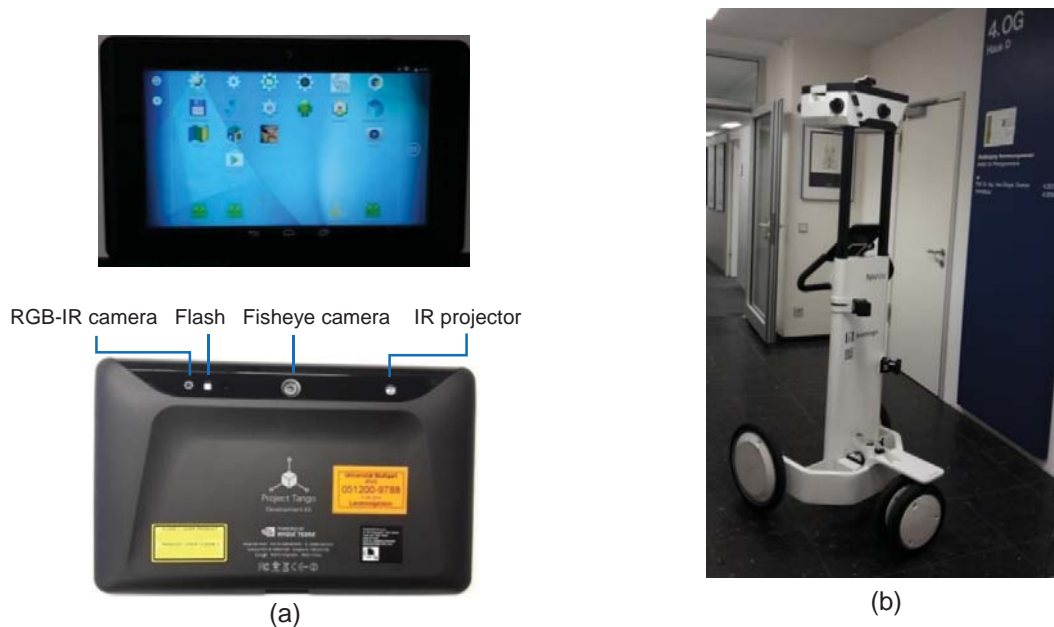


Fig. 1: (a) Google Tango tablet; (b) NavVis M3 trolley

4 Our Algorithm for 3D Indoor Interpretation

The goal of our approach is to automatically interpret 3D indoor point clouds. We decided to work with point clouds instead of meshes (which also implies the normal computation), in order to have a flexible algorithm, usable for a variety of indoor mapping systems, which usually deliver point clouds as a raw data. Therefore, the proposed algorithm consists of: (1) a 3D data pre-processing step and (2) a segmentation and semantic interpretation step. For enabling comparison between different measurement systems, the input data consists of point clouds of the same floor of one office building. In order to describe the algorithm, an indoor dataset acquired with Google Tango tablet is considered.

4.1 3D data pre-processing

The pre-processing step is preparing the 3D data for the segmentation. The floor and the ceiling are firstly removed by applying a height filtering. Considering that the errors coming with the data, acquired with low cost sensors, are strongly influencing the later processing, a sparse outlier removal algorithm is applied. This algorithm is considering for every point the mean distances to k -nearest neighbours (in our case $k = 20$) and is computing their mean μ and standard deviation σ . All the points having the mean distance outside the interval $\mu \pm \alpha \cdot \sigma$ are considered outliers. The standard deviation multiplier threshold is set to be $\alpha = 1$ after multiple experiments which proved its applicability. Fig. 2 (a) and (b) are showing the effect of applying the outlier filtering. For making the further process more efficient, a downsampling method based on octree voxel grid filtering, is used. This filtering is replacing all the points bounded by one voxel with the corresponding centroid. The selected size for the voxel grid was chosen to be 5 cm, considering the overall noise of the point cloud. This way, the size of the point cloud is reduced with a factor

of 5. Some data irregularities are eliminated with the help of a resampling algorithm, which implies a Moving Least Squares surface reconstruction method, firstly introduced by LANCASTER & SALKAUSKAS (1981). This method smooths and recreates the missing parts of a point cloud by using a higher order polynomial interpolation between the neighbouring points. Through this, part of the registration errors is eliminated and also the curvature of flat surfaces is reduced by local plane fitting, like Fig. 2 (c) and (d) are showing.

The 3D data pre-processing was implemented in C++ using the Point Cloud Library (PCL) (RUSU & COUSINS 2011).

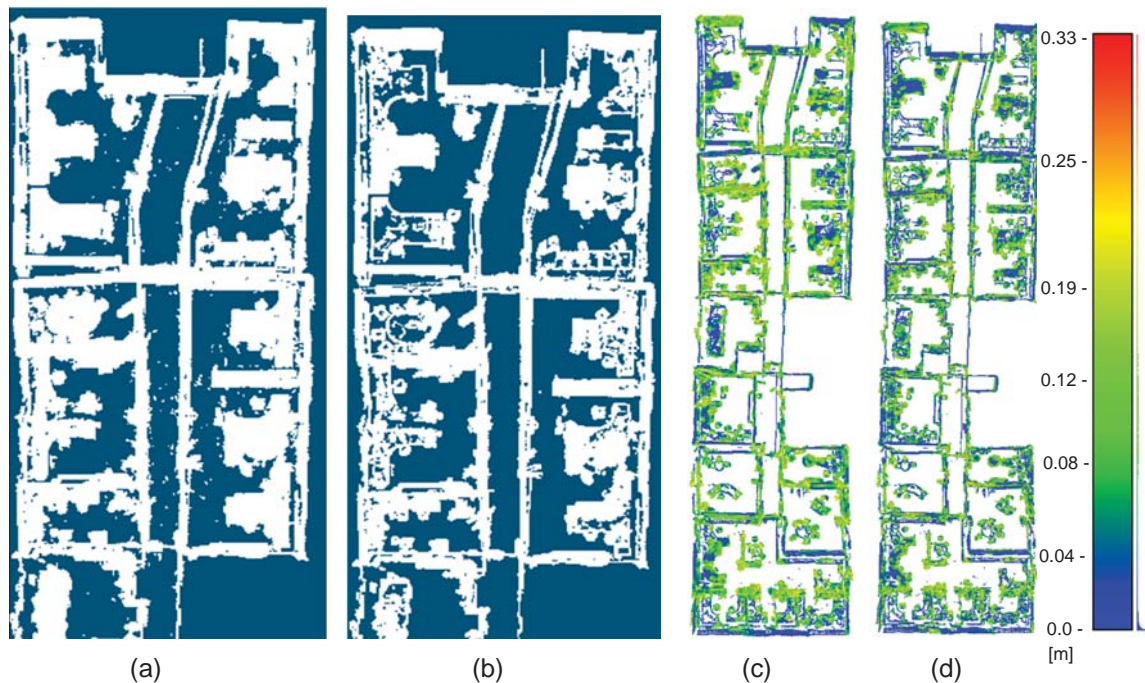


Fig. 2: (a) Original point cloud; (b) After sparse outlier removal; (c) Point cloud curvature after downsampling; (d) Point cloud curvature after Moving Least Square surface reconstruction [m]

4.2 Segmentation and semantic interpretation

Interpreting the 3D data semantically implies to distinguish between a large variety of indoor objects, such as walls, chairs, tables, cupboards, etc., and for this reason different algorithms are normally used, adapted to the object to be detected.

In this current work and as a first step in modelling the 3D interior, walls are aimed to be detected. Therefore, similar to KHOSRAVANI (2016), the 3D points are filtered in a height range in order to remove the furniture and to obtain the wall structure. The height thresholds should be selected according to the typical height of the furniture, which is in our case less than 1.7 m, and the ceiling decorations, which are approximately 0.5 m under the ceiling. Therefore the 3D points between 1.7 – 2.3 m were considered for further investigations. This filtering can be iteratively applied, for semantically interpreting different objects. After the walls are detected, the corresponding 3D points should be removed from the dataset and a new filtering range should be applied according to the new object aimed to be detected. For example, in a range of 0.5 m down from the previous maximum (2.3 m in this case) tall cupboards and bookshelves can be detected.

After the filtering, the points extracted in a height range between 1.7 m and 2.3 m are clustered by applying a model-based segmentation algorithm. We propose an iterative RANSAC method, which detects 3D lines describing the walls (see Fig. 3). The lines are then passed through a clipping and filtering process in order to retain only the line segments describing the reality. This is realized by following steps:

- Step 1: The filtered 3D point cloud and the 3D lines are transformed into a 2D map grid and 2D lines, respectively, by projecting them onto a horizontal plane, parallel with the floor. Each cell value represents the number of points falling into that cell. The grid cell dimensions are established according to the point sampling distance and the accuracy conditions.
- Step 2: Each projected 2D line is segmented by checking the corresponding grid values, in order to detect possible line breaks (segments of lines, which are not overlapping 3D point regions) and to save the continuous ones (Fig. 4 (a)). It is considered that the data may have interruptions, due to occlusions or low-reflectance surfaces, and for this reason, the continuous segments are accepted to have discontinuities less than 0.5 m.
- Step 3: The resulted segments are filtered by removing the ones smaller than 1 m resulted from the clipping process, as it can be seen in Fig. 4 (b).
- Step 4: The remaining segments are clustered according to their orientation and distance to the other segments (Fig. 4 (c)).
- Step 5: One line segment is fitted for each cluster in order to have the individual walls represented by individual line segments, like it is shown in Fig. 4 (d).
- Step 6: In order to divide the indoor space into rooms, the wall segments are intersected with the neighbouring ones and according to the specific situations they are extended or trimmed (Fig. 4 (e)).

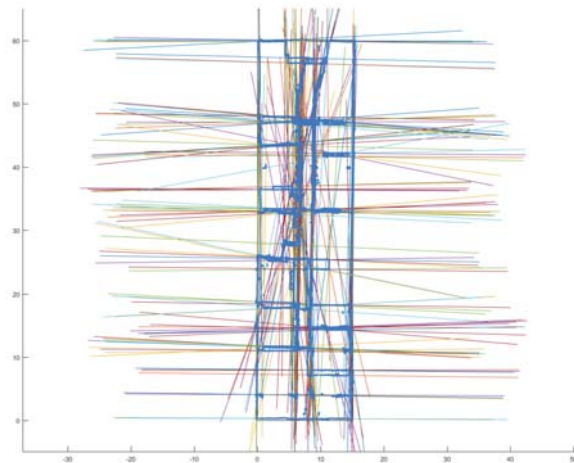


Fig. 3: 2D projection of the 3D point cloud (blue points) and the 3D lines resulting from the iterative RANSAC

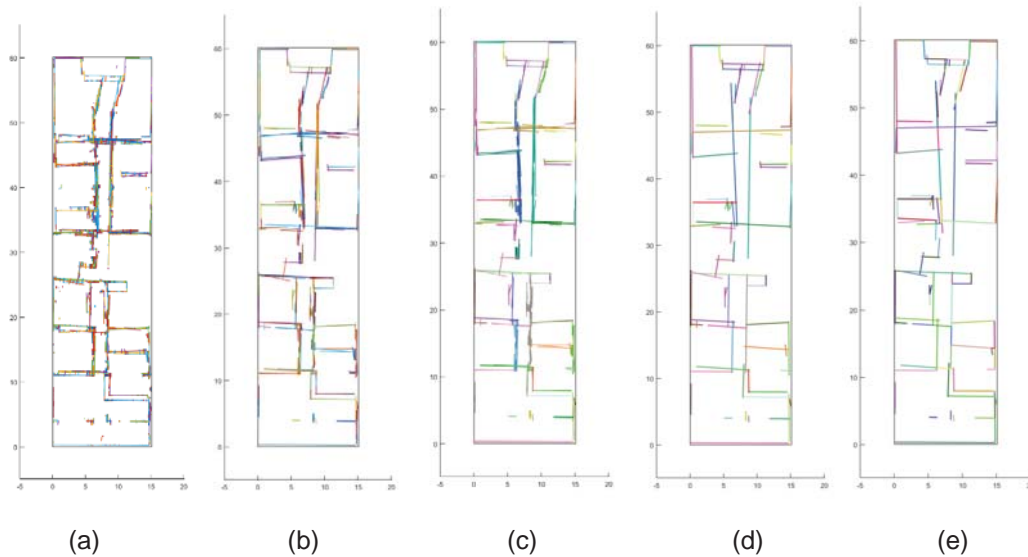


Fig. 4: Line clipping and filtering process for obtaining the wall segments (random colours are associated to different lines): (a) line segmentation; (b) filtering small segments; (c) segment clusters; (d) line fitting; (e) segment intersections

The proposed algorithm is very flexible, not being restricted to the Manhattan World constraints, and accurate enough to detect the main wall structures. However, some problems, addressed in section 5, appear due to the noisy and incomplete character of the data. By imposing topological constraints, the results could be further improved.

5 Accuracy evaluation and comparison

The algorithm presented in section 4 is designed for detecting wall structures in 3D point clouds coming from low cost sensors. A comparison between the 3D point clouds coming from different systems, i.e. the Tango tablet and the NavVis M3 trolley, reveals the challenges which were needed to be overcome and to which extent this succeeded.

5.1 Tango tablet versus NavVis M3 trolley for obtaining indoor models

For automatically interpreting 3D indoor data, it is considered an indoor dataset acquired with Google Tango tablet and it is analysed and interpreted in parallel with data, of the same indoor space, coming from NavVis M3 trolley.

Automatic point cloud registration was not subject of research of this work. However, a significant aspect is that the Tango data used for these investigations, representing the 4th floor of our office building, is composed of multiple small datasets, limited by the device overheating. Each individual dataset was captured by using simultaneous localization and mapping, integrated in the mode Area Learning provided by Google Tango (GOOGLE TANGO 2017). This enabled a relatively accurate capture of each individual indoor section. Nevertheless, their registration manually realized, depending on the overlapping regions, could have caused errors, resulting in double walls and object displacements. Even though, the NavVis M3 Trolley acquired also the data in different

sessions, for the individual rooms, the integrated graph-based SLAM algorithm (GRISSETTI et al 2010) made possible the automatic registration for all the datasets.

Though both systems are delivering 3D data useful for a large variety of applications, Fig. 5 is showing their noise with respect to a reference dataset, coming from the laser scanner Leica HDS3000. In the Tango tablet case (Fig. 5 (a)), the computed absolute distances reveal noise, (up to 0.3 m) also for the flat surfaces of non-moving objects, like walls or tall cupboards, which is not the case for the M3 trolley. It can also be observed, that the non-reflecting surfaces, like windows, are causing for both datasets missing parts in the point clouds, affecting especially the further reconstruction of the exterior walls of the building.

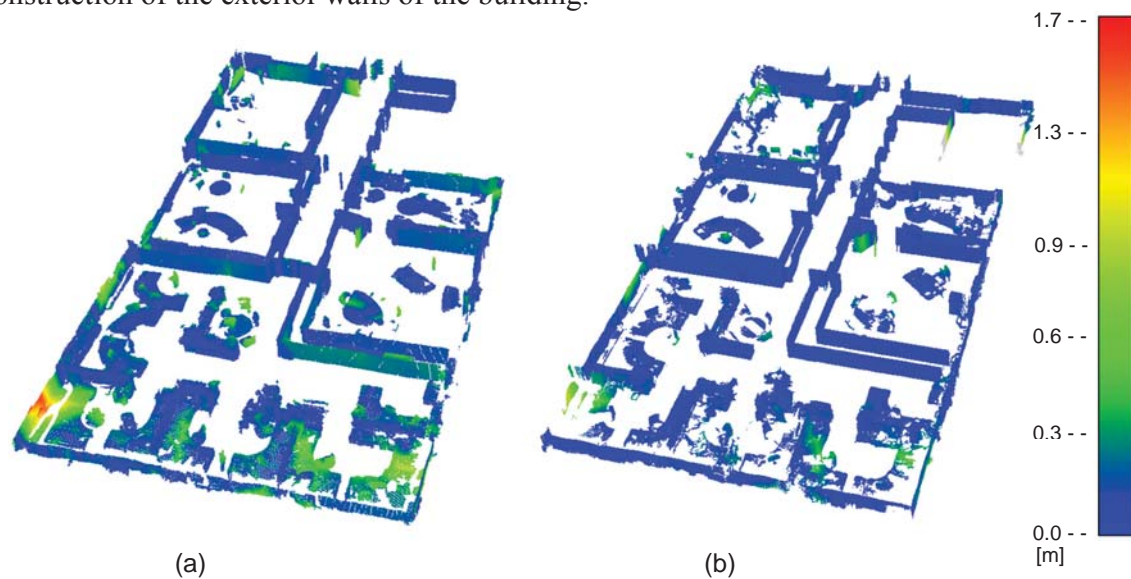


Fig. 5: Cloud to Cloud absolute distance in [m]: (a) between Tango and TLS point clouds; (b) between M3 trolley and TLS point clouds

The distance range of the sensors is a property which influences the accuracy and completeness of the acquired 3D data and therefore also the quality of the interpreted structures. By having the range distance up to 4 m, the Tango tablet is limiting the obtained point cloud to this range, while the NavVis M3 Trolley allows an acquisition up to 30 m. Therefore, for automatically detecting the wall structures in rooms with very high ceiling, one must consider during the processing that considerable big parts of those structures can be missing. This aspect was handled by the proposed algorithm from section 4, by filtering out the points over 2.3 m.

5.2 Quality evaluation of the interpreted wall structures

After applying the proposed algorithm, the resulted wall structures from both datasets are compared with a reference dataset coming from the laser scanner Leica HDS3000, acquired a couple of years ago. In this way it was possible to realize the influence of data quality onto the resulted interpreted structures. In Fig. 6, it is shown the 2D projection of the corresponding 3D point cloud, as blue points, and the detected wall segments, as red lines. It is considered that a wall is detected only when the corresponding segment is complete, from one wall junction to the other. Thus, 44 from 55 segments (80%) were detected for the Tango tablet dataset, and 47 from 55 segments (85.5%) for the M3 trolley. However, the main majority of the undetected wall segments

are located on the building shell, being mainly made of windows. As it was previously shown in section 5.1., non-reflecting surfaces, like windows are affecting the reconstruction process. This issue can be overcome by adding some knowledge, like the ground plan of the building. Furthermore, some wall segments could be misplaced due to some object parts, different than walls, still remained after the pre-processing step.

Both datasets made possible to automatically detect the main wall structures, but in some cases they were shifted from the reference position. For example, Fig. 7 shows a closer detail of Fig. 6, where, in the case of Tango dataset, the wall segments are shifted in centimetre range. This is visible only if they are compared with a reference dataset, coming from TLS. Thus, it is proved that the displacement is not influenced by the proposed algorithm, but of the noisy character of the data and possible remaining registration errors.

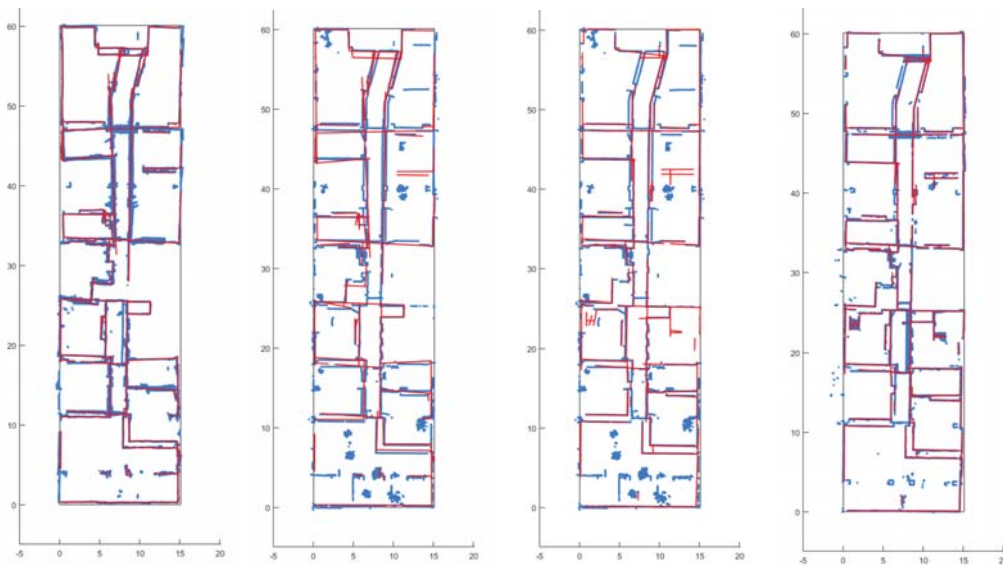


Fig. 6: Detected wall segments from: (a)Tango tablet and Tango tablet point cloud; (b) Tango tablet and TLS point cloud; (c)M3 trolley and TLS point cloud; (b) M3 trolley point cloud and M3 trolley point cloud

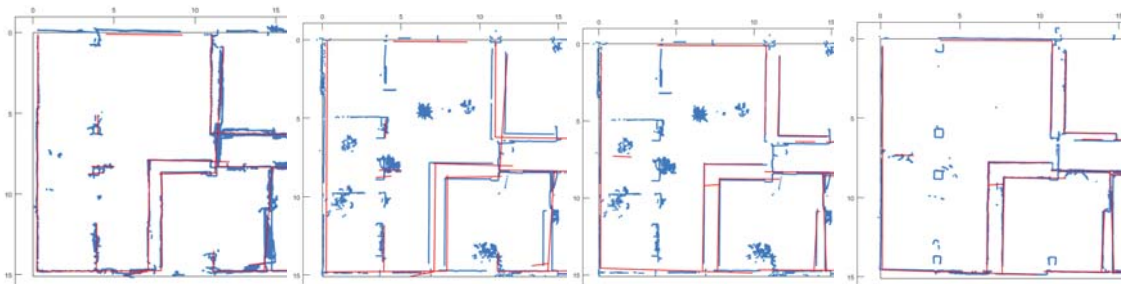


Fig. 7: Detail of detected wall segments from: (a)Tango tablet and Tango tablet point cloud; (b) Tango tablet and TLS point cloud; (c)M3 trolley and TLS point cloud; (b) M3 trolley point cloud and M3 trolley point cloud

6 Conclusions

We presented an evaluation of an IMMS system (NavVis M3 trolley) and a low-cost device (Google Tango tablet). Also, it was investigated their potential for delivering interpreted indoor structures. For doing that, we proposed a wall extraction algorithm, which proved to be robust and flexible, not being restricted to Manhattan World constraints. By comparing the results with a TLS reference dataset, it was proved that the main wall structures were detected, the missing ones being corresponding for surfaces which do not reflect IR light, like windows. Only few detected segments, coming from Tango tablet dataset have been displaced in centimetre range, due to possible registration errors. On the one side, there are applications which require accurate indoor models, like mapping indoors of industrial facilities and for this, a system like NavVis M3 trolley will be suitable. On the other side, for some other applications, like indoor navigation, a mobile device with Google Tango's capabilities will successfully perform the task. Furthermore, a way to overcome these possible displacement problems is to use a formal grammar for modeling the indoor environment, which is a subject of further work.

7 Acknowledgements

This work is supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), grants FR 823/25-3 and RO 1086/17-3, and the European Social Fund (ESF) as well as the Ministry Of Science, Research and the Arts Baden-Württemberg.

8 References

- BASSIER, M., YOUSEFZADEH, M. & VAN GENECHTEN, B., 2015: Evaluation of data acquisition techniques and workflows for Scan to BIM. Proceedings of Geo Business.
- BESL, P. J. & MCKAY, N. D., 1992: Method for registration of 3-D shapes. Robotics-DL Tentative, 586-606.
- BOEHM, J., 2014: Accuracy investigation for structured-light based consumer 3d sensors. Photogrammetrie Fernerkundung Geoinformation, **2014**(2), 117-127.
- BOSSE, M., ZLOT, R. & FLICK, P., 2012: Zebedee: Design of a Spring-Mounted 3-D Range Sensor with Application to Mobile Mapping. IEEE Trans. Robot. **28**(5), 1104-1119.
- BUDRONI, A. & BÖHM, J., 2009: Toward Automated Reconstruction of Interiors from Laser Data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38**(5).
- DIAKITÉ, A. A. & ZLATANOVA, S., 2016: First experiments with the Tango tablet for indoor scanning. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **III-4**, 67-72.
- DOTPRODUCT DPI-8, 2017: <https://www.dotproduct3d.com/DPI8.php>, last access 16.01.2017.
- FOFI, D., SLIWA, T. & VOISIN, Y., 2004: A comparative survey on invisible structured light. Electronic Imaging 2004, International Society for Optics and Photonics, 90-98.
- GEOSLAM, 2017: <http://geoslam.com/hardware-products/zeb1/>, last access 16.01.2017.

- GOOGLE TANGO, 2016a: <https://developers.google.com/tango/overview/depth-perception> (last updated 07.12.2016), last access 16.01.2017.
- GOOGLE TANGO, 2016b: <https://developers.google.com/tango/overview/motion-tracking> (last updated 10.06.2016), last access 16.01.2017.
- GOOGLE TANGO, 2017: <https://developers.google.com/tango/overview/area-learning> (last updated 05.01.2017), last access 16.01.2017.
- GRISSETTI, G., KUMMERLE, R., STACHNISS, C. & BURGARD, W., 2010: A tutorial on graph-based SLAM. *IEEE Intelligent Transportation Systems Magazine* **2**(4), 31-43.
- JENKE, P., HUHLE, B. & STRASSER, W., 2009: Statistical reconstruction of indoor scenes. In Proc. of 17. Int. Conf. in Central Europe on Comp. Graphics, Vis. and Comp. Vision (WSCG '09).
- KHOSRAVANI, A. M., 2016: Automatic Modeling of Building Interiors Using Low-Cost Sensor Systems. *Deutsche Geodätische Kommission, Reihe C, Nr. 767*, München.
- KURIAN, A. & MORIN, K. W., 2016: A fast and flexible method for meta-map building for ICP based SLAM. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B3, 273-278.
- LANCASTER, P. & SALKAUSKAS, K., 1981: Surfaces generated by moving least squares methods. *Mathematics of computation* **37**, 141-158.
- LASKAR, Z., HUTTUNEN, S., HERRERA, D., RAHTU, C. E. & KANNALA, J., 2016: Robust loop closures for scene reconstruction by combining odometry and visual correspondences. *IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, 2603-2607.
- LEONARD, J. J. & DURRANT-WHYTE, H. F., 1991: Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation* **7**, 376-382.
- LI, M. & MOURIKIS, A. I., 2012: Improving the accuracy of EKF-based visual-inertial odometry. *IEEE International Conference on Robotics and Automation (ICRA)*, 828-835.
- MSDN KINECT, 2017: <https://msdn.microsoft.com/en-us/library/jj131033.aspx>, last access 16.01.2017.
- NAVVIS DEMO DATA, 2017: <http://www.navvis.com/products/demodata/>, last access 16.01.2017.
- NAVVIS M3 TROLLEY, 2017: <http://www.navvis.com/products/m3-trolley/>, last access 16.01.2017.
- NEWCOMBE, R. A., DAVISON, A. J., IZADI, S., KOHLI, P., HILLIGES, O., SHOTTON, J., MOLYNEAUX, D., HODGES, S., KIM, D. & FITZGIBBON, A., 2011: KinectFusion: Real-time dense surface mapping and tracking. *10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 127-136.
- OKORN, B., XIONG, X., AKINCI, B. & HUBER, D., 2010: Toward automated modeling of floor plans. *Proceedings of the symposium on 3D data processing, visualization and transmission* **2**.
- PREVITALI, M., BARAZZETTI, L., BRUMANA, R. & SCAIONI, M., 2014: Towards automatic indoor reconstruction of cluttered building rooms from point clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **II-5**, 281-288.
- RUSU, R. B. & COUSINS, S., 2011: 3D is here: Point Cloud Library (PCL). *IEEE International Conference on Robotics and Automation (ICRA)*, 1-4.
- THOMSON, C., APOSTOLOPOULOS, G., BACKES, D. & BOEHM, J., 2013: Mobile Laser Scanning for Indoor Modelling. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **II-5/W2**, 289-293.

THRUN, S. & LEONARD, J. J., 2008: Simultaneous localization and mapping. Springer handbook of robotics, Springer, 871-889.

VALERO, E., ADÁN, A. & CERRADA, C., 2012: Automatic method for building indoor boundary models from dense point clouds collected by laser scanners. Sensors **12**, 16099-16115.

VIAMETRIS IMS 3D, 2017: <http://viametris.info/ims3d/en/>, last access 16.01.2017.

WIKIPEDIA TANGO, 2017: <https://en.wikipedia.org/wiki/Tango> (platform), last access 16.01.2017.

Calibration of a Vehicle Camera System with Divergent Fields-of-view in an Urban Environment

ALEXANDER HANEL¹ & UWE STILLA¹

Abstract: Multi camera systems are used in modern cars for environment perception. To combine information extracted from different cameras, their relative pose has to be known. Therefore, a multi camera system can be calibrated. In this contribution, a calibration approach using the structure-from-motion (SfM) method is described. From images of urban structures, 3d points are calculated using SfM. These 3d points are used as reference to estimate the pose of the vehicle cameras. Test images are taken in a car park with a DSLR camera, together with two action cameras mounted in a test vehicle recording the car park. More than 26,000 3d points representing the car park can be obtained. With these points, the poses of the vehicle cameras are estimated successfully. The distance between the vehicle cameras shows the highest deviation from laser reference measurements for images showing a low point density area of the car park.

1 Single and system camera calibration

The road the automotive industry is driving down currently is characterized by autonomously driving cars. To allow advanced driver assistance systems to control a car reliably, they need

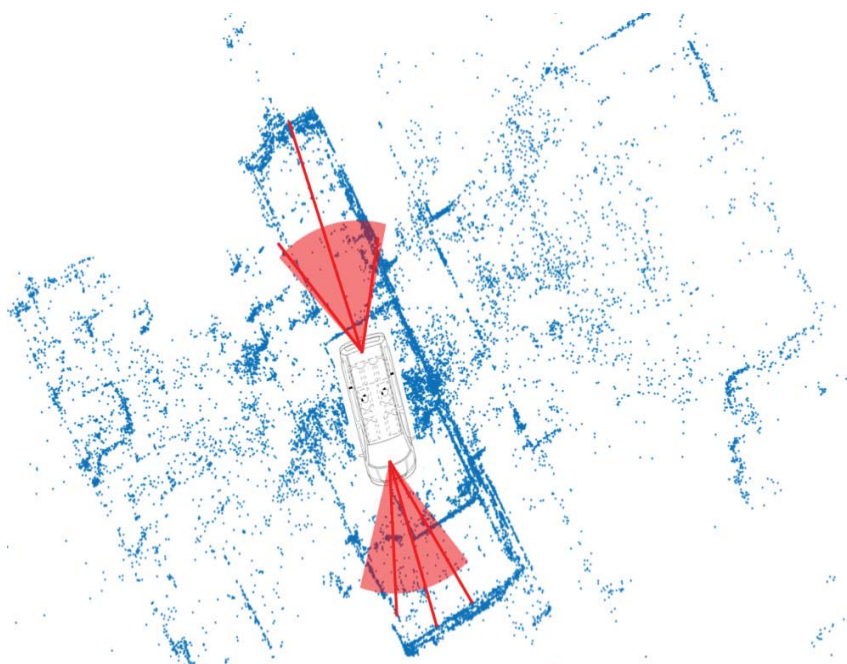


Fig. 1: Top-down view on a car, whose cameras with divergent fields-of-view (in red) can be orientated using 3d points (blue) generated for the calibration environment.

information about the car's environment. To capture this information, cars can be equipped with several cameras observing the environment. As the relative position of environment objects to the car is important, the positions and orientations (pose) of the vehicle cameras relative to the car have to be known.

To determine the poses of the vehicle cameras, a system calibration can be performed. To keep costs low, the number of cameras is kept small, too.

¹ Technische Universität München, Photogrammetry & Remote Sensing, Arcisstraße 21, D-80333 München, E-mail: [alexander.hanel, stilla]@tum.de

So, their fields-of-view are typically non-overlapping, making stereo camera calibration with the same planar calibration pattern seen in all cameras impossible.

Due to their wide field-of-view, so called “action cameras” can cover a huge part of the car’s environment despite their small number. Furthermore, a removable mount available for such cameras on the car windows with a suction pad can avoid in Germany problems with the car registration caused by additional “built-in” installations in the car.

Therefore, it has to be considered, that the wide field-of-view might cause large image distortions requiring reliable single camera calibration. In addition, the mounting on the windshield might not be rigid over time caused by mechanical movements during a car drive. To check the poses again and again, the system calibration has to be repeated from time to time.

To perform the calibration, a suitable calibration environment has to be selected. This environment has to be large enough to contain cars. Therefore, calibration laboratories are typically not feasible. Urban environments, as for example car parks, are large enough. As it might not be allowed by the owner of the car park and to keep costs and time effort low, the calibration procedure should work without manually placed reference marks. In addition, the number of placed reference marks is typically limited due to the available space. The number of reference marks extracted automatically using image feature descriptors can be higher. In contrast, it is not possible to provide 3d reference information for automatically extracted reference marks in advance.

All three mentioned tasks, single camera calibration, system calibration and generation of 3d reference information can be performed with the structure-from-motion method (Fig. 1). Therefore, in this contribution, a structure-from-motion software toolbox will be used to complete these tasks.

2 Related work

2.1 Camera calibration

Single camera calibration can be divided into photometric calibration (KINGSLAKE 1983; ASADA ET AL. 1996; KRAWCZYK ET AL. 2005), radiometric calibration (HEALEY & KONDEPUDY 1994; SEON & POLLEFEYS 2008; MITSUNAGA & NAYAR 1999) and geometric calibration (BROWN 1971; FRASER 2013; HARTLEY & ZISSERMAN 2003). Geometric camera calibration is used to estimate the interior orientation as well as image distortion parameters.

Geometric camera calibration performed simultaneously for a multi-camera system can be used to obtain the relative pose between the cameras (ZHANG & PLESS 2004). This requires typically a calibration pattern seen by all cameras. To calibrate a vehicle camera system, the car has to be placed in front of a calibration pattern to take the calibration images (DICK AND RICKS 2015). Such expensive calibration patterns are typically only available in calibration laboratories (SAFE CAR NEWS 2016).

There are different toolboxes for camera calibration available. The *camera calibrator app* for MATLAB (2016) bases on the algorithm of ZHANG (2000) and requires a set of images of a planar checkerboard pattern for calibration. Its stereo calibration function allows calibrating a rigid stereo camera system with overlapping fields-of-view. LI ET AL. (2013) provide another Matlab calibration toolbox using a feature descriptor-based calibration pattern, which can handle non-overlapping fields-of-view, as long as each camera looks onto a part of the planar calibration pattern.

To sum up, the mentioned solutions cannot handle the calibration of camera systems, where calibration patterns are not available or where the cameras cannot look onto the same planar calibration pattern.

2.2 Structure from motion

Structure-from-motion is a photogrammetric method to estimate 3d coordinates (‘structure’) from an image set. Opposite to rigid stereo camera systems (e.g. in HANEL ET AL. 2016), motion between the camera poses at different acquisition times creates the necessary stereo baseline to estimate the 3d coordinates (KOENDERINK & VAN DOORN 1991; STURM & TRIGGS 1996; HÄMING & PETERS 2010).

There are different software toolboxes for structure-from-motion, which use a similar pipeline (SCHÖNBERGER & FRAHM 2016) starting with feature extraction from images and their matching. After that, camera poses are estimated and 3d points triangulated. Finally, camera poses and 3d points are optimized using bundle adjustment. VisualSfM (WU 2007; WU 2011; WU ET AL. 2011) is a SfM toolbox well known in the literature.

3 Vehicle camera calibration in an urban environment

An overview over the method proposed in this contribution is given in **Fehler! Verweisquelle konnte nicht gefunden werden.** First, 3d points of the calibration environment have to be calculated. Second, these 3d points are used as reference to estimate the pose of the vehicle cameras as well as their interior orientation and distortion parameters.

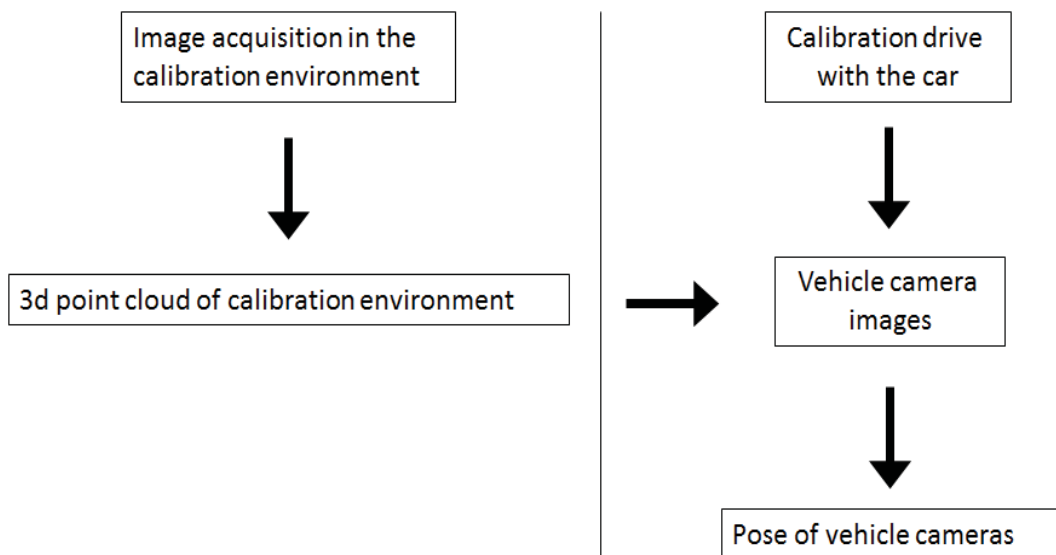


Fig. 2: Method for calibration of a vehicle camera system with divergent fields-of-view in an urban environment.

3.1 Creating a 3d point cloud of the calibration environment

In the first part, images of the urban calibration environment have to be acquired to obtain 3d points of the environment to use them as reference for the vehicle camera calibration.

To select a suitable calibration environment, the following aspects have to be considered. The objects within the environment have to have textured surfaces to allow a high number of 3d points and to avoid misassignments between different parts of the environment. Placing photogrammetric reference marks manually is therefore not necessary. The environment has to be large enough to allow a car to turn around, so that the vehicle cameras look at different reference points within the environment. For example, a car park can be used as calibration environment.

A professional camera with a fixed focal length lens is recommended to be used to ensure high quality images with minimal distortions. While the focus should be set to infinity to achieve a consistent interior orientation, aperture should have a medium value to get a wide depth of sharpness, but to avoid diffraction blur.

SIFT features are extracted from each image. Each image pair is checked for matching SIFT features. The pose of the used images is estimated as well as the 3d point coordinates are calculated for the matched image points using SfM. This results in a 3d point cloud of the calibration environment.

3.2 Estimating the pose of vehicle cameras using the point cloud

Having generated the 3d reference, the calibration of the vehicle cameras can be performed. Therefore, a car equipped with cameras looking at the area around the vehicle has to perform a calibration drive within the urban calibration environment. Images or videos of the vehicle cameras have to be recorded, and if necessary images extracted from the videos.

Temporarily corresponding images of all vehicle cameras are fed into the processing chain. SIFT features are extracted from the vehicle camera images and used to match them to the environment images (section 3.1).

A part of the matches between the vehicle camera images and the environment images relies on SIFT features, which are already used to calculate the 3d point cloud. These 3d points are now used as reference to estimate the pose of the vehicle camera images, as well as their interior orientation and distortion parameters. From a set of temporarily corresponding vehicle camera images, the mean relative pose between the vehicle cameras can be calculated.

4 Experiments

The calibration environment (Fig. 3) contains one parking lot on one side of a driveway and two parking lots on the other side. The environment consists of floor, ceiling and walls made of concrete. The walls and parts of the ceiling are painted white. Some signs are painted onto the walls. The experiments are conducted in the late evening, wherefore no cars are in the calibration environment at that time. The only illumination source in the car park is an artificial lamp, illuminating the environment quite darker than natural daylight.



Fig. 3: a) Sample image of the calibration environment taken with a Nikon D3. Two parking lots can be seen, b) Sample image of the front camera of the car. The sharp shadow on the left is caused by powerful external LED lights, which are not available in a). The reference marks at the wall are not used in this contribution.

A Volkswagen Transporter 5 van is used as test car. Two vehicle cameras are placed. One camera is mounted at the front of the car looking through the windshield to the area in front of the car (Fig. 4). The other camera is mounted at the rear of the car looking through the rear window to the area behind the car.

143 calibration environment images (Fig. 3a) are taken with a Nikon D3 camera with a 24 mm fixed focal length lens, acquiring 13 megapixels images. A tripod is used to keep ISO low, focus is set to infinite and the aperture to $f/8$. The lengths and height of the walls in Fig. 3a are measured with a laser distance meter to calculate ground control points for an Euclidean 3d reconstruction. Four ground control points are marked manually in seven images showing all four points together. The environment images are processed in VisualSfM. SIFT features are extracted and used to match the images. 3d point coordinates of parts within the TUM car park are calculated for the matching points.

To calibrate the vehicle cameras, they record videos with 4 megapixels resolution and 30 frames per second, while the car is driving slowly (< 5 km/h) through the calibration environment. 248 pairs of vehicle images (Fig. 3b) are extracted from the videos, one image pair per second. Videos instead of images are recorded, as the street scenes for later research will be recorded with videos, too. Garmin VIRB Ultra 30 action cameras with 2.73 mm focal length are used as vehicle cameras. Acquisition parameters are set automatically by the cameras. The recording is started and stopped with the Garmin VIRB app, allowing synchronous control of all cameras. During video recording,

the area near the vehicle path is illuminated with powerful permanent LED light used in photographic studios.

The vehicle camera images are processed in VisualSfM, too. They are linked to the environment images with common SIFT features, so that the 3d point coordinates can be used to estimate the pose of the vehicle cameras.



Fig. 4: a) Position (red rectangle) of the front camera, looking through the windshield to the area in front of the car, b) Position of the rear camera, looking through the rear window to the area behind the car. Image brightness strongly increased for better visualization.

5 Results

The 3d point cloud of the TUM car park is created with around 26,000 3d points (Fig. 5). All of the 143 calibration environment images are used and more than 600,000 SIFT features are extracted from them, in average around 4,800 SIFT features per image. More than 126 thousand combined matches (i.e., independent whether a SIFT feature is found in two or more images, it is always counted as one match) are calculated for all image pairs. That means, 3d points can be calculated for around 4 % of all matches.

In the following, the appearance of the point cloud is discussed. According to Fig. 5, the density of the point cloud is highest in the central part of the figure. The closer the 3d points to the camera positions are, the higher their density is. In areas further away from camera positions, the point density is lower. This observation correlates with few aspects. First, in closer areas, the building walls are shown with more details (e.g. wall signs), which can be seen as intensity differences in images. These differences are needed by the SIFT algorithm to extract features. Second, closer areas are pictured in a higher number of images than areas further away, leading to a higher number of matches in closer areas. Third, as the illumination in the car park is quite low, remote areas tend to appear dark in the images, which reduces the intensity differences needed by SIFT.

Accuracy information about the 3d points are not given by VisualSfM. Furthermore, as the only parameter of the interior orientation, the focal length is estimated with $f = 18.23 \text{ mm}$. For comparison purpose, the focal length calibrated in a laboratory environment is $f = 24.29 \text{ mm}$. The difference of around 6 mm might be interpreted, that VisualSfM might require calibration parameters as well as undistorted images as a-priori data.

In sum, as the area with the highest point density is the area used to record vehicle camera videos, the different point densities do not have negative consequences on the estimation of the vehicle camera poses.

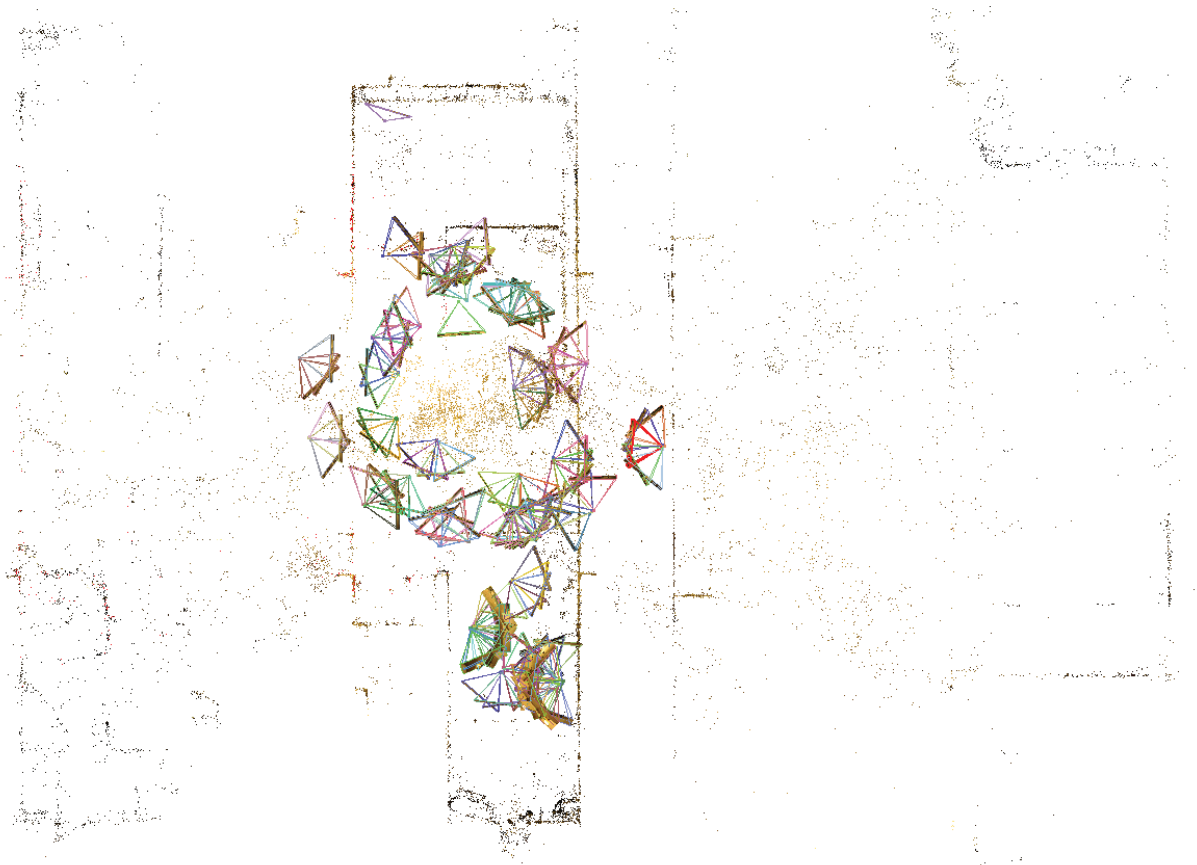


Fig. 5: Top-down view on the 3d point cloud of the TUM car park. Most points are along walls. The drive way goes horizontally from left to right in the middle of the figure, parking lots are orthogonal to it. The area close to the camera positions contains the most 3d points and is used to record the vehicle camera videos.

All 248 image pairs of the front and rear vehicle camera are matched to the environment images, so in total 639 (143 + 248 + 248) images are processed together. 378,867 combined matches are found. Around 10,000 3d points are calculated after adding the front and rear camera images. Previous tests of image pairs of the vehicle cameras have shown a maximum time gap between the two cameras of 0.1 second. As the car is standing or driving very slowly during recording, the image pairs are approximated as temporarily synchronized.

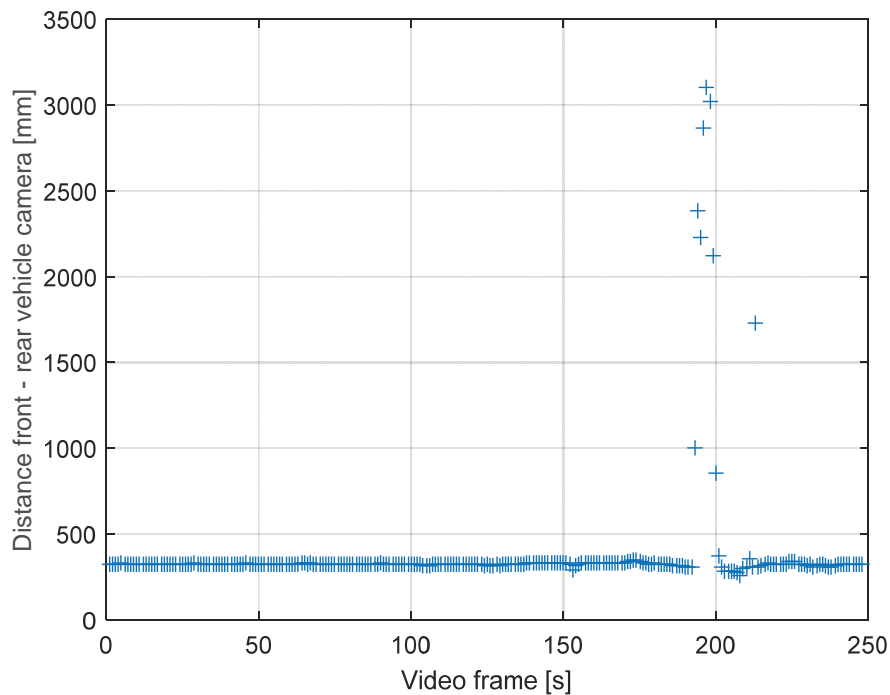


Fig. 6: Distance between the front and rear vehicle camera for the video recorded during the calibration drive. At most times, there is a deviation of the distance over time of smaller than 10 cm, but around frame 200, there is a large deviation of few meters. At that time, the vehicle cameras are looking at low point density areas.

The distance between the front and rear camera calculated for each of the 248 image pairs is used as evaluation metric for the estimated pose (**Fehler! Verweisquelle konnte nicht gefunden werden.**). The distance varies over time within an interval of 10 cm. Only around second 200, there is a way larger deviation of few meters compared to other points of time. At that time, the vehicle cameras are looking to an area (cf. Fig. 5: cameras looking at second 200 to the left and to right part of figure, respectively) of the calibration environment with low 3d point density. That might explain also, why the deviation is increasing to the maximum step by step and decreasing also step by step, as the vehicle moves slowly to and from the area with low point density.

The median distance is 3.26 m, while laser distance measurements show 3.22 m. These values vary only by a few centimetres, which might lead to the conclusion, that the estimated poses are reliable. Again, the missing accuracy information from VisualSfM prevents such conclusion.

To sum up, a car park built of concrete with relatively low structured surface allows to obtain a high number of several thousand 3d points. As these points are spread within the calibration area, they allow to calibrate the vehicle cameras even without overlapping fields-of-view. Neither standard planar calibration patterns, nor patterns designed specifically for the calibration environment are necessary.

As the generation of the 3d reference information and the vehicle camera calibration are two independent steps, the calibration can be repeated every time the car is driving through the part of the car park with the 3d information available.

Processing the images in VisualSfM is easy to use. But there are limitations in the estimated camera parameters (only interior, distortion) as well as there is no accuracy information about the obtained results.

6 Conclusion

In this paper, an approach to calibrate a vehicle camera system with divergent fields-of-view using urban structures has been proposed. It is shown, that several thousand 3d points can be obtained from images of urban structures using SfM. For images of vehicle cameras, taken during a drive through the urban structure, these 3d points can be used as reference to estimate their poses. The distance between the front and rear camera calculated from a set of image pairs, deviates in median only few centimetres from laser distance measurements. But it has been shown, that a high deviation correlates with a low number of 3d points.

For further investigation, the accuracy of the obtained 3d points as well as of the vehicle camera poses should be analysed. Furthermore, pre-processing steps for the images before using them for 3d point cloud generation should be considered to ensure a high accuracy pose estimation. This might involve separate single camera calibration, as well as the use of highly reliable ground control points, like for example obtained from tachymeter measurements.

7 References

- ASADA, N., AMANO, A. & BABA, M., 1996: Photometric calibration of zoom lens systems. *Proceedings of the 13th International Conference on Pattern Recognition* **1**, 186-190.
- BROWN, D. C., 1971: Close-range camera calibration. *Photogrammetric Engineering* **78**(8), 855-866.
- DICK AND RICKS AUTO INTERIORS INC., 2015: Windshield Replacement. <http://www.dickandricks.com/blog/windshieldsreplacement-not-the-same-glass-repair-as-yr-last-vehicle>, last access on 2015-12-08.
- FRASER, C., 2013: Automatic Camera Calibration in Close Range Photogrammetry. *Photogrammetric Engineering and Remote Sensing* **79**(4), 381-388.
- HANEL, A., HOEGNER, L. & STILLA, U., 2016: Towards the influence of a car windshield on depth calculation with a stereo camera system. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41**(B5), 461-468.
- HARTLEY, R. & ZISSERMAN, A., 2003: *Multiple View Geometry in Computer Vision*. 2nd edition. Cambridge University Press, New York, USA.
- HEALEY, G. E. & KONDEPUDY, R., 1994: Radiometric CCD camera calibration and noise estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**(3), 267-276.
- HÄMING, K. & PETERS, G., 2010: The structure-from-motion reconstruction pipeline – A survey with focus on short image sequences. *Kybernetika* **46**(5), 926-937.
- KINGSLAKE, R., 1983: *Optical system design*. Academic Press, Waltham, USA.
- KOENDERINK, J. J. & VAN DOORN, A. J., 1991: Affine structure from motion. *Journal of the Optical Society of America A* **8**(2), 377-385.

- KRAWCZYK, G., GOESELE, M., SEIDEL, H.-P., 2005: Photometric Calibration of High Dynamic Range Cameras. Research Report. Max-Planck Institute for Informatics, Saarbrücken, Germany.
- LI, B., HENG, L., KÖSER, K. & POLLEFEYS, M., 2013: A Multiple-Camera System Calibration Toolbox Using a Feature Descriptor-Based Calibration Pattern. 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 1301-1307.
- MATLAB, 2016: Single Camera Calibration App – MATLAB & Simulink. <https://de.mathworks.com/help/vision/ug/single-camera-calibrator-app.html?requestedDomain=www.mathworks.com>, last access on 2017-01-10.
- MITSUNAGA, T. & NAYAR, S. K., 1999: Radiometric self calibration. IEEE Computer Society Conference on Computer Vision and Pattern Recognition **1**, 374-380.
- SAFE CAR NEWS, 2016: National Windscreen launches UK-wide ADAS calibration service, Website, <http://safecarnews.com/national-windscreen-launches-uk-wide-adas-calibration-service-ma793/>, last access on 2017-01-10.
- SCHÖNBERGER, J. L. & FRAM, J.-M., 2016: Structure-from-Motion Revisited, IEEE Conference on Computer Vision and Pattern Recognition, 4104-4113.
- SEON, J. K. & POLLEFEYS, M., 2008: Robust Radiometric Calibration and Vignetting Correction. IEEE Transactions on Pattern Analysis and Machine Intelligence **30**(4), 562-576.
- STURM, P. & TRIGGS, B., 1996: A factorization based algorithm for multi-image projective structure and motion. European Conference on Computer Vision, 709-720.
- WU, C., 2007: SiftGPU: A GPU implementation of Scale Invariant Feature Transform (SIFT), <http://cs.unc.edu/~ccwu/siftgpu>, 2017-01-10.
- WU, C., 2011: VisualSFM: A Visual Structure from Motion System, <http://ccwu.me/vsfm/>, last access on 2017-01-10.
- WU, C., AGARWAL, S., CURLESS, B. & SEITZ, S. M., 2011: Multicore Bundle Adjustment, CVPR 2011, 3057-3064.
- ZHANG, Q. & PLESS, R., 2004: Extrinsic Calibration of a Camera and Laser Range Finder (improves camera calibration). IEEE/RSJ International Conference on Intelligent Robots and Systems, Proceedings **3**, 2301-2306.
- ZHANG, Z., 2000: A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence **24**(11), 1330-1334.

Hochauflösende Multisensor-Systeme

GERHARD KEMPER¹

Zusammenfassung: Die Entwicklung im Bereich der Sensorik ist rasant und höhere Leistung wird immer kompakter. Neben Laserscannern und RGB Flächensensoren werden zunehmend auch Thermalkameras, NIR-Sensoren aber auch Hyperspektralscanner eingesetzt. Auch UV-Sensoren sind für spezielle Aufgaben im Bereich des Monitoring zunehmend von Interesse. Die Integration solcher Sensoren in flugzeugbasierte Vermessungssysteme hat eine bisher nie erreichte Qualität und Flexibilität erreicht. Neben einfachen preiswerten Systemen für Aufgabenstellungen im Bereich Monitoring sind auch metrische Sensoren verfügbar, welche hochauflösende geometrisch präzise Daten zur Verfügung stellen. Die Kombination solcher Sensorköpfe ermöglicht es heute, nutzerspezifische Plattformen für flexible Aufgaben zu entwickeln.

1 Dual- und Triple RGB Sensoren

"Wide Corridor Mapping" ist insbesondere in strukturell wenig entwickelten Ländern eine wichtige Fragestellung. Nicht immer sind großformatige Kameras verfügbar oder für die Aufgabenstellung geeignet. Für die Befliegung von großen Gebieten im Sudan wurde basierend auf mittelformatigen Phase-One Kameras zunächst eine 2-Kopf-Lösung (KEMPER 2012) bestehend aus 2 PhaseOne iXA-180 und 110 mm Objektiven, entwickelt und erfolgreich in mehreren Projekten getestet. Für ein Eisenbahnprojekt wurde es notwendig eine noch breitere Trasse effizient abdecken zu können. So wurde eine 3-Kopflösung entwickelt und eingesetzt. Als Produkte sollte ein DGM abgeleitet sowie ein Orthophoto generiert werden. So wurden drei 80 MPix Kameras mit 150 mm Brennweite und einer Überlappung von 20% zwischen den Sensoren in Querausrichtung montiert. Daraus resultiert eine Flächendeckung von 26.800 * 7.752 Pixeln was einer Flächendeckung von 2.680m * 775m entspricht. Die Höhengenaugigkeit ist wegen der langen Brennweite nur 32 cm, was jedoch der geforderten Spezifikation des DGM von besser 50 cm lag. Somit war eine geeignete Sensorkombination gefunden.

Die Problematik der Passpunktsignalisierung in solchen Gebieten ist ja bekannt, so wurde auf direkte Georeferenzierung und wenigen Passpunkten in den urbanen Arealen gesetzt. Hierzu war ein Sensor zu Sensorkalibration sowie der Bestimmung der Boreside-Winkel notwendig. Dies wurde anhand eines Projektegebietes in Khartum gelöst.

Mit den aktuellen 100 MPix PhaseOne iXU Kameras könnten so Sensorköpfe mit 30.000 Pixeln quer zur Flugrichtung realisiert werden. Schon für unter 300 k€ kann so ein komplettes System mit Gyrostabilisierung, GNSS-INS und Flugführung bereitgestellt werden, eine preiswerte und flexible Alternative für Entwicklungsländer. Durch Auswahl von Objektiven von 32-240 mm sind auch höchste Auflösungen von geeigneten Flughöhen möglich.

¹ GGS-GmbH, Kämmererstr.14, D-67346 Speyer, E-Mail: kemper@ggs-speyer.de

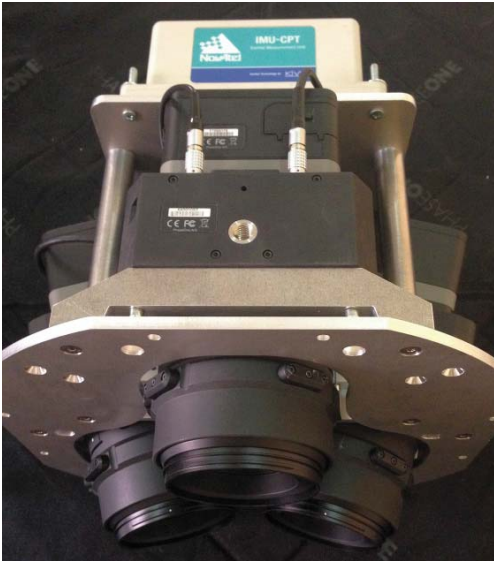


Abb. 1: Der Triple Sensorkopf aus 3*80 MPix iXA180 PhaseOne Kameras und GNSS-INS

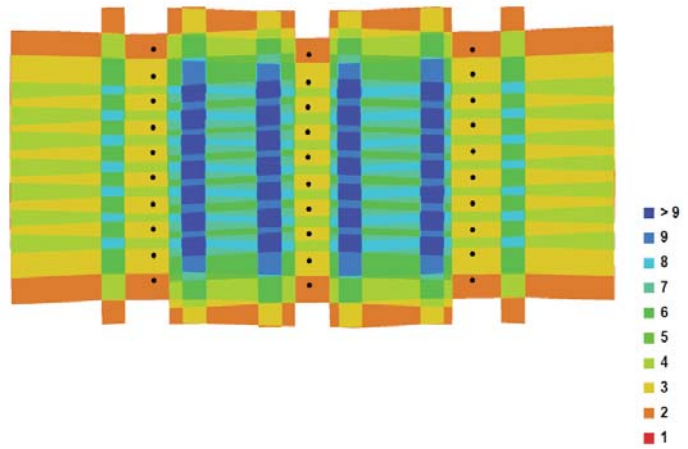


Abb. 2: Überlappungsanalyse des Kalibrierfelds in Nord-Khartum

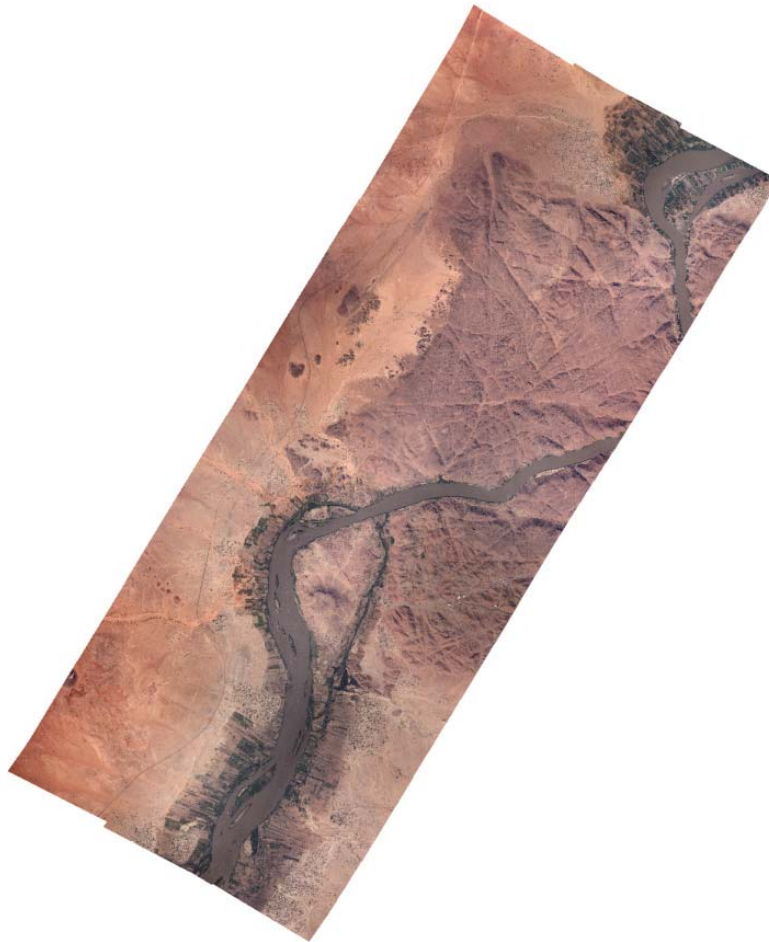


Abb. 3: Orthophoto-Mosaik des Wad al Basal Gebietes an Nil

2 DAISI - ein System für die Bestimmung von Vogelarten

Für das zu GICON Firmengruppe gehörende Institut für Angewandte Ökosystemforschung (IfAÖ) wurde ein Sensorsystem zur Erfassung von Seevögeln und Meeressäugern mit einer Auflösung von 2 cm GSD entwickelt. Die IfAÖ stellte am 6. März 2014 DAISI (Digital Aerial Imagery System by IfAÖ) vor (COPPACK et al. 2015). Mithilfe der in einem Flugzeug installierten hochauflösenden Doppelkamera kann DAISI Vögel und Meeressäuger an der Meeresoberfläche automatisiert erfassen. Dieses Verfahren muss seit Anfang des Jahres 2014 als Teil von Umweltverträglichkeitsprüfungen beim Bau von Offshore-Windparks digital und nicht länger manuell durchgeführt werden. Seit 2010 arbeitet das IfAÖ intensiv an der Entwicklung einer entsprechenden Methode. Was bisher von bis zu drei Beobachtern an Bord übernommen wurde, erledigt nun DAISI. Dafür wurde die 2-motorige Partenavia P68c mit zwei gekoppelten Phase One iXA180 Kameras mit einer Grundauflösung von 2cm und einer Abdeckung von 400m x 175m bei einer Flughöhe von 426m ausgestattet. Die Kameras sind mit einer minimalen Überlappung angeordnet um eine möglichst große Fläche abzudecken. Auf einer AeroStab-Twin sind die Kameras stabilisiert und werden durch das FMS getriggert. Der Import des Bildmaterials erfolgt teilautomatisiert mit Georeferenzierung über GPS-IMU. Durch eine spezielle Software werden Vogel- und Säugetierbeobachtungen kategorisiert und gezählt. Um Veränderungen zu dokumentieren, werden die Flächen in der Regel sieben Jahre lang – vor der Genehmigung, in der Bauphase und während der Betriebsphase – jeweils zehnmals jährlich überflogen.

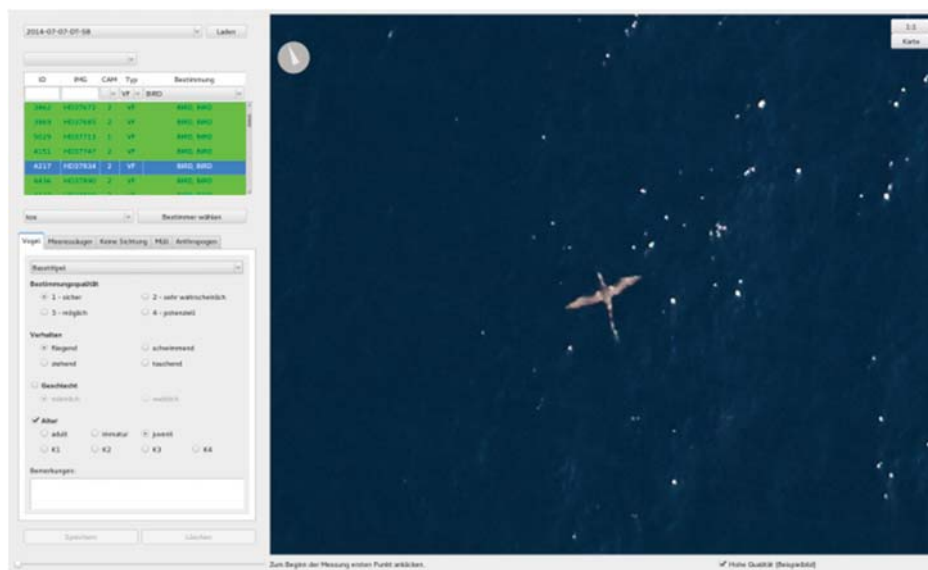


Abb. 4: Screen-shot des Auswerteprogramms zur automatisierten Vor-Bestimmung von Vogelarten (Beispiel zeigt einen jungen Töpel)

Bereits erfolgreich getestet wurde im letzten Jahr eine 3-Kameralösung (2*RGB und 1*NIR), die zur Glare-Kompensation (Reflektion in an der Wasseroberfläche insbesondere Wellen) dient. Glare-Effekte stören bei der automatischen Erkennung der Arten, deren Geschlecht und deren Anzahl. Durch das NIR-Band können Reflektionen in den RGB Bildern zugeordnet werden und so werden in den automatischen Vorselektionsschritten Interpretationsfehler reduziert. Dieses Jahr soll ein solches System basierend auf der iXU-1000 (100 MPix) installiert werden.

3 OIS- Oblique Imaging System

Oblique Kamerasysteme für die 3D Modellierung historischer Städte oder Objekte kann heute dank mittelformatiger metrischer 100-MPix Sensoren eine GSD von < 2 cm erreichen. Je nach Art der Integration können höchstgenaue Modelle gerechnet werden. Neben preisgünstigen Systemen (OIS-S) mit 5×36 MPix Nikon-Kameras oder bzw. mit der Canon 5DSR (50 MPix) welche alle auf der SLR Technologie basieren sind Systeme mit bis zu sieben 100 Mpix PhaseOne Kameras (OIS XL) möglich.

Um hohe geometrische Genauigkeiten zu erzielen sind nur Kameras mit Zentralverschluss geeignet (PIVNICKA et al. 2012). Es ist auch möglich SLR-Kameras mit Schlitzverschluss (Focal Plane Shutter) zu kalibrieren, jedoch ist die Entzerrung der Bilder zur Rückführung der durch den Verschluss verursachten Pixelverschiebungen gerade bei hohen Auflösungen komplex und gerade bei verschiedenen Einbaulagen obliquer Systeme eigentlich nicht zu realisieren.

Im Sommer wurde ein Kamerasystem mit einer iXU-180 als Nadir-Sensor und 4 iXU 150 als Oblique-Sensoren aufgebaut und kalibriert.

Die Kalibrierung beinhaltet neben der Kalibrierung der Kameras auch die Kalibrierung der Relativwinkel und Verschiebungen sowie die Boreside-Winkel. Im Sommer 2015 wurde die Befliegung mit einer Cessna 206 über Speyer durchgeführt. Es wurden 2 Flughöhen gewählt und bei großer Überlappung zudem Querstreifen in den Bildblock geplant. In Speyer besteht ein dichtes Netz von Passpunkten sowie eine GNSS Referenzstation.

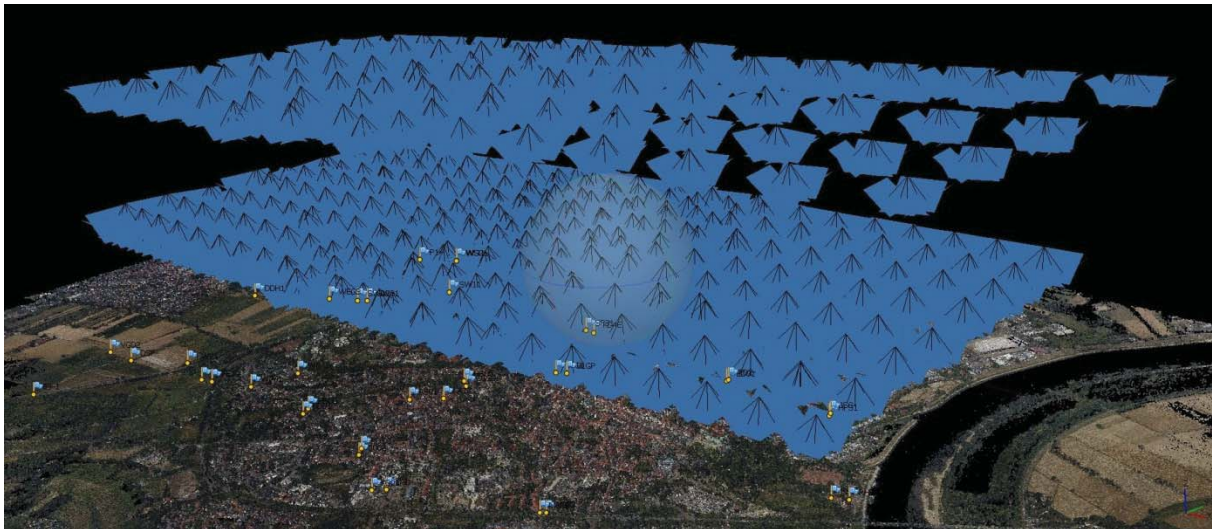


Abb. 5: Aerotriangulation aller Bilder und Flughöhen in PhotoScan

Bereits der erste Versuch einer Aerotriangulation aller Bilddaten war erfolgreich. Durch die große Überlappung wurden sehr viele Verknüpfungspunkte gefunden und eine robuste erste Lösung erreicht. Die Daten wurden dann in Bingo (KRUCK & MELYKUTI 2014). weiterverarbeitet. Hier wurden zunächst Fehlmessungen automatisch korrigiert und mit den Passpunkten eine präzise AT erzielt. Daraus wurden die einzelnen Kameras zunächst mit klassischen verfahren, dann zur Eliminierung der Restfehler mit zusätzlichen Parametern kalibriert.

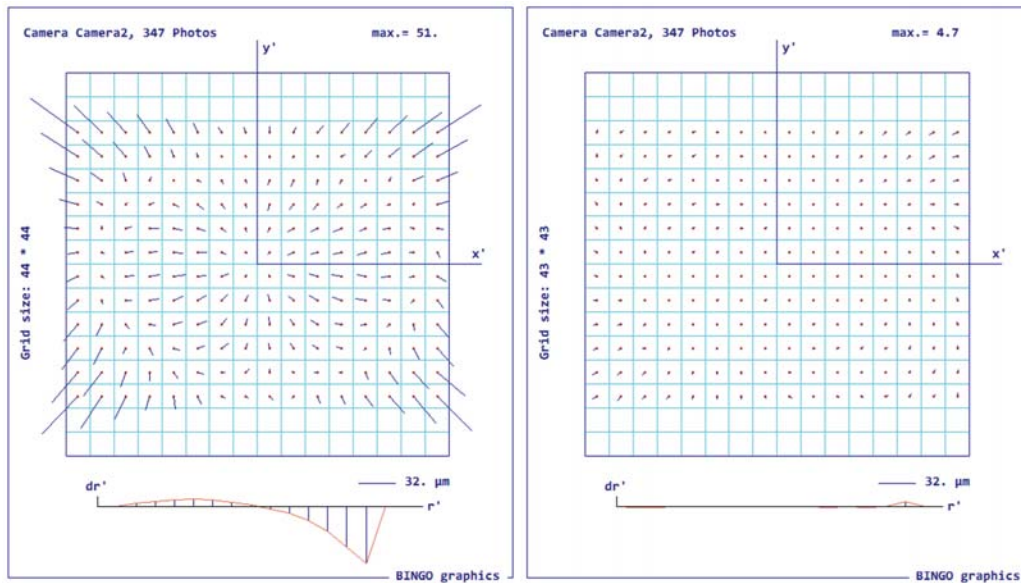


Abb. 6: Verbleibende radiale Verzeichnung vor und nach der Ausgleichung in Bingo

Nach der Kalibration konnten nun die relativen Winkel der obliquen Kameras zur Nadir-Kamera berechnet werden. Die Berechnung der Boresight-Winkel kann grundsätzlich für alle Kameras erfolgen oder zur Nadir-Kamera und von dort Co-Referenziert zu den obliquen Kameras. Beide Wege führen zu vergleichbaren Ergebnissen (KEMPER et al. 2016). Bei der Übernahme der GNSS-INS Daten ist noch die Lage der Nodalpunkte von Bedeutung, diese wird von dem Hersteller der Kamera/Objektive bereitgestellt. Diese ist mit den spezifischen Einbaulagen im Sensorkopf zu verrechnen.



Abb. 7: Ausschnitt des 3D Stadtmodells vom Testgebiet mit dem Dom zu Speyer

4 Multispektrale 4/5 Band-Systeme

Multispektrale Setups mit der Kombination von hochauflösenden 4 Kanal Kameras (PhaseOne 4-Band Solution) in Kombination mit thermalen Sensoren eröffnen neue Möglichkeiten auf der Suche nach Objekten im Untergrund. Die Spuren der menschlichen Kultur sind oft unter der Erde verborgen. Luftbilder hatten schon immer mit einer neuen Perspektive geholfen Strukturen zu erkennen die so im Gelände nicht sichtbar waren. Sei es durch leichte Deformation der Oberfläche, durch andere Bodenstruktur mit der Wirkung auf die Vegetation oder veränderte hydrologische Bedingungen.

Zusätzlich zu Farbbildern welche darstellen was auch das menschliche Auge wahrnimmt können CIR Daten (TÖLG et al. 2016) noch besser Hinweise auf den Zustand der Vegetation geben und somit Rückschlüsse auf Anomalien im Untergrund. Der Chlorophyllgehalt der Pflanzen ist ein Indikator deren Vitalität. Gerade in der Phase der Zustandsänderung ist dies im sichtbaren Band oft noch nicht erkennbar. Besser als mit reinen RGBI Daten können Feuchteunterschiede durch veränderte Evapotranspiration und der damit einhergehenden Temperaturänderung (VASEL et al. 2008) erfasst werden. Thermalsensoren erfassen die thermische Emission und zeigen bei feuchteren Stellen eine niedrigere Temperatur am Tag und höhere Temperaturwerte in der Nacht oder am frühen Morgen.

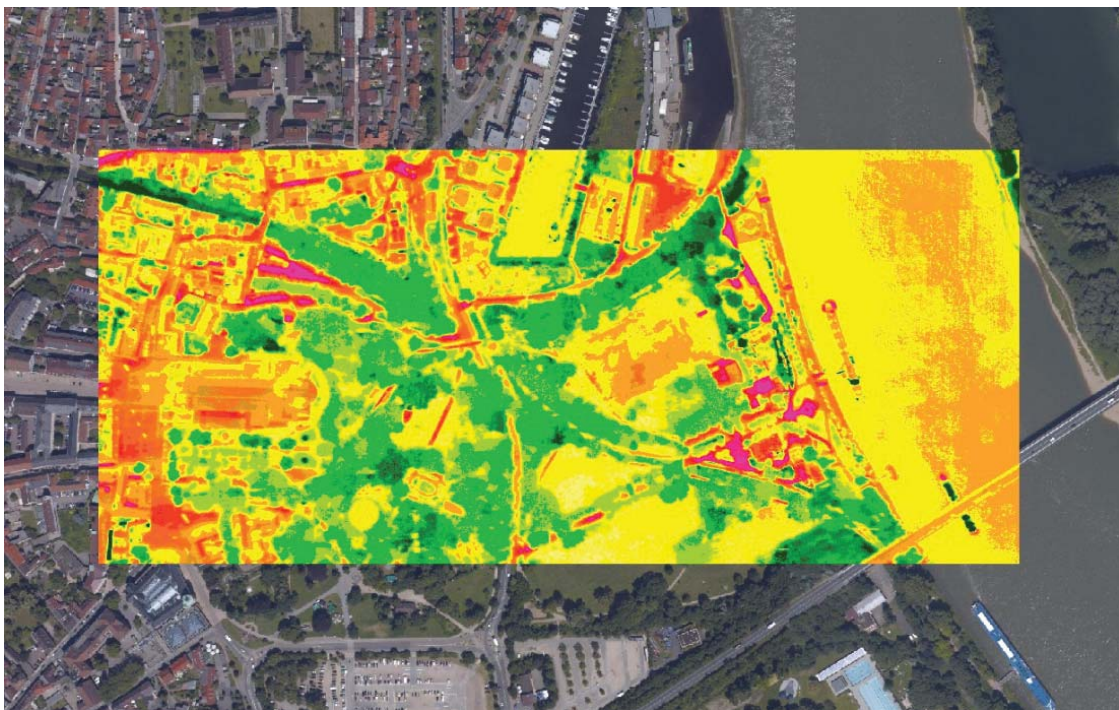


Abb. 8: Orthophoto von Speyer überlagert mit einem thermalen Orthophoto.

Thermaldaten helfen aber auch in Objekten Strukturen zu erkennen, welche sich thermisch durchzeichnen (KEMPER et al. 2013). Das können unterschiedliche Arten der Gesteine sein oder Gradienten des Feuchtetransportes.

5 OTS - Oblique Thermal Systeme

Thermalkameras haben noch immer das Problem der geringen Auflösung, Standard ist noch immer 640*512 Pixel. Damit sind höhere Auflösungen schwierig zu realisieren da sich der Flugaufwand potenziert. Typische GSD für thermale Anwendungen liegen bei 25-100 cm.

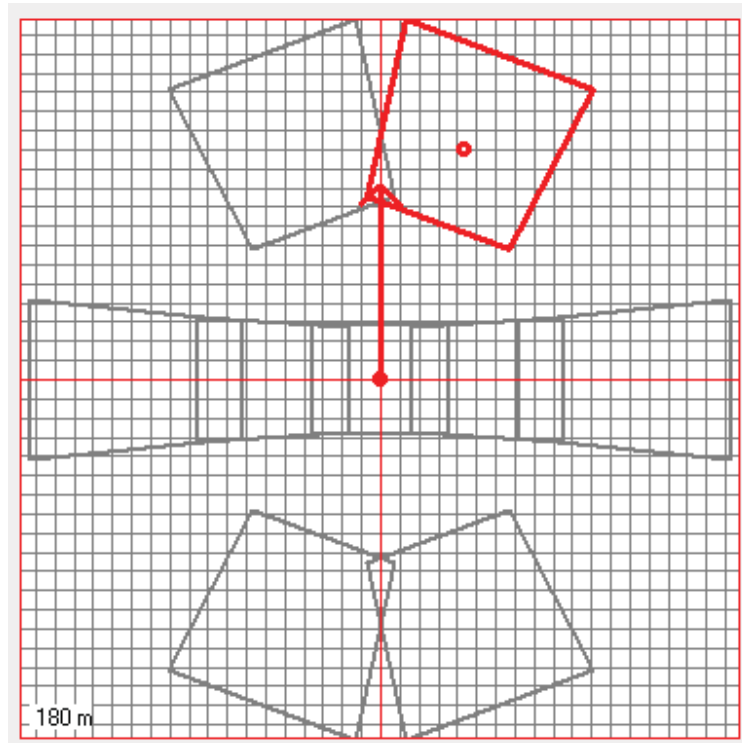


Abb. 9: Setup des OTS mit seinen 9 Thermalsensoren.

Durch die Kombination mehrerer Thermalsensoren kann die Erfassungsbreite vergrößert und sogar Schrägansichten in das Thermale Datenmanagement integriert werden (KEMPER 2016). Das Setup des OTS beinhaltet 9 Sensorköpfe. Dadurch werden 3200 Pixel quer erreicht, wobei die äußeren 6 Kameras mit 36 Grad schon klar oblique sind. Dadurch werden auch Fassaden und senkrechte Objekte direkt vermessen und nicht nur der konvektiv indirekt gemessene Aufstieg der erwärmten Luft.

Bei der Thermographie ist es besonders wichtig, das Gebiet zügig zu erfassen, da sich über die Zeit die Temperaturverhältnisse ändern. Was bei der klassischen Luftbildbefliegung ein Wandern der Schatten ist, stellt bei der Thermographie eine weitaus größere Problematik dar. Durch das OTS wird das Gebiet nicht nur schneller abgedeckt, sondern durch die unzähligen Überlappungen auch Möglichkeiten zur Korrektur der Temperaturdrift eröffnet. Dies führt nicht nur zu homogeneren Daten, sondern auch zu einer besseren Interpretierbarkeit.

6 Ausblick

Die Möglichkeiten heutiger Sensortechnik erlauben die optimale anwenderspezifische Adaption, sei es für höchste Auflösung z.B. bei Architektur und Denkmalschutz, oder großen Flächen z.B. zur

Erfassung von Pyramiden im Sudan. Die Sensoren werden immer leistungsstärker und kompakter was auch den Einsatz in UAVs ermöglicht. Allerdings sind nicht alle Sensoren einfach zu integrieren denn bei Multisensor-Systemen spielt die Synchronität eine wichtige Rolle für die automatisierte Auswerten und auch im Zusammenspiel mit GNSS-INS. Neben der stabilen geometrischen Integration ist also die Synchronisierung eine ernste Aufgabe. Während bei den PhaseOne Kameras schon von Beginn eine "Daisy Chain" Lösung angeboten wurde sind bei einfachen SLR Kameras solche Lösungen selbst zu erarbeiten.

Für Applikationen des hochgenauen Monitorings von Powerlines sind Lösungen mit Lidar, RGB, Thermal und UV in Entwicklung, also fast alles was es an Sensoren heute gibt. Eine solche Integration ist auch finanziell heute möglich.

7 Literatur

- COPPACK, T., WEIDAUER, A. & KEMPER, G., 2015: Erfassung von Seevogel- und Meeressäugerbeständen mittels georeferenzierter Digitalfotografie. AGIT, Journal für Angewandte Geoinformatik **1**, 358-367.
- KEMPER, G., & VASEL, R., 2009: Luftgestützte Thermographie zur Überwachung von Bauwerksschäden. Innovative Feuchtemessung in Forschung und Praxis 4, Berichtsband zum Workshop 2009, Karlsruhe.
- KEMPER, G., 2010: Neue luftgestützte Sensoren und Plattformen für verschiedenste Aufgaben in der Fernerkundung. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 19, G. KOHLHOFER & M. FRANZEN (Hrsg.), 650-657.
- KEMPER, G., 2012: New Airborne Sensors and Platforms for Solving Specific Tasks in Remote Sensing. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **39**(B5), 351-356.
- KEMPER, G., SCHUHMANN, R. & KÖNIGER, F., 2013: New airborne thermal infrared photogrammetric applications and sensors for moisture detection. Proceedings of the 10th International Conference on Electromagnetic Wave Interaction with Water and Moist Substances (ISEMA), KUPFER, K. (Hrsg.), Weimar, 25.-27. September 2013, 423-431.
- KEMPER, G., MELYKUTI, B. & YU, C., 2016: Calibration procedures on oblique camera setups. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences **41**(B1), 205-209.
- KEMPER, G. 2016: Multisensors and platforms for various airborne applications. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **41**(B1), 211-215.
- KEMPER, G., WEIDAUER, A. & COPPACK, T., 2016: Monitoring seabirds and marine mammals by georeferenced aerial photography. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **41**(B8), 689-694.
- KRUCK, E. & MELYKUTI, B., 2014: Kalibrierung von Oblique- und UAV Kameras. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 23, Potsdam, SEYFERT, E., GÜLCH, E., HEIPKE, C., SCHIEWE, J., SESTER, M. (Hrsg.), Beitrag 120.

- PIVNICKA, F., KEMPER, G. & GEISLER, S., 2012: Calibration Procedures in Mid Format Camera Setups. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39**(B1), 149-152.
- TÖLG, T., KEMPER, G., KALINSKI, D., 2016: Medium-format camera evaluation based on the latest PhaseOne technology. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41**(B1), 121-126
- VASEL, R., KEMPER, G., SCHUHMAN, R. & KÖNIGER, F., 2008: New airborne thermal infrared photogrammetric applications and sensors for moisture detection. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **37**(B8), 17-22.

Involving Different Neighborhood Types for the Analysis of Low-level Geometric 2D and 3D Features and their Relevance for Point Cloud Classification

MARTIN WEINMANN¹, CLÉMENT MALLET² & BORIS JUTZI¹

Abstract: In this paper, we address the automatic interpretation of 3D point cloud data in terms of associating a (semantic) class label to each 3D point. In particular, we aim at analyzing the behavior of standard handcrafted low-level geometric 2D and 3D features for different neighborhood types. For this purpose, we present a framework that considers four neighborhood definitions as the basis for calculating a set of 18 low-level geometric 2D and 3D features which, in turn, are provided as input to three classifiers relying on different learning principles. We demonstrate the performance of our framework on a benchmark dataset for which a labeling with respect to three structural classes (linear, planar and volumetric structures) as well as a labeling with respect to five semantic classes (wire, pole/trunk, façade, ground and vegetation) is available. The derived results clearly reveal that the suitability of the considered neighborhood types and thus the relevance of respectively extracted features with respect to the classification task varies significantly.

1 Introduction

The automatic analysis of 3D point cloud data is a topic of major interest in photogrammetry, remote sensing, computer vision and robotics. Among a variety of tasks, attention has been paid in particular to 3D point cloud classification, where the objective consists in automatically labeling the given 3D points with respect to pre-defined classes. Thereby, the main challenges are represented by the irregular point sampling, very different types of objects, the typically high complexity of observed scenes and the high computational burden arising from the consideration of large 3D point clouds in combination with a rich diversity of available features.

Many investigations on 3D point cloud classification rely on the use of geometric features and, consequently, a respective processing pipeline typically involves (1) the recovery of a local neighborhood for each point of the 3D point cloud, (2) the extraction of geometric features via a consideration of the spatial arrangement of all points inside the local neighborhood and (3) a supervised classification via a classifier trained on representative training data. Accordingly, the quality of derived classification results relies on the selected neighborhood definition, the extracted features and the involved classifier.

In this paper, we focus on analyzing the behavior of standard handcrafted low-level geometric 2D and 3D features for different neighborhood types. For this purpose, we consider four neighborhood definitions as the basis for feature extraction: (1) a cylindrical neighborhood with a fixed radius of 1m, (2) a spherical neighborhood with a fixed radius of 1m, (3) a spherical

¹ Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Englerstraße 7, 76131 Karlsruhe, Germany, E-Mail: [martin.weinmann, boris.jutzi]@kit.edu

² Université Paris-Est, LASTIG MATIS, IGN, ENSG, 73 avenue de Paris, 94160 Saint-Mandé, France, E-Mail: clement.mallet@ign.fr

neighborhood formed by the $k = 50$ nearest neighbors, and (4) a spherical neighborhood formed by the k_{opt} nearest neighbors whereby the optimal scale parameter k_{opt} is selected for each 3D point individually via *eigenentropy-based scale selection* (WEINMANN et al. 2015a). Based on these neighborhoods, we calculate a set of 18 low-level geometric 2D and 3D features and, subsequently, we evaluate their relevance for the classification task. For this purpose, we present classification results achieved for three classifiers relying on different learning principles (instance-based learning, probabilistic learning and ensemble learning) and for different classification tasks. By involving a benchmark dataset for which a labeling with respect to three structural classes (*linear structures*, *planar structures* and *volumetric structures*) as well as a labeling with respect to five semantic classes (*wire*, *pole/trunk*, *façade*, *ground* and *vegetation*) is available as shown in Fig. 1, we demonstrate that the relevance of extracted features with respect to a classification task may significantly vary depending on the selected neighborhood type.

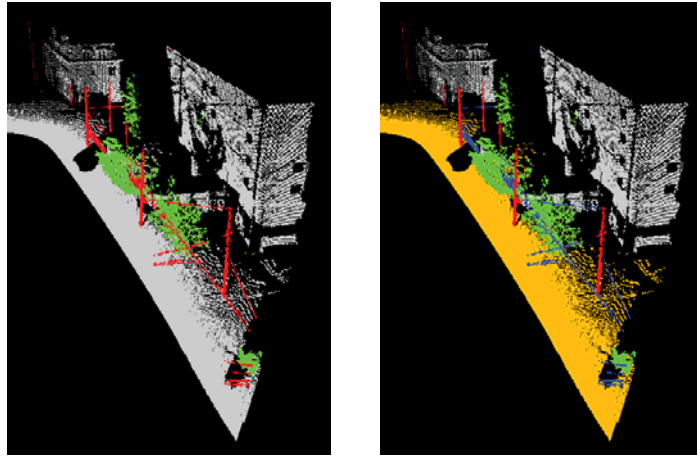


Fig. 1: Benchmark dataset with a reference labeling with respect to three structural classes (left), which indicate *linear structures* (red), *planar structures* (gray) and *volumetric structures* (green), and with a reference labeling with respect to five semantic classes (right), which are represented by *wire* (blue), *pole/trunk* (red), *façade* (gray), *ground* (brown) and *vegetation* (green).

This paper represents an extension of our previous work (WEINMANN et al. 2015a), where we presented a framework for semantic 3D point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. From this framework, we make use of (1) the eigenentropy-based scale selection in order to recover the optimal scale parameter k_{opt} for each 3D point individually and (2) the extraction of a set of 18 low-level geometric 2D and 3D features. In addition, we now consider further neighborhood types as a basis for extracting the same set of handcrafted features, and we evaluate their relevance with respect to two different classification tasks. Thus, our contribution consists of an extended framework which involves

- the recovery of local neighborhoods based on conceptually different neighborhood types,
- the extraction of low-level geometric 2D and 3D features on the basis of different neighborhood types, and
- a classification via standard classifiers relying on different learning principles.

After briefly summarizing related work (Section 2), we proceed with describing the proposed framework in detail (Section 3). Subsequently, we present the results derived for a benchmark dataset (Section 4), whereby we focus on both the behavior of geometric features for different neighborhood types and the classification results derived for two different classification tasks. Finally, we provide concluding remarks and suggestions for future work (Section 5).

2 Related Work

The semantic classification of 3D point cloud data typically relies on the use of geometric features. The extraction of such geometric features often exploits characteristics of the local 3D structure, i.e. local neighborhoods have to be recovered (Section 2.1). On the basis of the 3D points within the recovered local neighborhoods, a variety of geometric features can be calculated (Section 2.2). These features are then provided as input to the classifier that is trained on representative training data (Section 2.3).

2.1 Neighborhood Recovery

In general, there are very different strategies for recovering local neighborhoods for the points of a 3D point cloud. Among the most commonly used neighborhood types, one can find spherical neighborhoods parameterized by a radius (LEE & SCHENK 2002), cylindrical neighborhoods parameterized by a radius (FILIN & PFEIFER 2005), spherical neighborhoods parameterized by the number of nearest neighbors with respect to the Euclidean distance in 3D space (LINSEN & PRAUTZSCH 2001) and cylindrical neighborhoods parameterized by the number of nearest neighbors with respect to the Euclidean distance in 2D space (NIEMEYER et al. 2014). However, all these neighborhood types rely on a scale parameter which is represented by either a radius or the number of nearest neighbors. To define this scale parameter, prior knowledge about the scene and/or the data is typically involved. Furthermore, the same value of the scale parameter is commonly selected for all points of the 3D point cloud. To allow for more variability and thus account for the fact that structures of different classes might favor a different neighborhood size, data-driven approaches for optimal neighborhood size selection have been presented (MITRA & NGUYEN 2003; LALONDE et al. 2005; DEMANTKÉ et al. 2011; WEINMANN et al. 2015a).

Instead of describing the local 3D structure at a specific scale, it is possible to describe the local 3D structure at different scales and thus also how the local 3D geometry behaves across scales. Respective approaches for instance focus on the combination of multiple spherical neighborhoods with different radii (BRODU & LAGUE 2012) or the combination of multiple cylindrical neighborhoods with different radii (NIEMEYER et al. 2014; SCHMIDT et al. 2014). The combination of different entities such as voxels, blocks and pillars results in multi-type neighborhoods (HU et al. 2013) and the use of multi-scale, multi-type neighborhoods in the form of a combination of both spherical and cylindrical neighborhoods with different scale parameter has been proposed in (BLOMLEY et al. 2016).

2.2 Feature Extraction

Based on the recovered local neighborhoods, a variety of geometric features can be calculated. Among these, the eigenvalues of the 3D structure tensor (JUTZI & GROSS 2009) are often used to

derive the *local 3D shape features* presented in (WEST et al. 2004; PAULY et al. 2003) which are widely used for classifying 3D point cloud data. To address further characteristics of the given 3D point cloud data, it has been proposed to additionally extract features based on angular characteristics (MUNOZ et al. 2009), height and plane characteristics (MALLET et al. 2011), a variety of low-level geometric 3D and 2D features (WEINMANN et al. 2015a), moments and height features (HACKEL et al. 2016), or specific descriptors addressing surface properties, slope, height characteristics, vertical profiles and 2D projections (GUO et al. 2015).

2.3 Classification

The calculated features are provided as input to the classifier which may rely on different strategies. On the one hand, it is possible to involve standard classifiers such as a Random Forest classifier (CHEHATA et al. 2009) or a Support Vector Machine classifier (MALLET et al. 2011). A comprehensive comparison of a variety of standard classifiers relying on different learning principles is presented in (WEINMANN et al. 2015a) and reveals that a Random Forest classifier provides a good trade-off between classification accuracy and computational efficiency. On the other hand, it is possible to take into account that the derived labeling should be spatially regular which can be enforced via statistical models of context such as Associative Markov Networks (MUNOZ et al. 2009), non-Associative Markov Networks (SHAPOVALOV et al. 2010), or Conditional Random Fields (NIEMEYER et al. 2014; WEINMANN et al. 2015b).

3 Methodology

In the scope of this paper, we focus on a framework for point cloud classification, which allows statements about the suitability of involved neighborhood types and the relevance of respectively extracted features with respect to the considered classification task. This framework is illustrated in Fig. 2 and consists of three components addressing neighborhood recovery (Section 3.1), feature extraction (Section 3.2) and classification (Section 3.3).

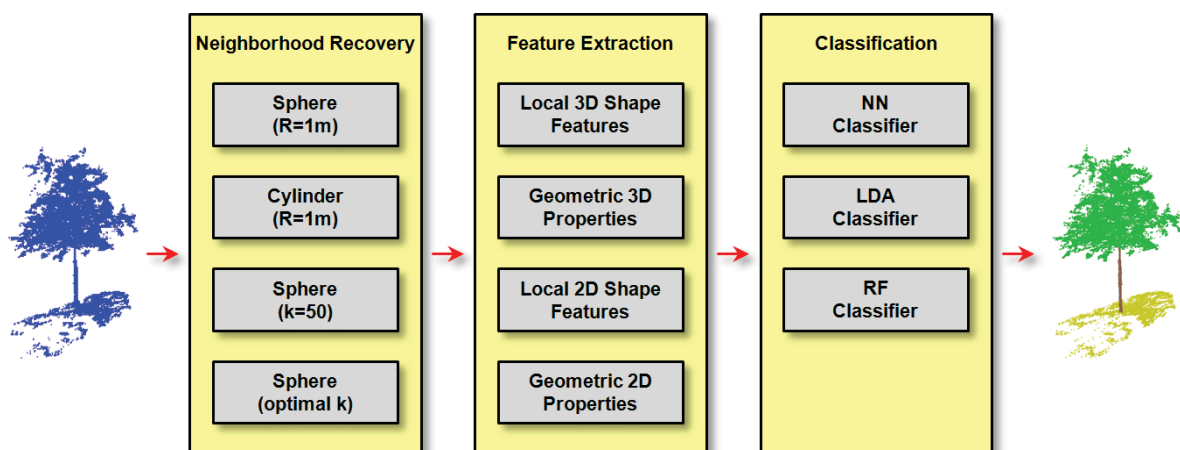


Fig. 2: The proposed framework for point cloud classification with four options for neighborhood recovery, four types of low-level geometric features and three options for supervised classification (NN: Nearest Neighbor; LDA: Linear Discriminant Analysis; RF: Random Forest).

3.1 Neighborhood Recovery

As geometric features describing a point \mathbf{X}_0 are typically derived by considering the local surrounding of \mathbf{X}_0 , a crucial prerequisite for calculating geometric features consists in the choice of an appropriate neighborhood definition. We therefore include four different neighborhood definitions in our framework:

- a spherical neighborhood $N_{s,1m}$, where a sphere is centered at \mathbf{X}_0 and has a radius of 1m,
- a cylindrical neighborhood $N_{c,1m}$, where a cylinder is centered at \mathbf{X}_0 , has a radius of 1m and is oriented along the vertical direction,
- a spherical neighborhood $N_{k=50}$ comprising the $k = 50$ nearest neighbors of \mathbf{X}_0 with respect to the Euclidean distance in 3D space, and
- a spherical neighborhood $N_{k,opt}$ comprising the optimal number k_{opt} of nearest neighbors of \mathbf{X}_0 with respect to the Euclidean distance in 3D space, whereby k_{opt} is selected for each point individually via *eigenentropy-based scale selection* (WEINMANN et al. 2015a).

3.2 Feature Extraction

Once a local neighborhood has been derived for each point \mathbf{X}_0 , the spatial arrangement of points within that local neighborhood can be described via geometric features. Here, we focus on the use of low-level geometric features that have been presented in (WEINMANN et al. 2015a). These features comprise the *local 3D shape features* of linearity, planarity, sphericity, omnivariance, anisotropy, eigenentropy, sum of eigenvalues and change of curvature (WEST et al. 2004; PAULY et al. 2003) that rely on the eigenvalues of the 3D structure tensor. In addition, features are defined based on *geometric 3D properties* that are represented by the height of \mathbf{X}_0 , the distance between \mathbf{X}_0 and the farthest point in the local neighborhood, 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). As the presence of different man-made objects can be assumed in urban scenes, some further features rely on a 2D projection of all points within the local neighborhood onto a horizontally oriented plane. Using the 2D projections, it is possible to define *local 2D shape features* in terms of the sum and the ratio of the eigenvalues of the 2D structure tensor (WEINMANN et al. 2015a), and *geometric 2D properties* are given by the distance between the projection of \mathbf{X}_0 and the farthest point in the local 2D neighborhood and the local point density in 2D space.

3.3 Supervised Classification

For classification, we integrate three classifiers relying on different learning principles into our framework. This allows us to draw more general conclusions about the suitability of different neighborhood types and thus the relevance of respectively extracted features with respect to the classification task:

(1) *Nearest Neighbor (NN) classifier:*

This classifier relies on the principle of instance-based learning, where the induction is delayed to the prediction stage. Accordingly, there is no training stage. In the prediction stage, each feature vector belonging to a sample in the test set is compared to the feature

vectors corresponding to all samples in the training data and the class label associated with the most similar feature vector in the training data is selected. Due to the typically large number of required comparisons of feature vectors (here with respect to the Euclidean distance), the prediction stage is likely to cause a higher computational burden.

(2) *Linear Discriminant Analysis (LDA) classifier:*

This classifier relies on the principle of probabilistic learning. In the training stage, a multivariate Gaussian distribution is fitted to the training data, i.e. the parameters of a Gaussian distribution have to be estimated for each class, assuming that the instances of each class follow a Gaussian distribution in the feature space. Thereby, the LDA classifier involves the simplifying assumption that the same covariance matrix can be used for all classes so that only their means vary. In the prediction stage, it is evaluated with which probability each feature vector belonging to a sample in the test set is assigned to the different classes, and the class with the maximum probability is selected.

(3) *Random Forest (RF) classifier:*

This classifier relies on the principle of ensemble learning (BREIMAN 2001). In the training stage, the strategy of bagging (BREIMAN 1996) is applied to generate an ensemble of randomly trained decision trees. 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. In the prediction stage, for each feature vector belonging to a sample in the test set, each of the randomly trained decision tree casts a vote for one of the defined classes and, typically, the majority vote is selected to obtain an appropriate prediction of the respective class label.

4 Experimental Results

In the following, we first present the involved benchmark dataset (Section 4.1). Subsequently, we focus on the behavior of geometric features for different neighborhood types and present the derived classification results (Section 4.2). Finally, we discuss the derived results with respect to different aspects (Section 4.3).

4.1 Dataset

We evaluate the performance of our framework on the *Oakland 3D Point Cloud Dataset* (MUNOZ et al. 2009), a widely used benchmark dataset acquired with a mobile laser scanning system in the vicinity of the CMU campus in Oakland, USA. The dataset is split into a training set comprising about 36.9k points, a validation set comprising about 91.5k points and a test set comprising about 1.3M points. For each point, a reference labeling with respect to three structural classes (*linear structures*, *planar structures* and *volumetric structures*) is available as well as a reference labeling with respect to five semantic classes (*wire*, *pole/trunk*, *façade*, *ground* and *vegetation*) as shown in Fig. 1. To distinguish between both classification tasks, we refer to *Oakland-3C* and *Oakland-5C*, respectively.

4.2 Results

First, we focus on the structural behavior of the low-level geometric features for the different neighborhood types. Thereby, we exemplarily consider the behavior of the three dimensionality features of linearity, planarity and sphericity which is illustrated in Fig. 3. It can be observed that the different neighborhood types have a significantly different impact on feature extraction.

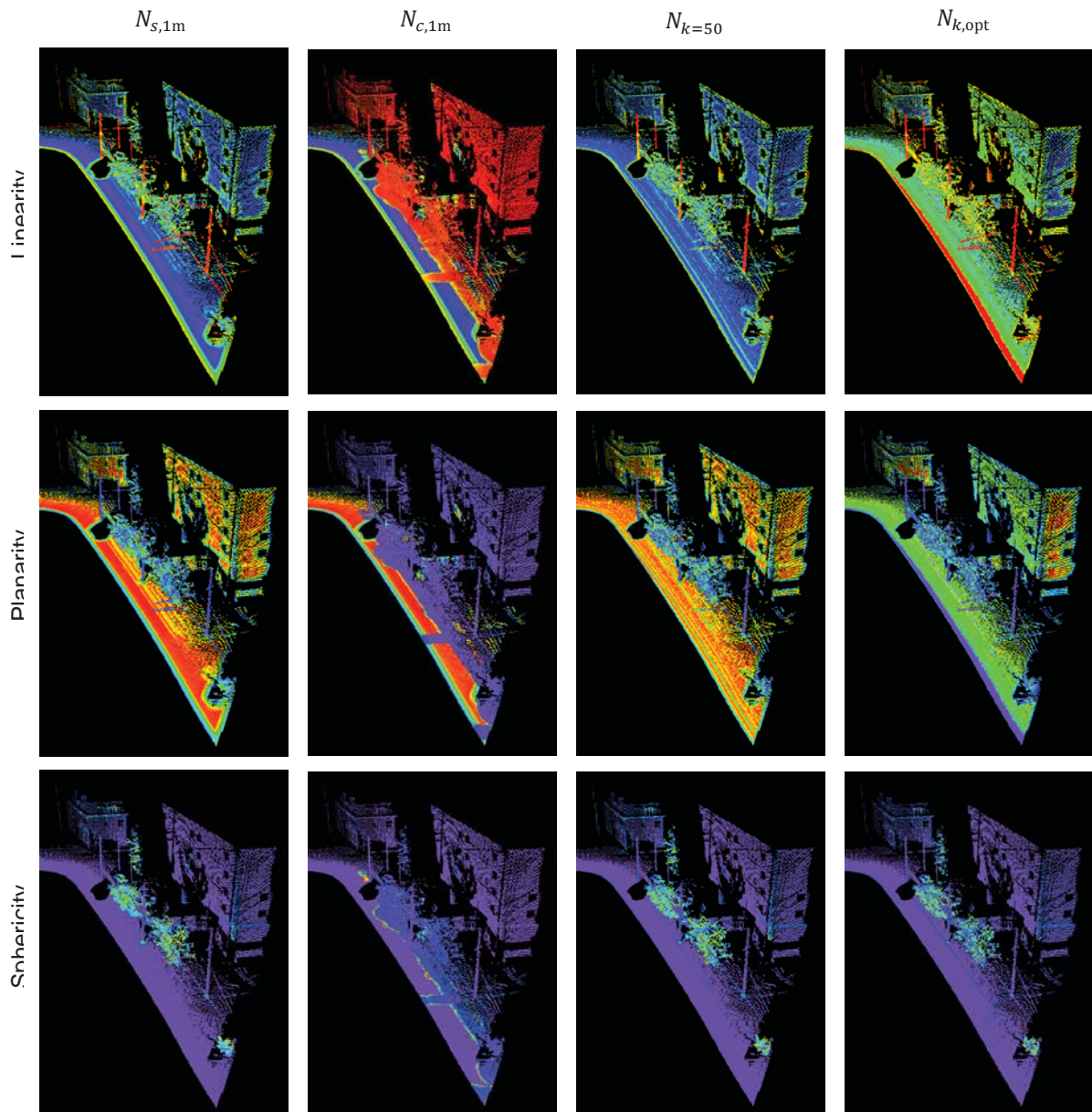


Fig. 3: Behavior of the dimensionality features of linearity, planarity and sphericity for the considered scene: the color encoding indicates high values close to 1 in red and reaches via yellow, green, cyan and blue to violet for low values close to 0.

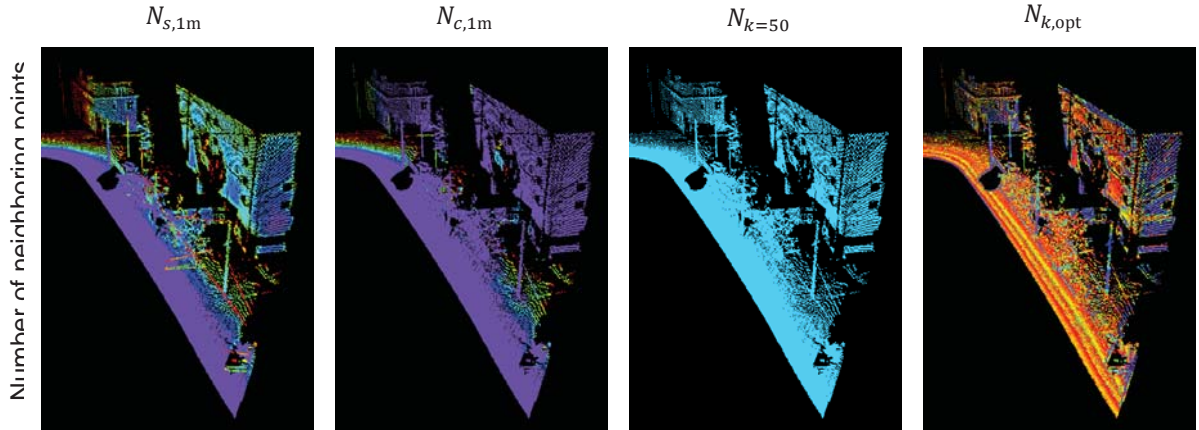


Fig. 4: Number of neighboring points for the different neighborhood types: the color encoding indicates neighborhoods with 10 or less points in red and reaches via yellow, green, cyan and blue to violet for 100 and more points.

Additionally, we consider the number of neighbors within the local neighborhood which is visualized in Fig. 4. Due to their significantly different impact on feature extraction, it can be expected that the different neighborhood types might also significantly differ in their suitability with respect to the classification task. To verify this, we use the different neighborhood types as the basis for extracting the 18 low-level geometric 2D and 3D features which, in turn, are provided as input to three classifiers relying on different learning principles. In the training stage, we take into account that an unbalanced distribution of training examples per class might have a detrimental effect on the classification results (CRIMINISI & SHOTTON 2013) and hence use a reduced training set comprising 1000 training examples per class. In the prediction stage, we estimate the class labels for the validation set. For *Oakland-3C*, the derived classification results are provided in Tab. 1 and visualized in Fig. 5. The quantitative results reveal an overall accuracy in the range of about 67-94% and a κ -value in the range of about 28-76%. For *Oakland-5C*, the derived classification results are provided in Tab. 2 and visualized in Fig. 6. Here, the quantitative results reveal an overall accuracy in the range of about 68-96%, whereas the κ -value is in the range of about 49-90%.

4.3 Discussion

The derived results clearly reveal that the spherical and cylindrical neighborhood types parameterized by a radius result in local neighborhoods that comprise a larger number of 3D points, while the neighborhoods recovered via eigenentropy-based scale selection tend to be comparably small (Fig. 4). Furthermore, it can be observed that the selected cylindrical neighborhood type is not well-suited for classifying mobile laser scanning data as extracted features exhibit a rather non-intuitive behavior with seam effects (Fig. 3) and do hence not lead to appropriate classification results (Tab. 1, Tab. 2, Fig. 5 and Fig. 6). In contrast, the use of the other neighborhood types results in adequate classification results for almost all cases, with one exception for the combination of using locally optimal neighborhoods and a RF classifier. For that case, misclassifications mainly occur in regions of low point density (Fig. 5 and Fig. 6).

Tab. 1: Classification results obtained for *Oakland-3C* when using four different neighborhood types.

Neighborhood Type	Overall Accuracy [%]			κ [%]		
	NN	LDA	RF	NN	LDA	RF
$N_{s,1m}$	90.80	92.78	91.74	66.09	72.30	69.47
$N_{c,1m}$	84.00	78.29	82.26	50.57	41.55	48.30
$N_{k=50}$	92.41	91.27	92.52	70.49	67.60	71.39
$N_{k,opt}$	67.86	90.91	93.83	28.33	67.47	75.24

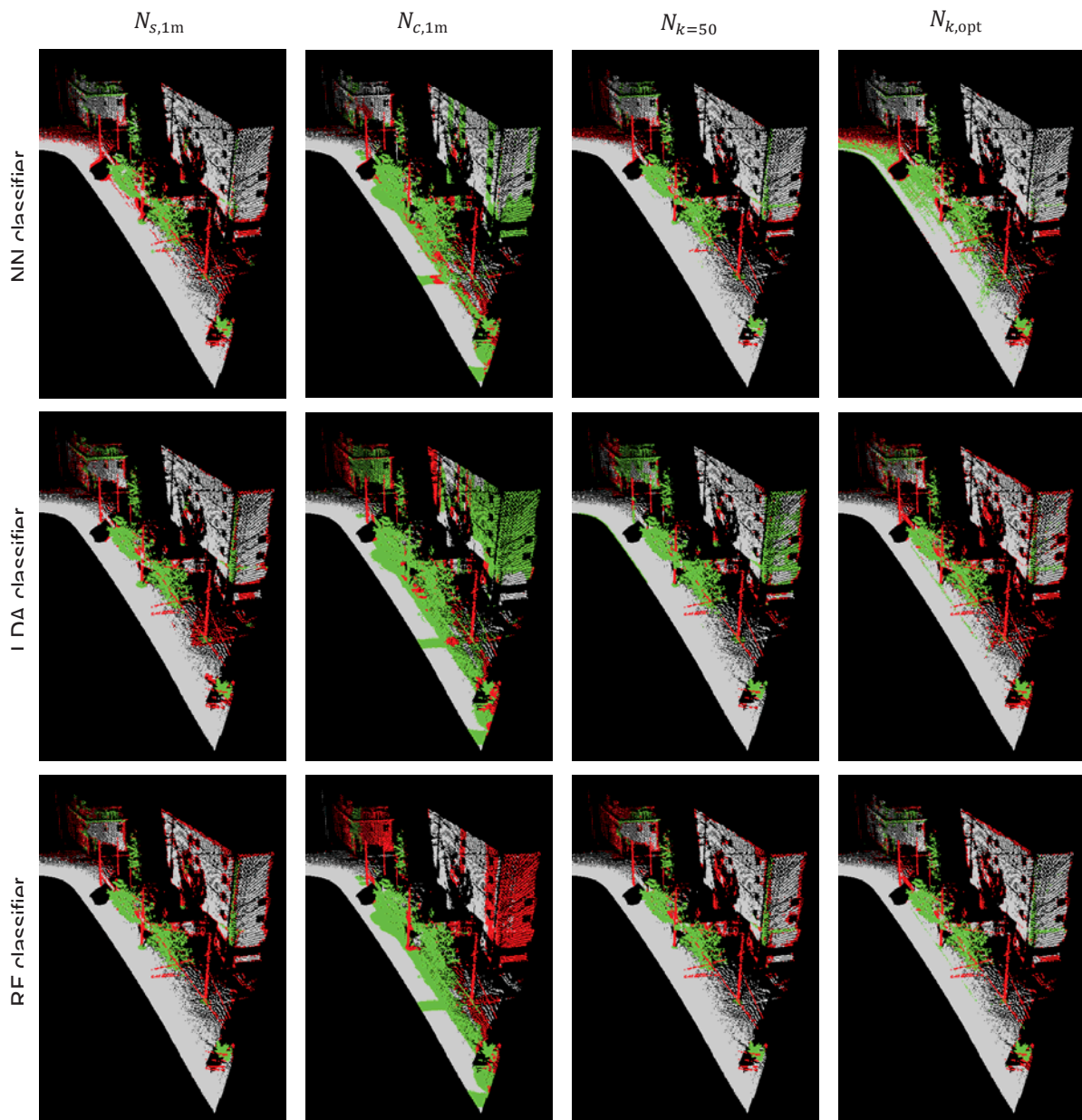


Fig. 5: The observed scene classified with respect to three structural classes which are represented by *linear structures* (red), *planar structures* (gray) and *volumetric structures* (green).

Tab. 2: Classification results obtained for *Oakland-5C* when using four different neighborhood types.

Neighborhood Type	Overall Accuracy [%]			κ [%]		
	NN	LDA	RF	NN	LDA	RF
$N_{s,1m}$	90.26	95.11	93.73	79.18	88.78	85.71
$N_{c,1m}$	85.22	86.73	85.89	70.01	72.68	71.18
$N_{k=50}$	91.09	95.28	95.36	80.61	89.06	89.26
$N_{k,opt}$	68.06	94.38	94.15	49.17	87.30	86.66

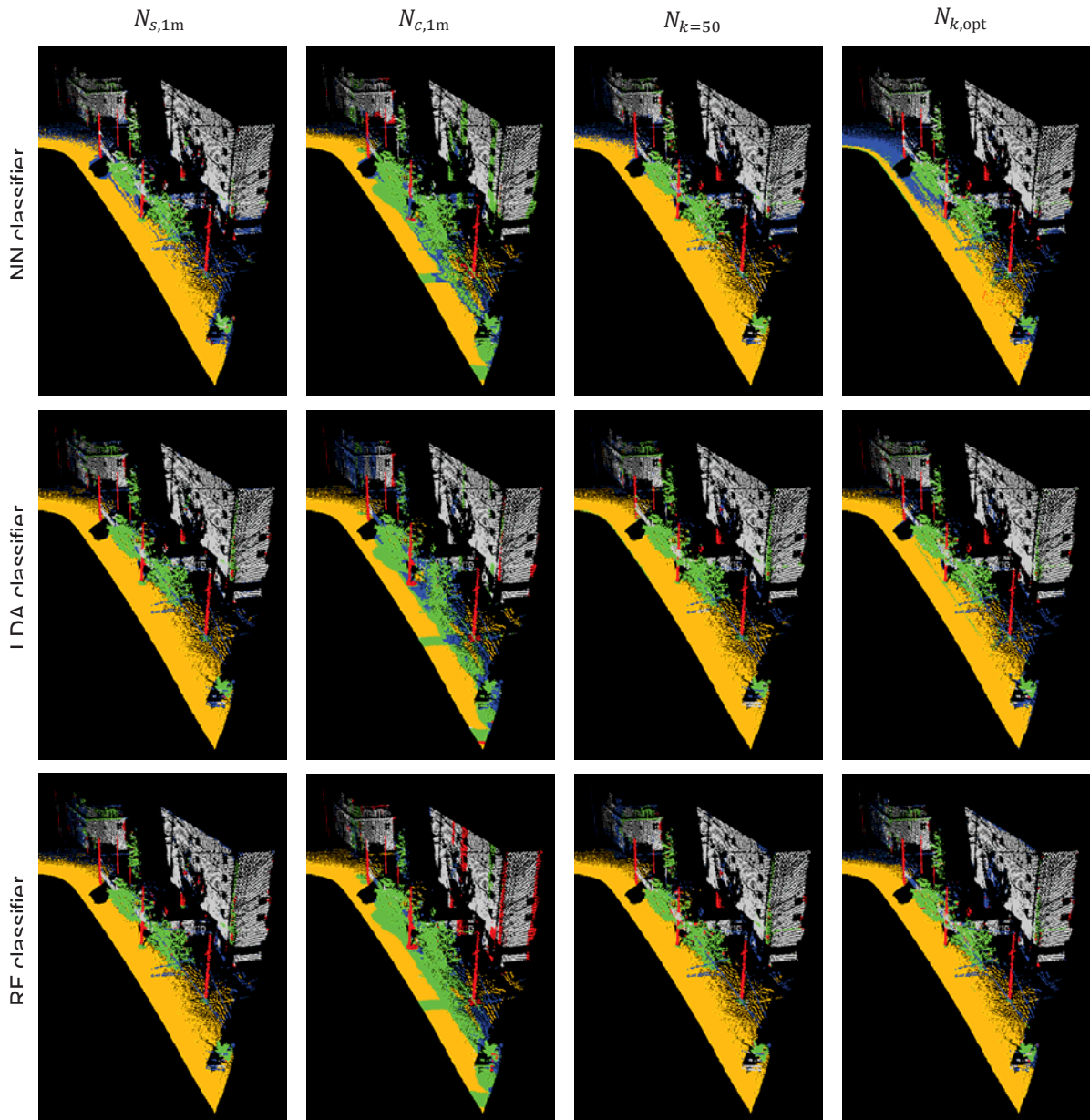


Fig. 6: The observed scene classified with respect to five semantic classes which are represented by wire (blue), pole/trunk (red), *façade* (gray), ground (brown) and vegetation (green).

In addition, it should be taken into account that the neighborhood types $N_{s,1m}$, $N_{c,1m}$ and $N_{k=50}$ rely on a scale parameter that is selected heuristically and kept identical for all points of the point cloud. For a 3D point cloud acquired with a different scanning device or within a different scene, the point sampling might exhibit different characteristics so that the selected scale parameters might not be that suitable anymore. In contrast, the neighborhood type $N_{k,opt}$ allows a data-driven selection of the scale parameter for each individual 3D point. Accordingly, it takes into account that different classes might favor a different neighborhood size. Furthermore, it better accounts for variations in point density (Fig. 5 and Fig. 6), particularly when using a LDA classifier or a RF classifier.

Among the considered classifiers, the LDA classifier and the RF classifier yield results of similar quality and tend to outperform the NN classifier for almost all cases. Due to the simplifying assumptions made with the LDA classifier – which cannot be guaranteed for the acquired data – using the RF classifier seems to be favorable.

Consequently, we conclude that local neighborhoods selected via *eigenentropy-based scale selection* are to be preferred as the basis for feature extraction, as respectively extracted features allow deriving appropriate classification results when using a RF classifier which we consider as the favorable option for classification.

5 Conclusions

In this paper, we have considered different neighborhood types and analyzed the respective behavior of low-level geometric 2D and 3D features and their relevance for point cloud classification. We have presented a framework allowing for the use of four neighborhood types as the basis for extracting a set of 18 low-level geometric 2D and 3D features which, in turn, are provided as input to classification. For the latter, we involved classifiers relying on three different learning principles. The derived results for a benchmark dataset with two reference labelings indicate that the relevance of extracted features with respect to a classification task may significantly vary depending on the selected neighborhood type. While the selected cylindrical neighborhood type clearly reveals less suitability with respect to the classification task, particularly the neighborhood types relying on a number of nearest neighbors lead to appropriate classification results, especially for the stronger classifiers based on probabilistic and ensemble learning.

In the scope of future work, it seems worth investigating the behavior of multi-scale neighborhoods in more detail. Thereby, the computational burden is typically significantly higher so that efficient approaches are required as for instance a recently proposed approach relying on a scale pyramid that is created by repeatedly downsampling a given 3D point cloud (HACKEL et al. 2016). Additionally, we plan to address the (partially) noisy behavior of the derived labeling. In this regard, we intend to use different approaches to impose spatial regularity on the classification results. A further aim consists in the transfer of the derived classification results to the level of single objects in the observed scene as a first step towards scene understanding.

6 References

- BLOMLEY, R., JUTZI, B. & WEINMANN, M., 2016: 3D semantic labeling of ALS point clouds by exploiting multi-scale, multi-type neighborhoods for feature extraction. Proceedings of the International Conference on Geographic Object-Based Image Analysis, 1-8.
- 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 & Remote Sensing* **68**, 121-134.
- CHEHATA, N., GUO, L. & MALLET, C., 2009: Airborne lidar feature selection for urban classification using random forests. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **38**(3/W8), 207-212.
- 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.
- GUO, B., HUANG, X., ZHANG, F. & SOHN, G., 2015: Classification of airborne laser scanning data using JointBoost. *ISPRS Journal of Photogrammetry & Remote Sensing* **100**, 71-83.
- 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.
- LALONDE, J.-F., UNNIKRISHNAN, R., VANDAPEL, N. & HEBERT, M., 2005: Scale selection for classification of point-sampled 3D surfaces. Proceedings of the International Conference on 3-D Digital Imaging and Modeling, 285-292.
- 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 & Remote Sensing* **66**(6), S71-S84.

- MITRA, N. J. & NGUYEN, A., 2003: Estimating surface normals in noisy point cloud data. *Proceedings of the Annual Symposium on Computational Geometry*, 322-328.
- 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 & Remote Sensing* **87**, 152-165.
- PAULY, M., KEISER, R. & GROSS, M., 2003: Multi-scale feature extraction on point-sampled surfaces. *Computer Graphics Forum* **22**(3), 81-89.
- SCHMIDT, A., NIEMEYER, J., ROTTENSTEINER, F. & SOERGEL, U., 2014: Contextual classification of full waveform lidar data in the Wadden Sea. *IEEE Geoscience and Remote Sensing Letters* **11**(9), 1614-1618.
- 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.
- 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.

Kalibrierung von Kameras mit allgemeinen, stetigen und nicht-symmetrischen Kameramodellen

DOMINIK RUEß¹ & RALF REULKE¹

Zusammenfassung: In diesem Beitrag stellen wir eine neue Kalibrierungsmethode vor. Diese Methode kann potentiell nicht-symmetrische Kameras modellieren, die möglicherweise auch keinen einzelnen Brennpunkt haben. In diesem Artikel vergleichen wir dieses Verfahren mit anderen Kalibriermethoden und beschränken uns wegen der Vergleichbarkeit dabei vorerst auf symmetrische Kameras mit einem Brennpunkt.

Wir verfahren klassisch: Zuerst initialisieren wir das System mit linearen Methoden. Danach wird das System nicht-linear optimiert. Wir vergleichen dazu verschiedene Möglichkeiten die sich aus dieser neuen Modellierung ergeben.

Wir zeigen, dass die Methode mit den klassischen Methoden mithalten kann und sich in extremen Situationen sehr gut verhält - z.B. bei kurzbrennweitigen Optiken.

1 Einleitung

Die geometrische Vermessung bzw. Kalibrierung von Kameras ist schon sehr lange etabliert. Etwas neuer sind die Forschungen bezüglich der Abweichungen von einem klassischen Lochkameramodell. Erste Verzeichnungsmodelle kamen früh auf, siehe zum Beispiel BROWN 1971. Stark verzeichnende Optiken sind Katadioptriken und Fischaugen, diese wurden mathematisch zum Beispiel in BAKSTEIN (2001), GONÇALVES (2004) und GONÇALVES (2008) beschrieben. Ein anderes Format ist das Panorama, es hat bereits früh begeistert (LUHMANN 2004; OLEKSIJCZUK 2011) bieten Rückblicke in die Panorama-Geschichte; diese wurden mathematisch ebenfalls schon gut untersucht von STURM et al. (2005), GENNERY (2006), MEI & RIVES (2007), LUHMANN (2010) und FRAHM et al. (2012).

Noch allgemeinere Kameras, bei denen jeder Pixel ein eigenes, mathematisches „Objektiv“ besitzt, wurden in FLOREANO et al. (2013), NAYAR et al. (2006) und ZAPPE (2010) untersucht. Dazu gehört zum Beispiel auch die Untersuchung von Insektenaugen in FRANCESCHINI et al. (1992) und NEUMANN et al. (2004).

Allgemeine, stetige und glatte Kameramodelle kamen in den letzten Jahrzehnten auch immer häufiger auf (GROSSBERG & NAYAR 2001 & 2005, STURM et al. 2005, BARRETO et al. 2010, MIRALDO & ARAUJO 2013). Dort wird kaum noch eine Vorannahme getroffen, zum Beispiel werden Forderungen nach Symmetrie oder einem einzigem Brennpunkt entfernt. Diese Methoden bedürfen aber großen Kalibrieraufwandes, Beispielsweise wurde in in der *Raxel*-Beschreibung von STURM et al. 2005 jeder Pixel per Laser in zwei Objektraum-Punkten vermessen, um eine Raum-Gerade für jeden Pixel bestimmen zu können. Dieses Prinzip wurde verbessert, um nicht mehr jeden einzelnen Lichtstrahl separat vermessen zu müssen (LI et al. 2008). Moderne Methoden wie in MIRALDO & ARAUJO (2013) benötigen aber immer noch eine separate Lagemessung zu jeder Aufnahme.

¹ Humboldt Universität zu Berlin, Unter den Linden 6, 10099 Berlin,
E-Mail: ruess@informatik.hu-berlin.de

In diesem Artikel beschreiben wir eine weitere Methode, die das Kalibrieren mit normalen, existierenden 3D-Kalibrier-Mustern erlaubt und allgemeine, unbeschränkte und glatte Kameras beschreiben kann. Die Basis dafür ist die mathematische Liniengeometrie, wie sie auch in POTTMANN & WALLNER (2001) beschrieben wird. Anhand von existierenden Lochkamera-ähnlichen Systemen haben wir dieses Modell untersucht. In zukünftigen Forschungen werden wir die Anwendbarkeit auf komplett unbeschränkte Systeme untersuchen, auch z.B. ohne Brennpunkt.

2 Theorie

Betrachtet man fokussierte Abbildungssysteme (wie Kameras), dann sind die allermeisten davon eine Projektion zwischen Geraden im Raum und einem Pixel auf dem Sensor. Im Gegensatz zur klassischen Photogrammetrie und zur Computer Vision stellen wir hier eine Abbildung vor, die zweidimensionale Sensorkoordinaten in Raumgeraden abbildet. Diese Abbildungsrichtung ist umgekehrt zur Computer Vision und Photogrammetrie. Das hat aber den Vorteil, dass man Bildkoordinaten direkt und analytisch in Raumgeraden abbilden kann, ohne ein iteratives bzw. numerisches Optimierungsverfahren.

2.1 Lineare Kalibrierung

Im Folgenden stellen wir eine Initialisierungsmethode vor, die es auch erlaubt nicht-lineare Basisfunktionen des Raumes stetiger Funktionen zu benutzen. Diese Methode kommt mit Methoden der linearen Algebra aus. Insbesondere sind die gesuchten Lösungen Teil von konvexen, und damit leicht zu lösenden Problemstellungen.

2.1.1 Plücker Koordinaten

Der Raum der Geraden aus $\mathbb{R}P^3$ ist eine vierdimensionale Mannigfaltigkeit. Es gibt verschiedene Darstellungen dieses Raums. Plücker-Koordinaten ist eine davon. Sie sind Koordinaten einer Quadrik im 5-dimensionalen projektiven Raum. Es gibt eine konkrete Darstellung einer Geraden, in Form eines 6-Vektors, dessen erster und zweiter 3-Teilvektor senkrecht aufeinander stehen:

$$L = (l, \bar{l}), \text{ mit } l^T \bar{l} = 0, \quad (1)$$

auch bekannt als *Plückerbedingung*. l und \bar{l} können als Richtung, respektive Momentenvektor (im Ursprung) interpretiert werden. Die Darstellung einer solchen Gerade ist eineindeutig -bis auf einen Skalierungsfaktor. Eine Linearkombination von Geradenkoordinaten ist nicht automatisch wieder eine Gerade, da die Plückerbedingung nicht automatisch erfüllt ist. Sei also A eine Matrix in der die Spalten jeweils eine Gerade repräsentieren,

$$C = \begin{pmatrix} A \\ \bar{A} \end{pmatrix} \in \mathbb{R}^{6 \times n}. \quad (2)$$

Das heißt also, dass nur dann sämtliche Linearkombinationen wieder eine Gerade ergeben, wenn $x^T \bar{A}^T A x = 0, \forall x \in \mathbb{R}P^n$. Aus der linearen Algebra folgt, dass $\bar{A}^T A$ schiefsymmetrisch ist.

Lässt man nun $n=3$ sein, dann ist C eine lineare Abbildung von $\mathbb{R}P^2$ in den Raum der (projektiven) Geraden. Also insbesondere auch eine Abbildung von Sensorkoordinaten in den Raum der Geraden des \mathbb{R}^3 - also eine lineare Kamera.

2.1.2 Lineare Kalibrierung

Die Punkt-Linien-Inzidenz-Relation einer mit Plücker-Koordinaten gegebenen Gerade L und eines mit homogenen Koordinaten gegebenen Punktes $\tilde{X} = (X, x_4) = (x_1, x_2, x_3, x_4)$ ist in folgenden vier linearen Gleichungen gegeben:

$$\begin{aligned} \bar{l}^T X &= 0, \\ -x_4 \bar{l} + X \times l &= 0 \quad (3) \end{aligned}$$

Betrachtet man eine Linearkombination von Geraden, siehe Gleichung (2), dann werden die Gleichungen (3) zu:

$$\begin{aligned} x^T \bar{A}^T X &= 0, \\ -x_4 \bar{A}x + X \times Ax &= 0. \quad (4) \end{aligned}$$

Dieses lineare Gleichungssystem kann benutzt werden, um eine Kalibrierung durchzuführen: \bar{A} und A sind die $6n$ Unbekannten der "Kameramatrix" C . Für eine Bestimmung von C benötigt man also mindestens 6 Punkte, da das Gleichungssystem von (4) einen Rang von 3 hat.

Zusätzlich sollten die Weltpunkte nicht-planar sein, da sonst eine Mehrdeutigkeit entsteht, die nur durch weitere Annahmen oder mehrere Ansichten reduziert werden kann.

In POTTMANN & WALLNER (2001) wurde gezeigt, dass C ein Linienbündel repräsentiert. Es existiert also ein Punkt in $\mathbb{R}P^3$, der in allen Geraden enthalten ist. Dieser Punkt ist hilfreich für die Bestimmung einer korrekten Matrix C , so dass die Plückerbedingung von Gleichung (1) eingehalten wird. Das Problem, dass C nach der Kalibrierung diese Bedingung nicht erfüllt, ist vor allem bei verrauschten bzw. echten Objekt- und Bilddaten gegeben. Mit den Gleichungen (4) kann also ein $\tilde{X} = (X, x_4)$ bestimmt werden, der algebraisch "bestmögliche" Schnittpunkt des Linienbündels gegenüber den Daten. Hat man \tilde{X} bestimmt, kann man mit Hilfe der Daten ein C bestimmen, das die Plückerbedingungen erfüllt und algebraisch am besten zu den gegebenen Daten passt.

Somit gibt es eine algebraisch bestmögliche lineare Initialisierung für die Kamera C aus den Gleichungen (4), die als Initialisierung in einer weiteren Optimierung verwendet werden kann.

2.1.3 Erweiterung auf beliebige Basisfunktionen

In diesem Abschnitt betrachten wir nun die Erweiterung der im vorherigen Abschnitt vorgestellten Methoden auf eine endliche Anzahl von beliebigen und stetigen Basisfunktionen des differenzierbaren Funktionenraums. Damit können wir eine direkte, nicht-lineare Kamerainitialisierung durchführen, was insbesondere für stark verzeichnende Kameras von Vorteil ist.

Sei also $B = \{b_1(x), b_2(x), \dots, b_n(x)\}$ eine Menge von Basisfunktionen, für $x \in \mathbb{R}^2$. Nun lässt sich eine Kamera C bestimmen, mit einer Linearkombination aus den Basisfunktionen, also $C \in \mathbb{R}^{6 \times n}$. Das heißt also die Abbildung von Sensorkoordinaten in den Geradenraum wird mit $= C \cdot (b_1(x), \dots, b_n(x))$ beschrieben.

Als Basisfunktionen bieten sich Polynome an, zum Beispiel: $B = \{1, x_1, x_2, x_1 x_2, x_1^2, x_2^2\}$. Aber auch orthogonale Basisfunktionen funktionieren hier gut, vor allem da sie sich durch ihre höhere numerische Stabilität und geringere Approximationsfehler in den Randbereichen besser für die Funktionen-Approximation eignen. Ein Beispiel hierfür sind Legendre-Polynome.

Die Kalibrierung erfolgt dann wie in Gleichung (4), vom vorherigen Abschnitt.

2.2 Nicht-Lineare Optimierung

Die lineare Approximation funktioniert zwar sehr gut birgt aber auch Nachteile: Die Optimierung ist nicht global und der Raum der Geraden ist nicht orientiert. Letzteres bedeutet, dass bei der Approximation der Kamera C die Geraden sowohl nach vorne als auch nach Hinten zeigen können, dazu gibt es möglicherweise Nulldurchgänge im Richtungsvektor.

Für eine globale Optimierung entschieden wir uns für eine andere Beschreibung. Eine zweidimensionale Mannigfaltigkeit im Raum der Orientierten Geraden, auch Linien-Kongruenz genannt, kann man auch mit zwei Oberflächen bzw. zweidimensionalen Mannigfaltigkeiten im \mathbb{R}^3 beschreiben: Die erste Oberfläche, $O(x) \in \mathbb{R}^3$ beschreibt den Ort der Geraden, ist also eine Menge von Punkten, die jeweils einer Geraden zugeordnet werden. Die zweite Oberfläche ist eine Richtungsbeschreibung, $D(x) \in \mathbb{R}^3$. Letztere Oberfläche hat aber die Möglichkeit, dass die Richtung 0 sein kann, was für alltägliche Kameras unüblich ist. Auch ist die Richtung nicht eindeutig, da der Richtungsvektor eine unterschiedliche Länge haben kann. Gewünscht ist also eine beliebige Funktion auf der Kugeloberfläche $D(x) \in S^2$. Im Folgenden Abschnitt stellen wir solche Funktionen vor.

Da die Abbildung $D(x)$ beliebig ist, entfallen auch alle notwendigen Symmetrien wie sie in bisherigen Modellen existierten, insbesondere die radiale Symmetrie. Aber gleichzeitig sind wohl mehr Parameter notwendig, um bei klassischen Kameras eine ähnliche Genauigkeit wie bei radialsymmetrischen Modellen zu erhalten.

2.2.1 Kartenprojektionen

Im vorherigen Abschnitt 2.2 haben wir eine allgemeine Beschreibung einer nicht-linearen Kamera anhand von Oberflächen dargestellt. Hier wollen wir zeigen, wie eine Richtungsoberfläche effizient und mit obigen Bedingungen formuliert werden kann. Die Einheitssphäre S^2 kann topologisch nicht auf die Ebene projiziert werden. Aber bis auf einen Punkt der Sphäre ist es möglich. Klassischerweise wird das Problem bei diversen Kartenprojektionen bereits gelöst. Für unsere Problematik bieten sich verschiedene solcher Projektionen an, wir haben die folgenden drei ausgewählt: Stereographische Projektion, flächentreue Azimutalprojektion und die mittabstandstreue Azimutalprojektion. Diese Projektionen haben den Vorteil gegenüber den meisten anderen Projektionen, dass sie - bis auf einen einzigen Punkt der Sphäre - alle anderen Punkte auf die Eben glatt projizieren. Andere Projektionen haben oft mehr als nur eine einzige Diskontinuität, wie zum Beispiel die Darstellung nach Längen- und Breitengraden, wo es eine Winkel-Diskontinuität von π auf $-\pi$ entlang des gesamten Null-Meridians gibt.

Die Abbildung von Sensorkoordinaten in Richtungsvektoren ist also eine Abbildung von Sensorkoordinaten in Kartenkoordinaten.

Ähnlich wie oben können nun zweidimensionale Oberflächen benutzt werden, wir wählen bivariate Polynome. In zukünftigen Arbeiten werden wir den Einsatz von anderen Approximationsmethoden untersuchen, zum Beispiel Radiale-Basis-Funktionen.

Die Abbildung in den orientierten Geradenraum lautet also

$$L = O(x) + \alpha \cdot D(x); \forall \alpha \in \mathbb{R}.$$

Dabei ist $D(x) = (D_1(x), D_2(x))$ eine zweidimensionale Abbildung, von jeweils der gewählten Funktionsart.

Für diesen Artikel benutzen wir momentan nur konstante Ortsfunktionen, d.h. Kameras mit einem einzigen Brennpunkt: $O(x) = (c_1, c_2, c_3); \forall x \in \mathbb{R}^2$, d.h. mit den c_i konstant. In zukünftigen Arbeiten werden wir die Wirkung von nicht-zentralen Modellen weiter untersuchen.

2.3 Kalibrierung Nichtsymmetrischer Kameras

Die im Abschnitt 2.2 beschriebene nichtlineare Abbildung von Sensorkoordinaten in den orientierten Geradenraum hat eine Menge von Unbekannten: Einerseits die Konstanten c_i der Ortsfunktion und andererseits die Koeffizienten der Richtungspolynome von $D(x)$. Abgesehen davon kommen, wie in jeder Kalibrierung die unbekannte Translation und Rotation hinzu.

Die initiale Rotation kann aus der QR-Zerlegung von A aus Gleichung (2) gewonnen werden. Die Translation ergibt sich aus den linearen Kamerazentren.

Für die initialen Richtungskoeffizienten müssen die nicht-orientierten Geraden gegenseitig angepasst werden, so dass benachbarten Geraden in die gleiche Richtung zeigen und alle linearen Kameras die gleiche Hauptrichtung haben (z.B. nach vorne oder nach hinten). Die Richtungen können mit den Rotationen in ein gemeinsames System transformiert werden. Damit kann ein initiales Parameterset für $D(x)$ geschätzt werden, zum Beispiel mit der Vandermonde-Matrix der Polynominterpolation.

Mit dem Levenberg-Marquart-Algorithmus kann die gesamte Parameterfamilie optimiert werden. Da L aus Gleichung (1) im Objektraum definiert wurde, kann eine Optimierung ohne die Bestimmung von Pixel Fehlern im Bildraum durchgeführt werden. Ein Fehlermaß, welches uns sehr gute Ergebnisse liefert, ist der Fehlerwinkel φ_i zwischen Brennpunkt $O(x_i)$, der Geradenrichtung des Modells $D(x_i)$ und dem Objektpunkt X_i für jedes Punktpaar (X_i, x_i) .

Die Optimierung minimiert also folgende Energie:

$$P = \operatorname{argmin}_P \sum_i \varphi_i^2,$$

wobei P die Menge alle Unbekannten beschreibt.

Die Fehler in Bildkoordinaten können nun am Ende berechnet werden. Das bedeutet, dass die teure, nicht-analytische Inversion des Modells nur einmal für alle Punktpaare durchgeführt werden muss.

3 Ergebnisse

In diesem Kapitel beschreiben wir die Ergebnisse von Initialisierung und Kalibrierung verschiedener Kameraarten. Die Ergebnisse der Kalibrierung deutlich nicht-symmetrischer Kameras werden in einem späteren Bericht folgen.

3.1 Lineare Initialisierung

Die lineare Initialisierung mit bivariaten Legendre-Polynomen vom Grad 3 lieferte sehr gute Ergebnisse für alle Kameraklassen. Man benötigt hierzu aber mindestens 30 Punkt-Paare, um eine Unterbestimmung der Gleichungssysteme zu vermeiden.

Tab. 1: Initialisierungsfehler für die verschiedenen Kameras aus Abb. 1. Benutzt wurden bivariate Legendre-Polynome vom Grad 3. Weitere Details bitte aus der Abbildung entnehmen.

Kamera-Sensor-Daten Brennweite Vollformat (FF), Auflösung, Sensorgröße	Initialisierung; RMS [px]
f=1,4mm (FF); 1280 x 960 px; 8,3 x 6,2 mm	1,50
f=4,6mm (FF); 1392 x 1040 px; 8,8 x 6,6 mm	6,97
f=13,6mm (FF); 768 x 488 px; 5,1 x 3,3 mm	1,56
f=17,7mm (FF); 2592 x 1944 px; 3,2 x 2,4 mm	44,9
f=28,3mm (FF); 5472 x 3648 px; 13,2 x 8,8 mm	2,95
f=146,9mm (FF); 1360 x 1024 px; 8,8 x 6,6 mm	0,51
f=150,2mm (FF); 720 x 576 px; 4,6 x 3,7 mm	9,29

3.2 Kalibrierung von Kameras und Vergleich mit anderen Methoden

Für eine Auswahl von Kameraklassen (Katadioptrik, Fischaugen, kurze, mittlere und lange Brennweiten) haben wir Vergleiche von verschiedenen Kalibrierungsmethoden durchgeführt. Zu den Vergleichsmethoden gehören unser eigenes Modell in den verschiedenen Kartenprojektions-Varianten und verschiedene Polynomgraden, eine Photogrammetrische Kalibrierung mit dem Australis-Modell, dem Allgemeinen und radialsymmetrischen Modell von LUBER et al. (2012), und der Computer-Vision-Kalibrierung mit dem OpenCV-Modell.

Die Ergebnisse sind in Abb. 1 illustriert. Im Allgemeinen funktioniert unser Modell sehr gut und kann mit den anderen Methoden mithalten. Im Speziellen, vor allem für stark verzeichnende Kameras, funktioniert unser Modell, ähnliche wie das Modell von LUBER et al. (2012), deutlich besser als die beiden anderen Varianten. Die verschiedenen Kartenprojektionen liefern ähnlich gute Ergebnisse. Alle Methoden entfernen Ausreißer: Australis z.B. entfernt Ausreißer, die einen Fehler von einem vielfachem der Standardabweichung aufweisen. Wir entfernen die Abweichler ähnlich, schauen aber nur in der Nachbarschaft des Punktes, da die Verzerrung in unterschiedlichen Bildbereichen unterschiedliche Fehler-Niveaus haben kann.

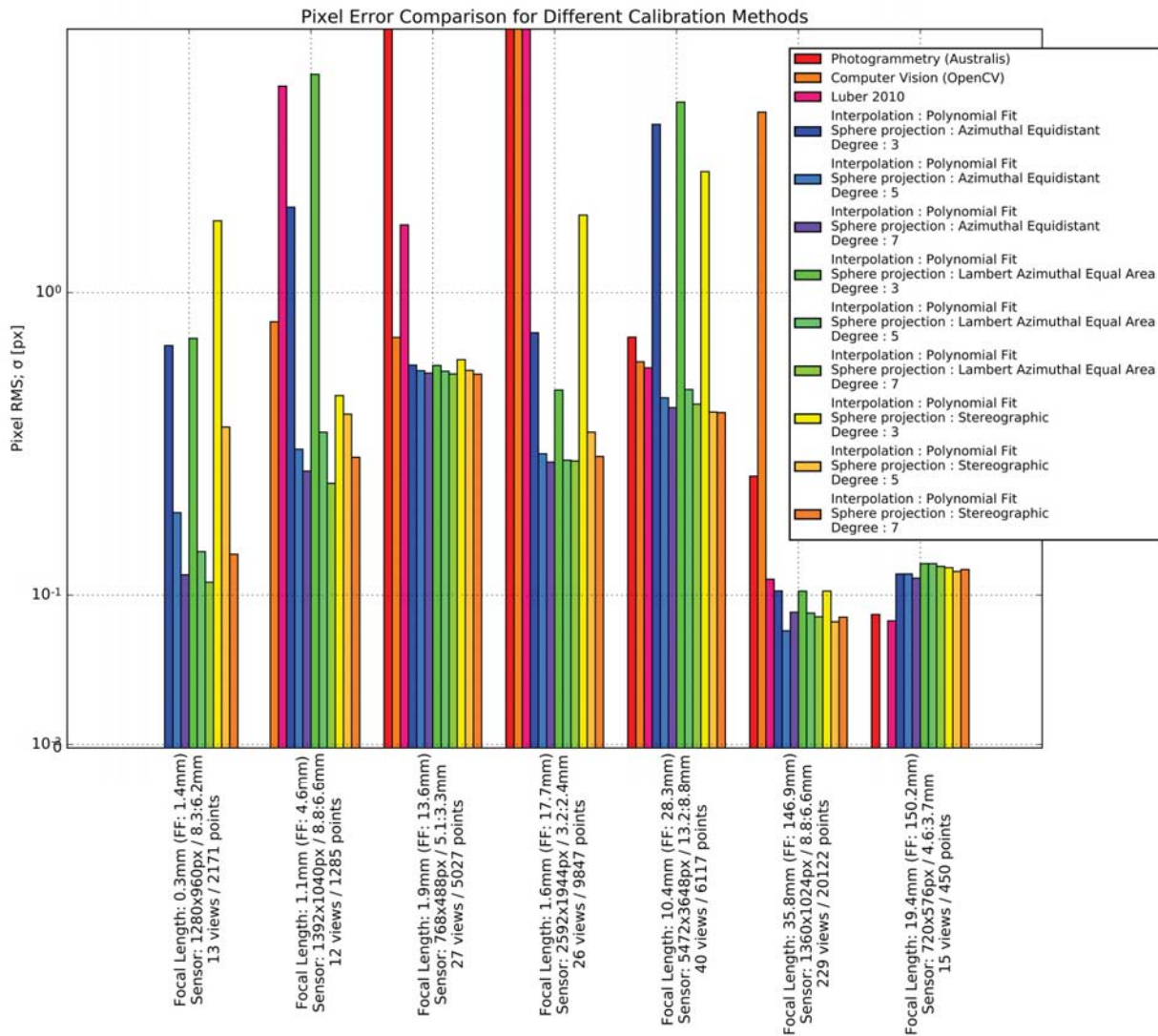


Abb. 1: Vergleich verschiedener Kalibriermethoden mit mehreren Varianten unserer Methode, siehe Legende. Nicht jede Methode konnte alle Kameras kalibrieren, dann fehlt der entsprechende Balken. Die Kameras sind nach Kleinbildäquivalent-Brennweite – Full Frame (FF) – sortiert.

3.3 Low-Cost Kalibrierung einer 360-Grad Kamera

Für eine Demonstration der Vielseitigkeit der hier eingeführten Kalibriermethode wollen wir die Ergebnisse der Kalibrierung einer 360 Grad Kamera, der Samsung Gear 360™ vorstellen, siehe Abb. 2. Die Kalibrierung führten wir mit Materialien aus dem Baumarkt für etwa 50€ durch. Ein dreiseitiges, etwa rechtwinklig aufeinander geklapptes Kalibriermuster beklebten wir mit gedruckten Kreisen. Als Orientierung druckten wir verschiedene Pfeile in die Ecke, damit die Zuordnung der Kreise möglich wurde, siehe Abb. 4.



Abb. 2: Samsung Gear 360™ Kamera mit 360 Grad Rundumsicht und 2x15 Megapixel.

Durch die starke Verzerrung der Kreise wird ein reiner Ellipsenfit eher ungenau. Um die Genauigkeit zu erhöhen, passten wir den Ellipsenfit mit einem bivariaten Flächen-Polynom an, siehe Abb. 3.

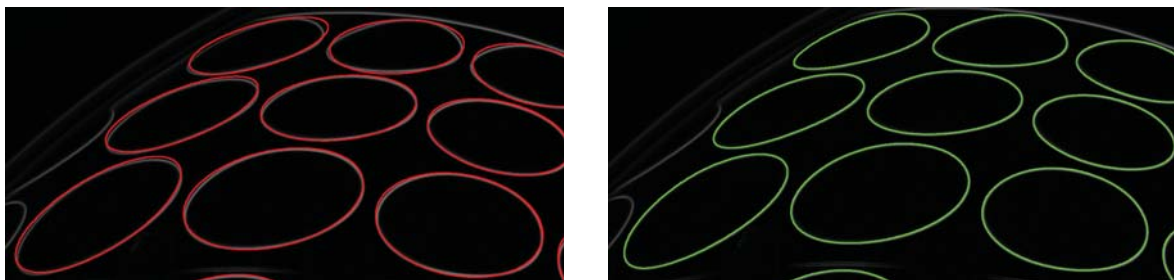


Abb. 3: Schätzung von Ellipsen an verzerrt abgebildete Kreise (li) und bivariate, polynomieller Optimierung derer Form.

Die Punkte des Kalibrieremusters haben wir grob abgeschätzt, sie wurden dann in der Kalibrierung selbst mitoptimiert. Die beiden Teil-Kameras haben wir separat kalibriert, anhand von jeweils 14 bzw. 17 Aufnahmen. Die Ergebnisse sind unter **Fehler! Verweisquelle konnte nicht gefunden werden.** aufgelistet. Die Kalibrierung funktioniert so immerhin bis auf ein Pixel-RMS von 3-5 Pixeln. Das ist nicht perfekt aber für approximative Anwendungen vielleicht in Ordnung. In zukünftigen Untersuchungen werden wir die Kamera auch mit genaueren Kalibrieremustern vermessen.

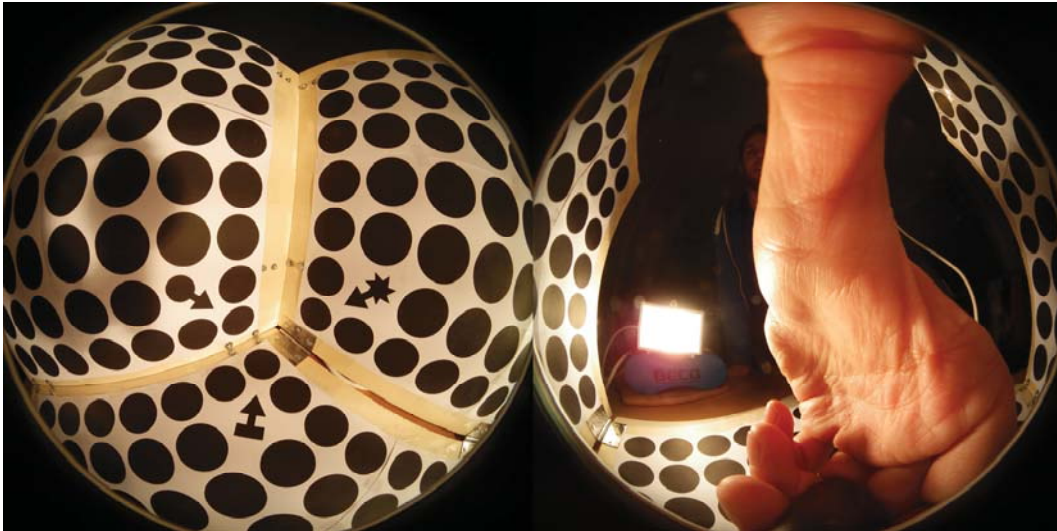


Abb. 4: Ein Beispielbild der Samsung Gear 360 Kamera mit dem Low-Cost-Kalibriermuster

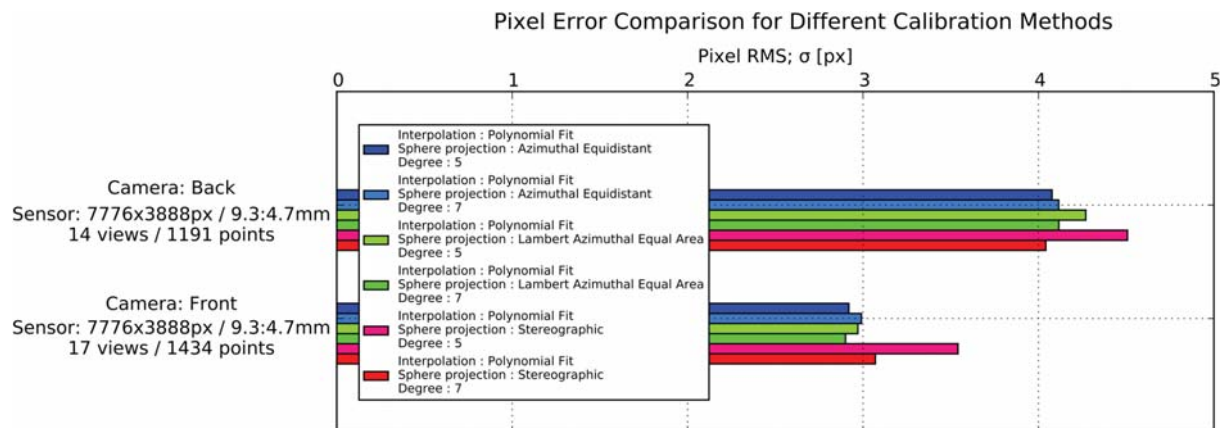


Abb. 5: Ergebnisse der Kalibrierung beider (Teil-)Kameras des 360Grad-Systems

4 Fazit & Ausblick

In diesem Beitrag haben wir eine neue, allgemeine Methode zur Beschreibung von Kamerageometrien vorgestellt. Sie funktioniert, wie wir demonstriert haben, für klassische Kameras, kann aber auch nicht-symmetrische Kameras ohne Brennpunkt beschreiben. Das werden wir in folgenden Artikeln weiter untersuchen. Die Modelle werden im Objektraum optimiert. Wir stellten eine erste Abschätzung vor, die auch für stark verzeichnende Kameras eine sehr gute Initialisierung mittels linearer Gleichungssysteme bieten. Die darauf folgende, nicht-lineare Optimierung kann nun eine große Breite von Kameramodellen potentiell kalibrieren.

Die vorgestellten Methoden sollen die etablierten Methoden nicht ersetzen, denn gerade für wenig verzeichnende Kameras sind die etablierten Methoden sehr genau und sehr schnell. Wir konnten aber zeigen, dass unsere Methoden für diese Kameras durchaus mithalten können.

In den folgenden Untersuchungen werden wir weitere Kameras kalibrieren, auch sehr asymmetrische Modelle, zum Beispiel Spiegelnde Systeme die über klassische Katadioptriken

hinausgehen. Auch ist noch offen, ob eine Erweiterung des einzelnen Brennpunktes, hin zu Mannigfaltigkeiten von Brennpunkten, vielleicht eine noch genauere Kalibrierung klassischer Kameras ermöglicht. Vielleicht gibt es damit die Möglichkeit, die Brennpunktgerade von Fischaugen-Optiken besser zu beschreiben.

5 Literaturverzeichnis

- BAKSTEIN, H., 2001: Non-central cameras for 3D reconstruction. PhD thesis, Czech Technical University, Prague.
- BARRETO, J.P., PAJDLA, T., SUGIMOTO, A., RAMALINGAM, S., STURM, P. & LODHA, S.K., 2010: Generic self-calibration of central cameras. *Computer Vision and Image Understanding* **114**(2), 210-219.
- BROWN, D., 1971: Close range camera calibration. *Photogrammetric Engineering* **37**, 855-866.
- FLOREANO, D., PERICET CAMARA, R., VIOLLET, S., RUßER, F., BRÜCKNER, A., LEITEL, R., BUSS, W., MENOUNI, M., EXPERT, F., JUSTON, R., DOBRZYNSKI, M.K., L'EPLATTENIER, G., RECKTENWALD, F., MALLOT, H.A. & FRANCESCHINI, N., 2013: Miniature curved artificial compound eyes. *Proceedings of the National Academy of Sciences* **110**(23), 9332-9337.
- FRAHM, J.M., POLLEFEYS, M., DELLAERT, F., KOSECKA, J., PUIG, L., BERMÚDEZ, J., STURM, P. & GUERRERO, J., 2012: Calibration of omnidirectional cameras in practice: A comparison of methods. *Computer Vision and Image Understanding* **116**(1), 120-137.
- FRANCESCHINI, N., PICHON, J.M., BLANES, C. & BRADY, J., 1992: From insect vision to robot vision [and discussion]. *Philosophical Transactions of The Royal Society Of London. Series B: Biological Sciences* **337**(1281), 283-294.
- GENNERY, D.B., 2006: Generalized Camera Calibration Including Fish-Eye Lenses. *International Journal of Computer Vision* **68**(3), 239-266.
- GONÇALVES, N., 2004: Projection Model, 3D Reconstruction and Rigid Motion Estimation from Non-central Catadioptric Images. *Proceedings of 2nd International Symposium on 3D Data Processing, Visualization and Transmission*.
- GONÇALVES, N., 2008: Noncentral catadioptric systems with quadric mirrors: geometry and calibration. PhD thesis, Universidade de Coimbra.
- GROSSBERG, M. & NAYAR, S., 2001: A general imaging model and a method for finding its parameters. *Proceedings Eighth IEEE International Conference on Computer Vision* **2**, 108-115.
- GROSSBERG, M. & NAYAR, S., 2005: The Raxel Imaging Model and Ray-Based Calibration. *International Journal of Computer Vision* **61**(2), 119-137.
- LI, W.S., BOTHE, T., SCHULTE, M., KOPYLOW, C. v., KÖPP, N. & JÜPTNER, W., 2008: Sichtstrahlkalibrierung für optisch abbildende Systeme. *DGaO-Proceedings*.
- LUBER, A., RUEB, D., MANTHEY, K. & REULKE, R., 2012: Calibration and epipolar geometry of generic heterogenous camera systems. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39**(B5), 363-368.
- LUHMANN, T., 2004: A historical review on panorama photogrammetry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **34**(5/W16), 8.

- LUHMANN, T., 2010: Panorama photogrammetry for architectural applications. *Mapping*, 40-45.
- MEI, C. & RIVES, P., 2007: Single View Point Omnidirectional Camera Calibration from Planar Grids. *Proceedings on IEEE International Conference on Robotics and Automation*, 3945-3950.
- MIRALDO, P., ARAUJO, H., 2013: Calibration of smooth camera models. *IEEE transactions on pattern analysis and machine intelligence* **35**(9), 2091-2103.
- NAYAR, S.K., BRANZOI, V. & BOULT, T.E., 2006: Programmable Imaging: Towards a Flexible Camera. *International Journal of Computer Vision* **70**(1), 7-22.
- NEUMANN, J., FERMLER, C., ALOIMONOS, Y. & BRAJOVIC, V., 2004: Compound eye sensor for 3d ego motion estimation. *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 3712-3717.
- OLEKSIJCZUK, D., 2011: *The First Panoramas: Visions of British Imperialism*. University of Minnesota Press.
- POTTMANN, H. & WALLNER, J., 2001: *Computational Line Geometry*. Springer-Verlag Berlin.
- STURM, P., RAMALINGAM, S. & LODHA, S., 2005: On calibration, structure-from-motion and multi-view geometry for panoramic camera models. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**(5/W8).
- ZAPPE, H.P., 2010: *Fundamentals of Micro-optics*. Cambridge University Press.

KuLaDig – eine Kooperationsplattform zur Erfassung und Bewahrung der Kulturlandschaft und ihrer Entwicklung

PETER BURGGRAFF¹, ELMAR KNEIPS², JÖRN SCHULTHEISS¹ & MICHAEL TEMPEL¹

Zusammenfassung: Die vom Landschaftsverband Rheinland (LVR) gemeinsam mit dem Landesamt für Denkmalpflege Hessen und dem Rheinischen Verein für Denkmalpflege und Landschaftsschutz e.V. betriebene Plattform KuLaDig (Kultur. Landschaft. Digital) stellt eine öffentliche, GIS-basierte Kooperationsplattform dar, die zur Bewahrung des kulturlandschaftlichen Erbes beiträgt. Mit Hilfe von Kooperationen wurden bisher über 8.300 Einträge eingereicht, redaktionell überarbeitet und öffentlich über KuLaDig zugänglich gemacht. Die Daten stehen somit sowohl der breiten Öffentlichkeit als auch Entscheidungsträgern der politischen und raumplanerischen Ebene zur Verfügung.

1 Einleitung

Bei der Erfassung und Bewahrung des kulturellen Erbes der Menschheit müssen auch Aspekte der Kulturlandschaft Berücksichtigung finden. Um die Kulturlandschaft und ihre Entwicklung dauerhaft zu dokumentieren und diese Daten zugleich einer breiten Öffentlichkeit im Sinne der Umweltbildung zugänglich zu machen, wurde und wird die Plattform KuLaDig (Kultur. Landschaft. Digital) entwickelt. Hierbei handelt es sich um eine GIS-gestützte Kooperationsplattform, die ebenso der Berücksichtigung kulturlandschaftlicher Werte bei Planungsvorhaben dient. Betrieben wird KuLaDig vom Landschaftsverband Rheinland (LVR) gemeinsam mit dem Landesamt für Denkmalpflege Hessen und dem Rheinischen Verein für Denkmalpflege und Landschaftsschutz e.V.; aufrufbar ist die Anwendung über www.kuladig.lvr.de bzw. www.kuladig.de.

2 Technische Voraussetzungen der Plattform KuLaDig

Das Informationssystem KuLaDig wird von LVR-InfoKom, dem Systemhaus des Landschaftsverbandes Rheinland, entwickelt und betrieben. Die aktuelle Programmversion 4 des Portals und die die Anwendung sowie ihre Nutzerinnen und Nutzer begleitende und betreuende Redaktion beim LVR sind das Ergebnis eines etwa 15 Jahre währenden Lern- und Entwicklungsprozesses. Die aktuelle Anwendung ist auf Basis modernster Microsoft Technologie realisiert. Die Grundlage bildet das Microsoft ASP.NET MVC Framework, die Daten werden in einer Microsoft SQL Server 2012 Datenbank gespeichert. Zur geräteübergreifenden Darstellung (Desktop, Tablets, Smartphones) wird Twitter Bootstrap eingesetzt. Darüber hinaus werden spezielle Kartendienste vom Ar-

¹ Universität Koblenz-Landau, Campus Koblenz, Abteilung Geographie, Universitätsstr. 1, D-56070 Koblenz, E-Mail: burggra@uni-koblenz.de, jschultheiss@uni-koblenz.de, mtempel@uni-koblenz.de

² Landschaftsverband Rheinland, Dezernat Kultur und Landschaftliche Kulturpflege, Redaktion KuLaDig / Abteilung Digitales Kulturerbe, Ottoplatz 2, 50679 Köln, E-Mail: Elmar.Knieps@lvr.de

cGIS-Server der Firma ESRI geliefert. Die Kartendarstellung basiert ebenfalls auf ESRI-Technologie. Die Kartenkomponente mit GIS-Funktionalität stellt über Webdienste eingebundene topographische Karten sowie Luftbilder und historische Karten als Hintergrundkarten für die Objekterfassung und zur Präsentation zur Verfügung. Die Kartenansicht, die Expertensuche und die Editierfunktionalitäten werden aktuell durch das Browser-PlugIn Microsoft-Silverlight unterstützt. Hierdurch stehen in der Web-Anwendung komfortable Funktionen zur Verfügung, die ein Arbeiten mit der Anwendung ohne GIS- oder Datenbank-Kenntnisse ermöglichen.

Ausgewählte Inhalte werden in Form von Webdiensten (WMS) extern zur Verfügung gestellt. Über Webdienste werden Medien einer Bilddatenbank dynamisch in Form einer Mediengalerie und ein extern gepflegter Thesaurus für Zwecke der Recherche und Erfassung in die Anwendung eingebunden. Es existiert eine mobile App für Android und iOS. Der Einstieg in die Nutzung von KuLaDig ist über die Startseite der Anwendung (Abb. 1) bewusst sehr niedrighschwellig angelegt. Ein jeweils aktuelles „Objekt des Monats“ leitet Gelegenheitsbesucher zu einem Objekteintrag (Abb. 2), von dem aus es sehr einfach ist, sich in einen Bereich der heutigen Kulturlandschaft zu bewegen und ihre Strukturen zu erkunden. Etwas weniger prominent, aber mit der gleichen Intention, wird eine „Auswahl interessanter Objekte“ ebenfalls auf der Startseite unterhalb des Objekts des Monats präsentiert. Im Kopf des Startbildschirms steht eine einfache Suchmöglichkeit „Was?“ / „Wo?“ zur Verfügung, die auch kombinierte Einträge erlaubt. In beiden Feldern sind Autocomplete-Funktionen hinterlegt, die zum einen passende Objektnamen und zum anderen Ortsbezeichnungen (Gazetteer) der Verwaltungsgliederung Deutschland, der naturräumlichen Einheiten und der Kulturlandschaften in Nordrhein-Westfalen zur Auswahl anbieten. Die Suche nach Ortsbezeichnungen nutzt als Gazetteer einen Shape-Layer, der ebenfalls in der Karte sowohl für die Anzeige als auch für die Zuordnung erfasster Objekte zu den hinterlegten Orts-Entitäten genutzt wird. Bereits auf der Startseite, wie auch bei jedem Objekteintrag, ist der Aufruf einer Hinweisfunktion möglich. Mithilfe dieser Funktion können Nutzerinnen und Nutzer der KuLaDig-Redaktion Nachrichten allgemeiner Art oder zu einem spezifischen Objekteintrag zukommen lassen.

Die Expertensuche (Abb. 3) erlaubt komplexe Abfragen über quasi alle Datenbankfelder, also über Text-, Listen- und Datumsfelder, über die Schlagwörter und Datierungsangaben, die die über einen Webdienst angebundene Thesaurus-Anwendung liefert, und schließlich über den Raum, den bereits erwähnten Gazetteer nutzend. Die Objektansicht, also die Darstellung eines einzelnen KuLaDig-Objektes, hat einen definierten Aufbau (Abb. 2). Unter dem Namen des Objekts werden eventuell vorhandene alternative Bezeichnungen/Namen gelistet, darunter die Schlagwörter als Links zu einem Public-Viewer, die Fachsicht(en) als Angabe aus welche(r/n) fachlichen Perspektive(n) das Objekt beschrieben wird, und Informationen zur Verortung (Gemeinde, Kreis, Bundesland). Es folgen unterhalb links die Mediengalerie, die ein Scrollen durch die dem Objekt zugewiesenen Medien und deren Aufruf in eine vergrößerte Ansicht ermöglicht, und rechts die Darstellung der Objektgeometrie auf einem Luftbildhintergrund. Dort, wo kein Luftbilddienst zur Verfügung steht, wird Open-Street-Map als Hintergrund verwendet. Unter der Mediengalerie folgt die textliche Beschreibung des Objekts mit der Angabe zur Autorenschaft sowie zu Quellen und Literatur, ausgewählten Metadaten zum Objekteintrag und einer empfohlenen Zitierweise. Rechts unter dem Luftbild wird die Struktur abgebildet, in der das Objekt im KuLaDig-Datenbestand eingebunden ist.

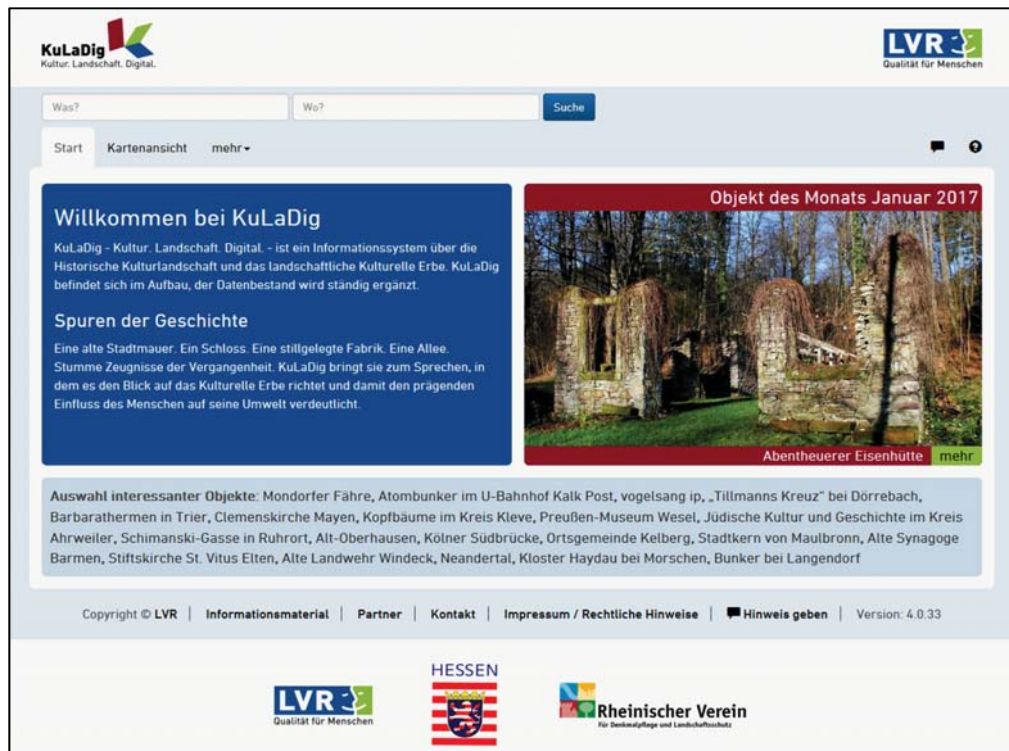


Abb. 1: Startbildschirm der Web-Anwendung KuLaDig

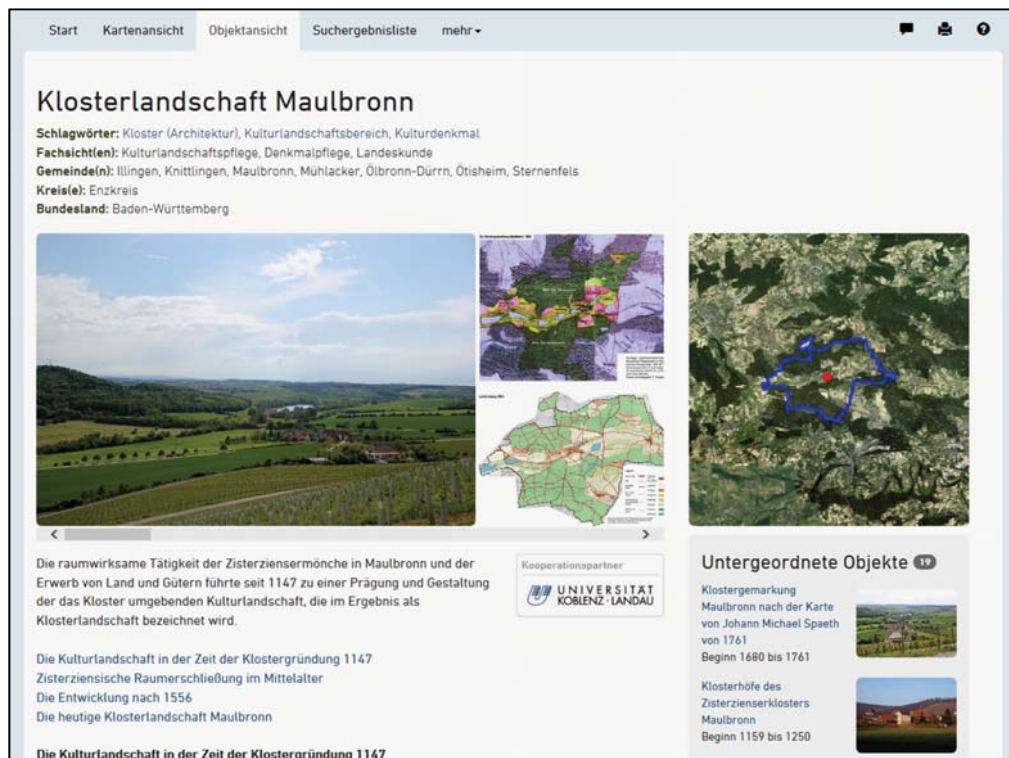


Abb. 2: Ausschnitt der Objektansicht eines KuLaDig-Eintrags

The screenshot displays the 'Maske der Expertensuche in KuLaDig' interface. It is organized into several sections:

- Suche in Listefeldern:** Includes a search criterion dropdown (currently 'Erfassungsmaßstab') and a search text input field. A search logic selector is set to 'ODER'.
- Suche in Datumsfeldern:** Features a search criterion dropdown (currently 'Erterfassung'), 'von' and 'bis' date input fields, and a search logic selector set to 'ODER'.
- Suche nach Schlagwörtern:** Contains a search concept dropdown, an 'Übernehmen in Liste der Suchbegriffe' button, and a list of search concepts. A search logic selector is set to 'ODER'. A 'Liste leeren' button is also present.
- Suche im Zeitraum:** Includes 'Jahr (von)' and 'Jahr (bis)' input fields, a search logic selector set to 'ODER', and checkboxes for 'Beginn im Zeitraum' and 'Ende im Zeitraum'. An 'Epochenauswahl' dropdown is set to 'Geologische Zeitschnitte'.
- Suche mit räumlicher Beschränkung:** Offers radio buttons for 'Suche nach Orten' (selected) and 'Suche in Karte', with a 'Wo?' input field.

Abb. 3: Maske der Expertensuche in KuLaDig

KuLaDig bietet die Möglichkeit, Objekte zueinander in Beziehung zu setzen (Abb. 4). Es lassen sich definierte Bezüge zwischen Objekten herstellen. Diese sind in hierarchische Bezüge („gehört zu“ / „ist Bestandteil von“ bzw. „hat Bestandteile“ / „setzt sich zusammen aus“) und in sogenannte Verwandtschaftsbeziehungen untergliedert. Letztere können inhaltlich ähnliche Objekte gruppieren, z.B. Naturdenkmale in einem bestimmten Landkreis, Objekte der jüdischen Kultur in einer bestimmten Stadt oder Wassermühlen entlang eines Flusses (Abb. 4). Es wird ein Phänomen oder das Gemeinsame einer Gruppe von Objekten beschrieben und die zugehörigen Einzelobjekte werden als verwandte Objekte gelistet.

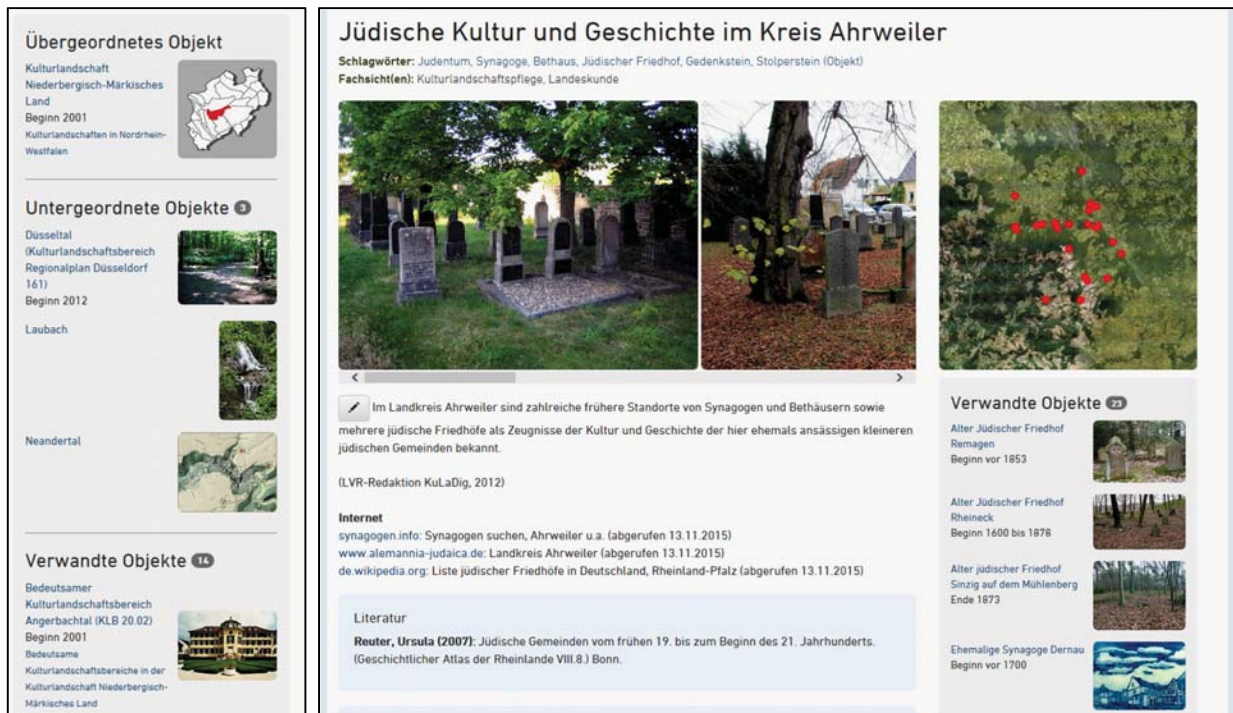


Abb. 4: Möglichkeiten zur Strukturbildung in KuLaDig (links) und Objektansicht einer Gruppe verwandter Objekte (rechts)

Die hierarchischen Bezüge orientieren sich an Planungshierarchien und an räumlich-funktionalen Gegebenheiten, etwa der Zugehörigkeit zu einer Siedlung oder einem Kloster, einem bestimmten Landschaftsausschnitt oder einer Struktur in der Kulturlandschaft. Da jeder Objekteintrag eine feste URL aufweist, lassen sich Objekteinträge individuell adressieren; dies ist sowohl im objektbeschreibenden Text im System KuLaDig selbst möglich (im Sinne von „siehe auch“) als auch auf jeder beliebigen Website. Hieraus ergibt sich die Option, eine flächendeckende Beschreibung der Kulturlandschaften eines Bundeslandes ebenso im System vorzuhalten, wie die Beschreibung der einzelnen Grenzsteine eines ehemaligen Klosterbezirks oder die eines einzelnen Wegekreuzes. Kartenfunktionalitäten stehen in KuLaDig an drei Stellen in ähnlichem Umfang zur Verfügung: bei der bildschirmfüllenden Kartenansicht (Abb. 5 links), der Expertensuche sowie bei der Editierfunktion in einem etwas kleineren Fenster. Die Kartendarstellung generiert sich aus übereinanderliegenden Layern. Als Basislayer werden Luftbilder, historische Karten und aktuelle Kartendarstellungen als WMS-Dienste eingebunden; Datenlieferanten sind die Landesvermessungsämter, das Bundesamt für Geodäsie und Kartographie, private Dienstleister, die öffentliche Vermessungsdaten aufbereiten, und die Open Street-Map-Community. Darüber folgen Verwaltungsgrenzen und Blattschnitte, weitere externe Fachdarstellungen, wie z.B. Schutzgebiete, und schließlich die aus dem KuLaDig-Datenbestand generierten Layer. Diese lassen sich z.B. nach Themen (analog zur Landnutzung), nach den bereits erwähnten Fachsichten sowie nach dem Planungsbezug unterscheiden und können sowohl gruppiert als auch einzeln zur Anzeige ausgewählt werden. Per Schieberegler sind Überblendungen möglich. Die Steuerung der Anzeige erfolgt über eine Kartensteuerung (Abb. 5 rechts).

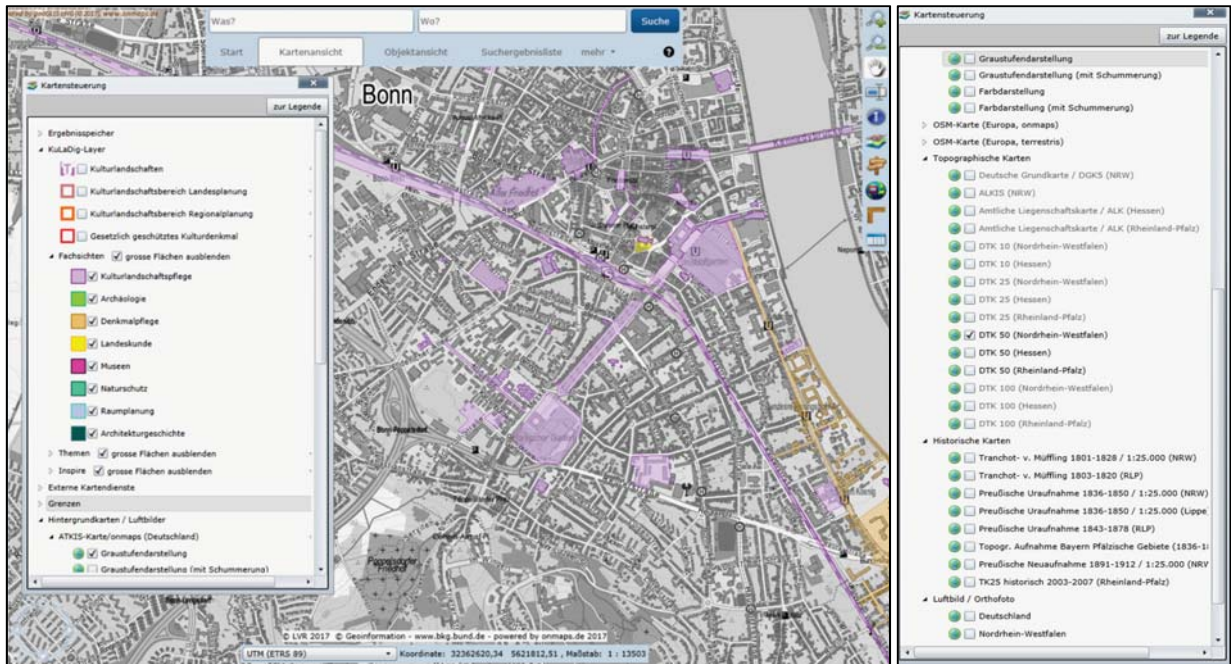


Abb. 5: Kartenansicht der Anwendung KuLaDig (links) und Ausschnitt der Kartensteuerung der Anwendung KuLaDig (rechts)

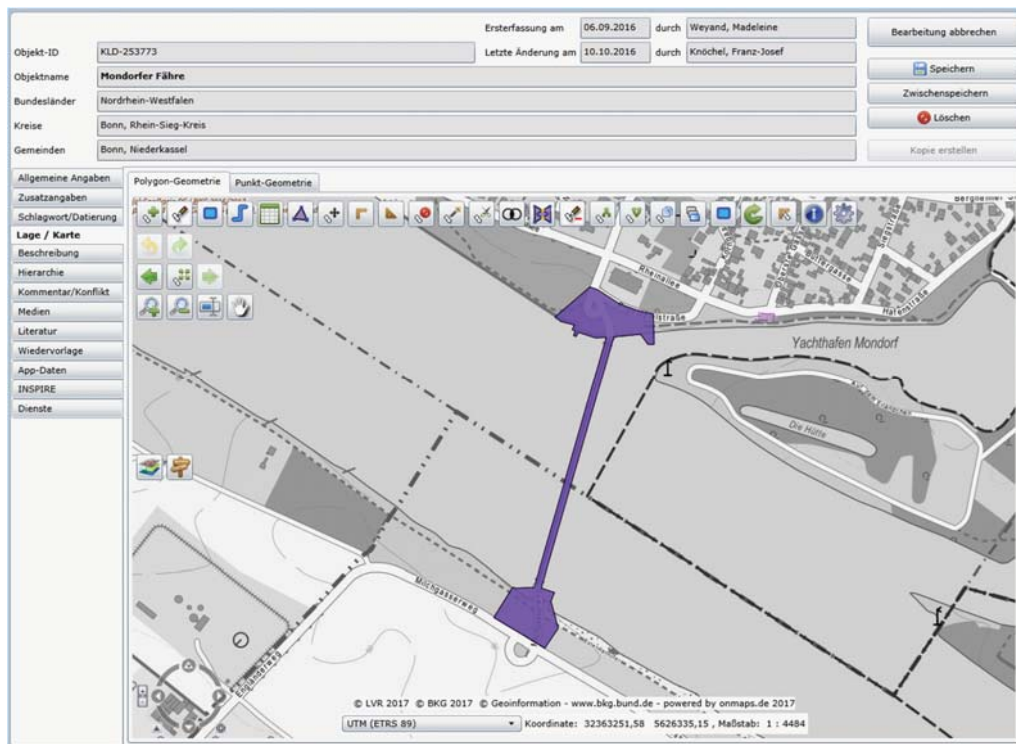


Abb. 6: Ausschnitt aus der Editierfunktionalität von KuLaDig

Die Editierfunktionen sind, entsprechend der dargestellten Such- und Anzeigemöglichkeiten, vielfältig (Abb. 6). Weiterhin ist ein Thesaurus-Dialog implementiert, der sowohl beim Editieren als auch bei der Expertensuche zur Verfügung steht. Die Vokabulardatenbank LVR-Wortnetz Kultur auf Basis der Software digiCULT.xTree der digiCULT-Verbund eG in Kiel, eingebunden per Web-Dienst, liefert zu einem ausgewählten Begriff Ober- und Unterbegriffe, assoziierte Begriffe, für den ausgewählten Begriff eine Definition, eine ergänzende Erläuterung sowie einen Verwendungshinweis, sofern vorhanden. Einen Einblick in die Vokabulardatenbank ermöglicht der bereits erwähnte Public-Viewer, der aus der Objektansicht von KuLaDig aufrufbar ist. Eine in KuLaDig integrierte Literaturdatenbank liefert Literaturangaben, die manuell oder über eine Schnittstelle aus einem der großen Verbundkataloge erfasst bzw. übernommen werden können.

Öffentlich sichtbar werden KuLaDig-Einträge erst dann, wenn ein entsprechend berechtigtes Mitglied der KuLaDig-Redaktion die Freigabe dazu erteilt. Bis dahin sind Einträge nur für Nutzerinnen und Nutzer mit einer entsprechenden Zugangsberechtigung intern sichtbar.

3 Erfassung der Kulturlandschaft

Die heutige Kulturlandschaft ist das Ergebnis regional unterschiedlicher Landnutzungs- und Siedlungsformen. Sie ist dabei einer steten Entwicklung und Veränderung unterworfen. Die Methode der Kulturlandschaftsanalyse dient dazu, die Kulturlandschaftsentwicklung fassbar zu machen, sie kartographisch auszuarbeiten und darzustellen. Ziel der Kulturlandschaftsanalyse ist die Erhebung und Bewertung kulturlandschaftlich prägender Strukturen und Objekte. Sie umfasst folgende Schritte:

1. Erstellung von Querschnittskarten auf Grundlage der vorliegenden Landesaufnahmen seit dem frühen 19. Jahrhundert für die Darstellung der Landnutzung der gewählten Zeitschnitte
2. Erstellung einer chronologischen Kulturlandschaftswandelkarte
3. Erstellung einer Karte der historischen Kulturlandschaftselemente und –strukturen
4. Erstellung einer Strukturkarte
5. Erstellung einer Bewertungskarte

Diese Arbeitsschritte sind aufeinander abgestimmt und haben sich in vielen Projekten bewährt. Die Kulturlandschaftsentwicklung wird vor allem anhand der gewählten Landnutzungsstadien vergangener Perioden, auf Grundlage vorhandener Altkarten als Querschnittskarten und mit dem chronologischen Kulturlandschaftswandel als Längsschnittkarte auf der aktuellen topographischen Karte dargestellt (Abb. 7). Der Betrachtungszeitraum kann für die zu untersuchende Region bis ins 18. Jahrhundert zurückreichen. Den Ausgangspunkt bilden die ältesten vorliegenden, präzisen Landesaufnahmen, die meistens aus dem ersten Viertel des 19. Jahrhunderts stammen. Vereinzelt existieren präzise Landesaufnahmen aus dem 18. Jahrhundert, beispielsweise die klevischen Katasterkarten (1732-1737) aus den heutigen niederrheinischen Kreisen Kleve und Wesel. Auf Basis

dieser Karten konnte in verschiedenen Projekten die Kulturlandschaftsentwicklung der Region Kleve seit 1732 erarbeitet werden.³

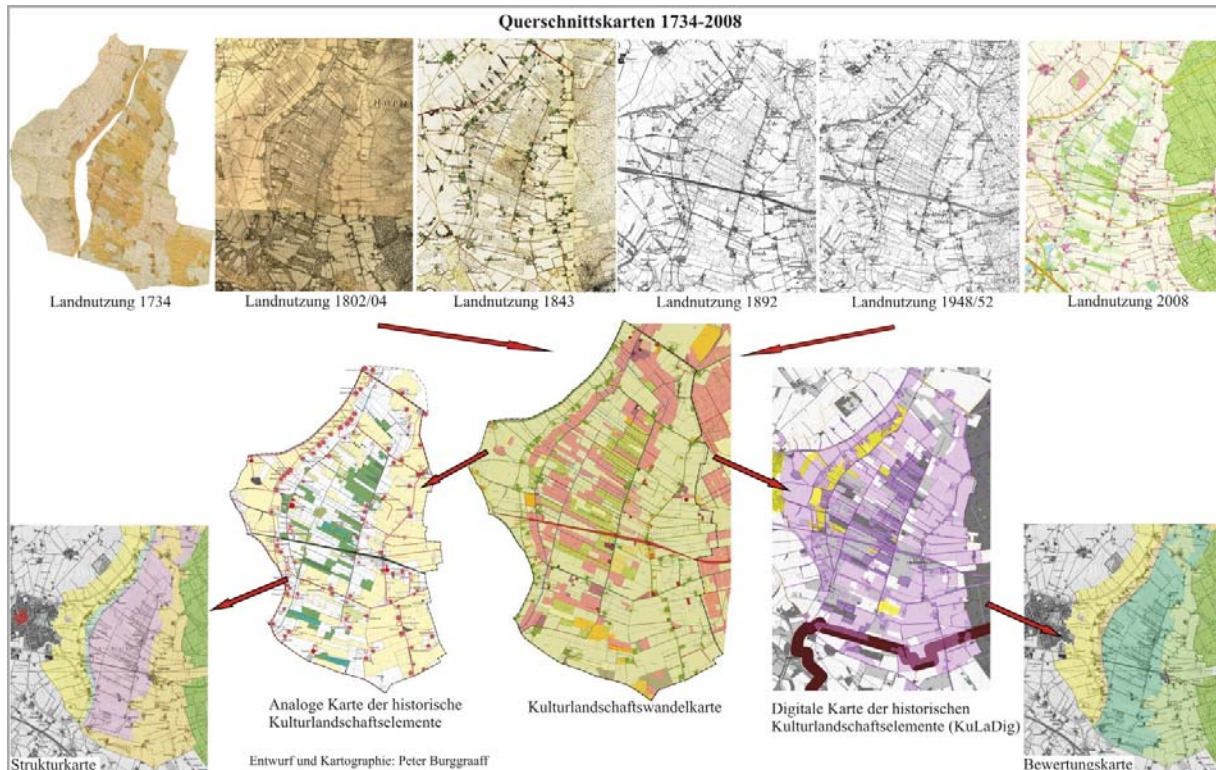


Abb. 7: Schematische Darstellung der Kulturlandschaftsanalyse im Projektgebiet Uedemerbruch (Burggraaff 2013 und 2016; Kartengrundlagen: Landesarchiv Nordrhein-Westfalen und Bezirksregierung Köln, Abt. 7: Geobasis NRW)

Die gewählten Zeitschnitte für die Querschnittskarten und der Wandelkarte hängen ebenfalls mit älteren Landesaufnahmen und älteren Ausgaben der topographischen Karten zusammen. Im Rheinland handelt es sich um die linksrheinische französische Tranchot- (1801-1813) und die rechtsrheinische preußische Von Müfflingkarte (1814-1828), die preußische Uraufnahme (1842-1850), die preußische Neuaufnahme (1893-1901) sowie die nachfolgenden Fortschreibungen beziehungsweise umfassenden Bearbeitungen der Blätter der Neuaufnahme und der aktuellen topographischen Karten (TK 25).

In den Querschnittskarten sind, neben den topographischen Darstellungen, die jeweiligen Landnutzungsstrukturen und konstituierenden Merkmale zu beachten. Eine weitere Rückschreibung der Kulturlandschaft bis ins Mittelalter ist durch das Fehlen von zeitgenössischen, flächendeckenden

³ Das klevische Kataster ist im Rahmen einer Steuerreform auf Anordnung des preußischen Königs Wilhelm I. seit 1731 erstellt worden. Die Kartierungen sind von Ingenieur-Offizieren nach den damaligen modernsten Methoden und Techniken mit Dreieckvermessung durchgeführt worden. Dies bedeutet, dass die Karten für diese Zeit sehr genau sind und deswegen als Ausgangspunkt für die Kulturlandschaftswandelkartierung dienen (AYMANS et al. 1988).

und genauen Altkarten sehr problematisch und kann nur auf der Grundlage von Archivquellen lückenhaft erfolgen.

Für die chronologische Darstellung der Kulturlandschaftlichen Entwicklung seit dem Beginn des 19. Jahrhunderts wurde die Methode der Kulturlandschaftswandelkartierung entwickelt. Sie stellt die Entwicklung durch Vergleich und Interpretation der Landnutzungskarten (Querschnittskarten) chronologisch dar, indem Kulturlandschaftselemente nach ihrer erstmaligen Darstellung in Altkarten auf der Grundlage der aktuellen Ausgabe der topographischen Karte kartiert werden. Im Drittmittelprojekt „Uedemerbruch. Naturschutz und Kulturgeschichte – gemeinsam erleben und Entwickeln“ (2009-2012) wurde eine Wandelkarte für das Projektgebiet Uedemerbruch erstellt. Sie gewährt Einblicke in die Datierung der Kulturlandschaftselemente und verdeutlicht die Chronologie der erfassten Elemente seit dem Jahr 1734 auf Basis der klevischen Katasterkarte. Die Wandelkarte beruht auf Querschnitten der Jahre 1734, 1802/1804, 1843, 1894, 1948/1950 und 2008. Das Besondere dieser Kartierungsmethode ist der unmittelbare Bezug auf die Topographie. Eine Kulturlandschaftswandelkarte lässt sich allerdings nicht automatisch digital generieren, sondern erst nach der quellenkritischen Auswertung der vorliegenden Altkarten.

Während Veränderungen der Kulturlandschaft im Kartenbild visuell erfasst werden, müssen Prozesse und Faktoren, die zur Herausbildung des Landschaftsbildes mit seinen erhaltenen Elementen und Strukturen führten, beschrieben und analysiert werden. Veränderungen und Persistenz sind in der Wandelkarte durch die chronologisch angeordneten Farben deutlich ablesbar. Sie hilft bei der Einschätzung von Landschaften hinsichtlich ihrer strukturellen historischen Bewahrung oder zeitgenössischen Überformung. Abbildung 8 stellt die methodische Vorgehensweise bei der Erstellung der Kulturlandschaftswandelkarte dar. Die Festlegung der Epochen mit den zugeordneten Farben hängt mit den damals neu erstellten Landesaufnahmen und der umfassenden Berichtigung bzw. umfassenden Bearbeitungen der topographischen Karten zusammen. Auf der Grundlage der Wandelkarte und Geländearbeit werden die überlieferten historischen Kulturlandschaftselemente gesichtet und kartiert. Seit 2008 werden die Ergebnisse dieser Kartierungen in KuLaDig textlich und kartographisch erfasst.

In der Strukturkarte werden die kulturlandschaftlichen Hauptstrukturen anschließend zusammengefasst. Dabei spielt neben der Landnutzungsentwicklung, die Siedlungsentwicklung und die infrastrukturelle Entwicklung, in Zusammenhang mit den naturräumlichen Rahmenbedingungen, eine bedeutende Rolle. In der Bewertungskarte werden die unterschiedlichen Kulturlandschaftsbereiche nach verschiedenen Kriterien wie historischem Wert, künstlerisch-gestalterischem Wert, Erhaltungswert, Seltenheitswert, regionaltypischem Wert, Wert der räumlichen Zusammenhänge und Beziehungen, Nutzungswert und Schutzstatus bewertet.

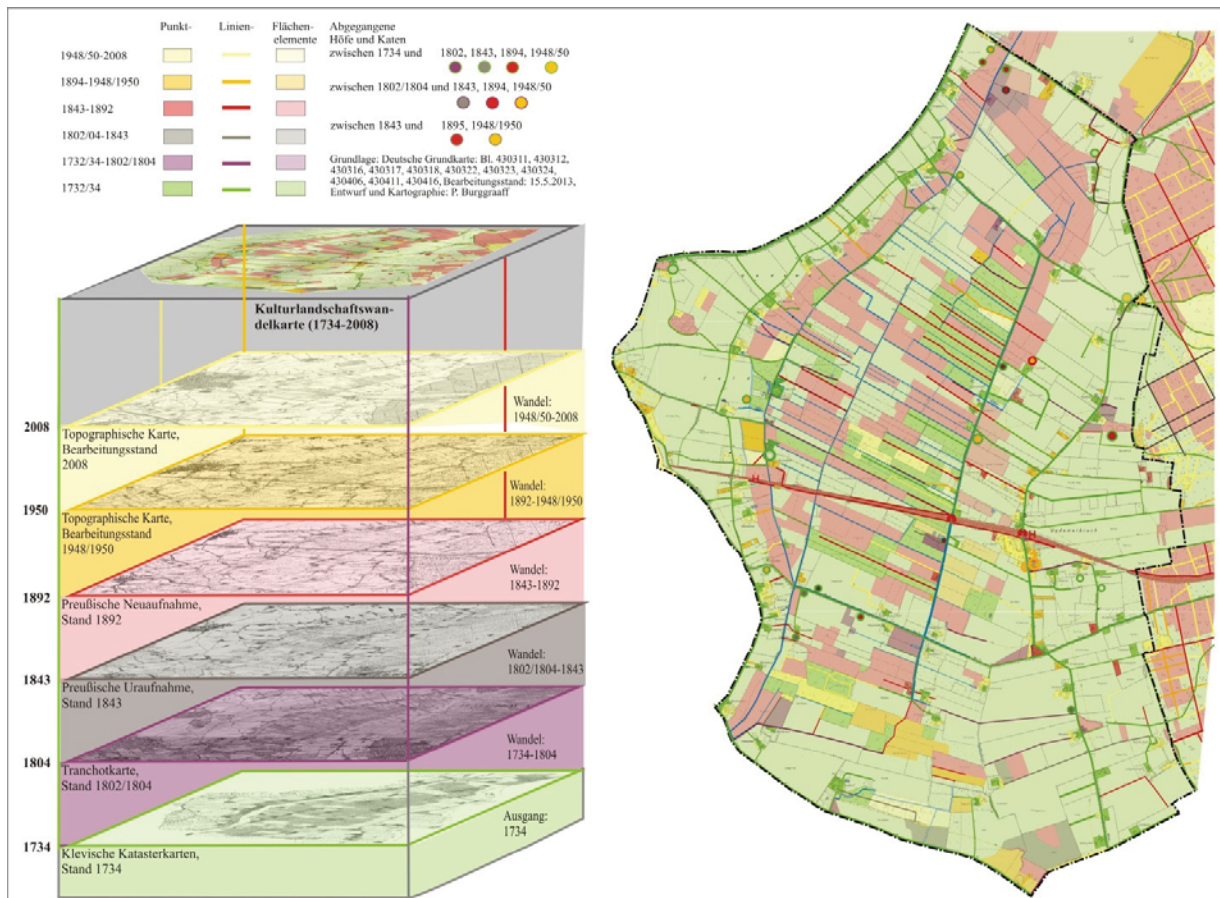


Abb. 8: Schematische Darstellung des Kulturlandschaftswandels und Kulturlandschaftswandelkarte 1734 bis 2008 (Kartengrundlagen: Deutsche Grundkarte, Bezirksregierung Köln, Abt. 7: Geobasis NRW)

4 Datenbestände in KuLaDig und Kooperationen

Der LVR ist auf die Zusammenarbeit mit externen Projektpartnern angewiesen, um das System KuLaDig mit Informationen zu versehen. Zwar nimmt der LVR auch selbst Objekte in das System auf, aufgrund der Vielzahl an Kulturlandschaftselementen und Fachrichtungen muss er jedoch mit Projektpartnern zusammenarbeiten. Diese Projektpartner beteiligen sich aktiv und pflegen nicht nur neue Objekte in das System ein, sondern betrachten bestehende Einträge kritisch und korrigieren und erweitern diese. Projektpartner sind einerseits Fachleute, die ihr Wissen in einem bestimmten Themengebiet der Öffentlichkeit zugänglich machen möchten, sowie Personen, die ortsbezogene Kulturlandschaftselemente in KuLaDig einpflegen. Zusätzlich arbeitet der LVR auf Basis vertraglich festgelegter Kooperationsvereinbarungen mit einigen Gebietskörperschaften und Universitäten zusammen. Diese Zusammenarbeit ist sehr fruchtbar, ein Großteil der vorhandenen KuLaDig-Beiträge stammt aus diesen Kooperationen.

Der LVR ist mit KuLaDig bundesweit in Nordrhein-Westfalen am stärksten aktiv. Von über 8300 freigegebenen Einträgen stammen knapp 6500 aus diesem Bundesland. Innerhalb von Nordrhein-

Westfalen bildet der Verwaltungsbereich des LVR ein Schwerpunktgebiet. In Nordrhein-Westfalen arbeitet der LVR darüber hinaus im Rahmen von Kooperationsvereinbarungen mit einer Vielzahl von Partnern zusammen. Dazu zählen unter anderem der Rheinische Mühlenverband, der Rheinische Verein für Denkmalpflege und Landschaftsschutz (RVDL), verschiedene Biostationen und das Geographische Institut der Universität Bonn.

Seit einigen Jahren ist KuLaDig über Kooperationsvereinbarungen mit Institutionen in weiteren Bundesländern aktiv. Im Rahmen des durch die Deutsche Bundesstiftung Umwelt (DBU) geförderten Projektes „Kulturlandschaft Kommunal“ wurden in Zusammenarbeit mit dem Landesamt für Denkmalpflege des Landes Hessen beispielsweise im Rheingau-Taunus-Kreis über 140 Kulturlandschaftselemente erfasst. Seit dem Jahr 2012 besteht zwischen der Abteilung Geographie der Universität Koblenz-Landau und dem LVR ein Kooperationsvertrag, der die Zusammenarbeit der beiden Institutionen regelt. KuLaDig hat damit einen rheinland-pfälzischen Projektpartner und wird seither dort rege angewandt. Die Universität ist bisher der einzige Projektpartner von KuLaDig in Rheinland-Pfalz, weshalb ein Großteil der über 1300 rheinland-pfälzischen KulaDig-Einträge dieser Kooperation entstammen. KuLaDig wird bisher insbesondere als Teil von Lehrveranstaltungen sowie im Rahmen von Abschlussarbeiten eingesetzt. Die Aufnahmen erfolgen in Rheinland-Pfalz anhand regionaler Schwerpunktgebiete, eine Ortschaft oder Region wird ausgewählt und anschließend kulturlandschaftlich untersucht. In den ersten Jahren der Kooperation konzentrierten sich die KuLaDig-Aktivitäten der Abteilung Geographie vor allem auf den nördlichen Bereich des Bundeslandes. Städte wie Koblenz, Andernach und Neuwied erfuhren dabei umfassende Kulturlandschaftsaufnahmen. Ein sehr umfangreiches Projekt läuft seit dem Jahr 2012 in der Verbandsgemeinde Kelberg im Kreis Vulkaneifel. Der aktuelle Erhebungsstand liegt dort bei derzeit über 140 Objekten. Der Bestand enthält unter anderem Informationen der Kelberger Geschichtsstraße, die vor einigen Jahren aus touristischen und heimatkundlichen Zwecken angelegt wurde. Weitere Schwerpunktgebiete im nördlichen Rheinland-Pfalz waren bisher die Moselorte Zell, Bullay, Alf und Pünderich, für die bisher mehr als 60 Objekte freigegeben sind. Seit dem Jahr 2015 dehnen sich die KuLaDig-Aktivitäten auf weitere Regionen von Rheinland-Pfalz aus. Hier ist zunächst das Mittelrheintal zu nennen, wo insbesondere St. Goar, St. Goarshausen mit der Loreley sowie Boppard untersucht wurden. Derzeit sind in diesen Orten über 60 Objekte freigegeben. Die Stadt Bad Kreuznach stellt ein weiteres Schwerpunktgebiet dar, für das derzeit Kulturlandschaftserfassungen laufen.

Seit dem Jahr 2015 besteht eine Kooperation der Universität Koblenz-Landau mit dem im selben Jahr gegründeten Nationalpark Hunsrück-Hochwald. Die Nationalparkleitung ist sehr an der kulturlandschaftlichen Erforschung des Nationalparks und der Nationalparkregion interessiert. Im Rahmen dieser Zusammenarbeit werden insbesondere in der inmitten des Nationalparks gelegenen Ortschaft Börfink Kulturlandschaftselemente aufgenommen. Auch die durch die Nationalparkgründung unter Schutz gestellten forstlich stark überprägten Waldbereiche werden seitdem kulturlandschaftlich erfasst. In der Nationalparkregion wurde bisher insbesondere die Orte Abentheuer und Rinzenberg in die Untersuchungen mit einbezogen. In Abentheuer konnten auf dem Gelände einer ehemaligen Eisenhütte umfassende Untersuchungen durchgeführt werden. In enger Zusammenarbeit mit dem heutigen Eigentümer war es möglich, einen Hauptbeitrag zur Hütte zu verfassen, der 24 Unterobjekte sowie sechs Verwandtschaftsobjekte enthält. Dieser Beitrag war im Ja-

nuar 2017 KuLaDig-Objekt des Monats und wurde damit prominent auf der Startseite der Webanwendung angezeigt. Die derzeitige Zwischenbilanz dieser Kooperation zeigt sich in über 100 freigegebenen Objekten für die Verbandsgemeinde Birkenfeld sowie über 10 freigegebenen Objekten für die Verbandsgemeinde Thalfang am Erbeskopf. Die Zusammenarbeit mit dem Nationalpark Hunsrück-Hochwald wird, in enger Kooperation mit dem LVR sowie regionalen Ortsgemeinden und weiteren Gebietskörperschaften, weiter fortgeführt.

Im rheinland-pfälzischen Teil des Hunsrücks existieren über die Universität Koblenz-Landau weitere KuLaDig-Kooperationen. Im östlichen Hunsrück trat im Jahr 2015 die im Landkreis Bad Kreuznach gelegene Ortsgemeinde Dörrebach an die Universität heran, mit der Bitte um eine kurlandschaftliche Aufnahme des Ortes. Die Erfassung der circa 700 Einwohner zählenden Gemeinde wird im Jahr 2017 abgeschlossen, bisher wurden über 70 Objekte aufgenommen, von denen knapp über 50 freigegeben sind. Die erhobenen Daten sollen für einen bestehenden Geschichts-Wanderweg genutzt werden. An erwähnenswerten Objekten ermöglichen auf Hinweisschilder aufgeklebte QR-Codes die direkte Verbindung des jeweiligen Objektes mit der Website von KuLaDig. Die Codes können per Smartphone eingescannt werden. Dadurch wird der Nutzer direkt auf die Website geleitet, wo er weitere Informationen zum Objekt erhält sowie verwandte und übergeordnete Objekte betrachten kann. Auf Basis der durch die Universität getätigten Aufnahmen werden in Dörrebach weitere, themenbezogene Wanderwege durch den Gemeinderat und die zuständigen Ausschüsse ausgewiesen. Die Kulturlandschaftsaufnahmen in Dörrebach führten dazu, dass an die Ortsgemeinde angrenzende Gemeinden bei der Abteilung Geographie um vergleichbare Maßnahmen auch in ihren Ortschaften bitten. Aufgrund dessen wird im Jahr 2017 die Ortsgemeinde Seibersbach im Rahmen verschiedener Semesterarbeiten kurlandschaftlich aufgenommen. In Rheinland-Pfalz soll KuLaDig in Zukunft eine größere Rolle spielen, da ortsbezogene Kulturlandschaftsaufnahmen auch in südlicheren Landesteilen geplant sind. Eine KuLaDig-Analyse der Stadt Wachenheim an der Weinstraße ist für das Jahr 2017 angesetzt und für Teile der Südwestpfalz existieren ebenfalls Projektplanungen.

Die Aktivitäten von KuLaDig haben seit dem Jahr 2015 auch das Saarland erreicht. Neben mehreren Abschlussarbeiten (Gemeinde Rehlingen-Siersburg) werden dort auf Basis der Zusammenarbeit mit dem Nationalpark Hunsrück-Hochwald Kulturlandschaftsaufnahmen in der Gemeinde Nonnweiler durchgeführt. Die Anzahl redaktionell freigegebener Objekte beziffert sich Saarland weit auf circa 60, wobei sich der weitaus größte Teil in den Gemeinden Nonnweiler und Rehlingen-Siersburg befindet.

Aufgrund der großen Nachfrage nach Zusammenarbeit durch verschiedene Gebietskörperschaften mit der KuLaDig-Redaktion bleiben die Hauptaktivitäten des Programms zunächst auf Nordrhein-Westfalen, Hessen, Rheinland-Pfalz und das Saarland beschränkt. Eine verstärkte Ausdehnung in andere Bundesländer wird angestrebt, ist jedoch für die nahe Zukunft nicht geplant. KuLaDig beschränkt sich jedoch nicht mehr nur auf Deutschland als Kerngebiet. Seit Anfang 2016 wird das Programm im Rahmen eines länderüberschreitenden Projektes „Baalsbrugger Mühle“ zwischen dem LVR und der niederländischen Provinz Zuid-Limburg angewandt. Innerhalb dieses Projektes wird die Kulturlandschaft des historischen Dreiecks um die Burg Rode in Herzogenrath, das Kloster Rolduc und die Baalsbrugger Mühle in Kerkrade erforscht und teilweise in KuLaDig eingepflegt. Die Übersetzung des Programms in die niederländische Sprache und das Integrieren niederländischer Texte ist aufgrund dessen fester Bestandteil der künftigen Programmentwicklung.

Zusätzlich besteht von Seiten der Deutschsprachigen Gemeinschaft in Belgien eine Nachfrage nach Zusammenarbeit.

5 Literatur

- AYMANS, G., BURGGRAAFF, P. & JANSEN, W., 1988: De regio Gennep aan de ketting. Gennep, Heijen, Milsbeek, Oeffelt, Ottersum, Ven-Zelderheide in kadasterkaarten (1731–1732), Venlo.
- BURGGRAAFF, P., 2013: Uedemerbruch. Naturschutz und Kulturgeschichte gemeinsam erleben und entwickeln. Ein Projekt des Naturschutzzentrums im Kreis Kleve in Kooperation mit der Universität Koblenz (Institut für Integrierte Naturwissenschaften, Geographie), Koblenz.
- BURGGRAAFF, P., 2016: Uedemerbruch. Naturschutz und Kulturgeschichte gemeinsam erleben und entwickeln. In: Heimat- und Verkehrsverein Uedem e.V. (Hrsg.): Uedem – ein Ort reich an Geschichte, 93-128.
- BURGGRAAFF, P. & KLEEFELD, K.-D., 2010: Vom Kataster zum Informationssystem. Digitale Inventarisierung historischer Kulturlandschaften. Historische Kulturlandschaft und Denkmalpflege. Definition, Abgrenzung, Bewertung, Elemente, Umgang, FRANZ, B. & HUBEL, A. (Hrsg.), 183-195.
- BURGGRAAFF, P. & KLEEFELD, K.-D., 2011: Historisch-geographische Forschung im digitalen kulturlandschaftlichen Informationssystem KuLaDig. Historisch geographische Konzepte und Fragestellungen. Gestern – heute – Morgen, HAFFKE, J., KLEEFELD, K.-D. & SCHENK, W. (Hrsg.), 185-206.

Die automatisierte Auswertung von LiDAR-Daten in der Archäologie

M. FABIAN MEYER¹, INGO PFEFFER² & CARSTEN JÜRGENS¹

Zusammenfassung: Während die LiDAR-Technologie in den letzten Jahren einige Probleme der archäologischen Prospektion gelöst hat, steckt die systematische und effiziente Auswertung der LiDAR-Daten immer noch in den Anfängen.

In Kooperation mit der LWL-Archäologie für Westfalen wurde daher im Rahmen einer Masterarbeit am Geographischen Institut der Ruhr-Universität Bochum eine Möglichkeit gesucht, die großen LiDAR-Datenmengen effizient auszuwerten. Bei der entwickelten Methode kommt u.a. die Software eCognition zum Einsatz, die automatisch nach Bodendenkmälern sucht. Um die anschließende Interpretation zu erleichtern, sind die Verdachtsflächen danach sortiert, wie sehr sie ihrem jeweiligen Idealtyp entsprechen. Auf diese Weise können die interessantesten Flächen zuerst begutachtet und uninteressante zunächst vernachlässigt werden.

1 Einleitung – vom Luftbild zum DGM

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. Bei der dafür nötigen, flächendeckenden Erfassung spielt in einer ganzen Reihe von Prospektionsmethoden die Luftbildarchäologie seit langem eine wichtige Rolle. Sie ist durch die große Distanz zur Erdoberfläche in der Lage, sowohl einzelne Bodendenkmäler als auch vollständige (historische) Landschaften zu erfassen. Damit schafft sie die Grundlage für die wortwörtlich tiefergehende Prospektion, z.B. mittels Georadar oder schließlich einer Grabung.

Klassischen (Schräg-)Luftbildern sind dabei gewisse Grenzen gesetzt. So hat vor allem der Reifegrad der Vegetation entscheidenden Einfluss darauf, ob ein Bodendenkmal erkannt wird oder nicht. Auch andere Aspekte wie die Tages- und Jahreszeit sind entscheidend. Geht es schließlich an die Erkundung von Waldgebieten, stoßen Luftbilder ebenso wie Feldbegehungen noch schneller an ihre Grenzen und nur in Ausnahmefällen können während der vegetationslosen Zeit Strukturen am Waldboden fotografiert werden. Da Bodendenkmäler jedoch unter der relieferhaltenden Wirkung von Vegetation relativ gut erhalten bleiben und zudem seltener durch landwirtschaftliches Gerät überformt werden, sind gerade diese Bereiche von großem Interesse.

Dort hat die LiDAR-Technologie in den letzten 15 Jahren die Werkzeugpalette der archäologischen Prospektion ergänzt und einen Beitrag dazu geleistet, die Grenzen ein Stück weit ‚in den Wald hinein‘ zu verschieben, indem sie kaum von Vegetation beeinträchtigt wird und es ermöglicht, digitale Geländemodelle (DGM) auch von Waldgebieten zu erstellen. Zumindest

¹ Ruhr-Universität Bochum, Geographisches Institut – AG Geomatik, Universitätsstraße 150 / Gebäude NA, 44780 Bochum, E-Mail: matthias.meyer@rub.de, carsten.juergens@rub.de

² LWL-Archäologie für Westfalen, Außenstelle Münster, An den Speichern 7, 48157 Münster, E-Mail: ingo.pfeffer@lwl.org

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.



Abb. 1: Landwehr in Dülmen in Luftbild und DGM; Quelle: Bezirksregierung Köln 2016 (Hrsg.): TIM-online. <http://www.tim-online.nrw.de>

2 Problemstellung – die manuelle Auswertung

Bei allem Erfolg leidet die Schummerungsdarstellung unter derselben Beleuchtungsproblematik wie Luftbilder, die Bodendenkmäler über Schattenwürfe preisgeben (Abb. 2). Deswegen wurden erfolgreich verschiedenste Visualisierungen entwickelt, die ohne eine simulierte Beleuchtungsquelle auskommen. Eine vergleichende Darstellung von Visualisierungen ist bei DONEUS (2013) zu finden. Am Ende stehen dennoch die manuelle Erfassung der Bodendenkmäler und eine Reihe von möglichen Fehlerquellen, angefangen von der Bildschirmdarstellung und -einstellung über das damit verbundene Übersehen von Verdachtsflächen bis hin zum großen Zeitaufwand. Letzteres wird einmal mehr deutlich, wenn man die enorme Zahl zu erwartender Verdachtsflächen berücksichtigt.

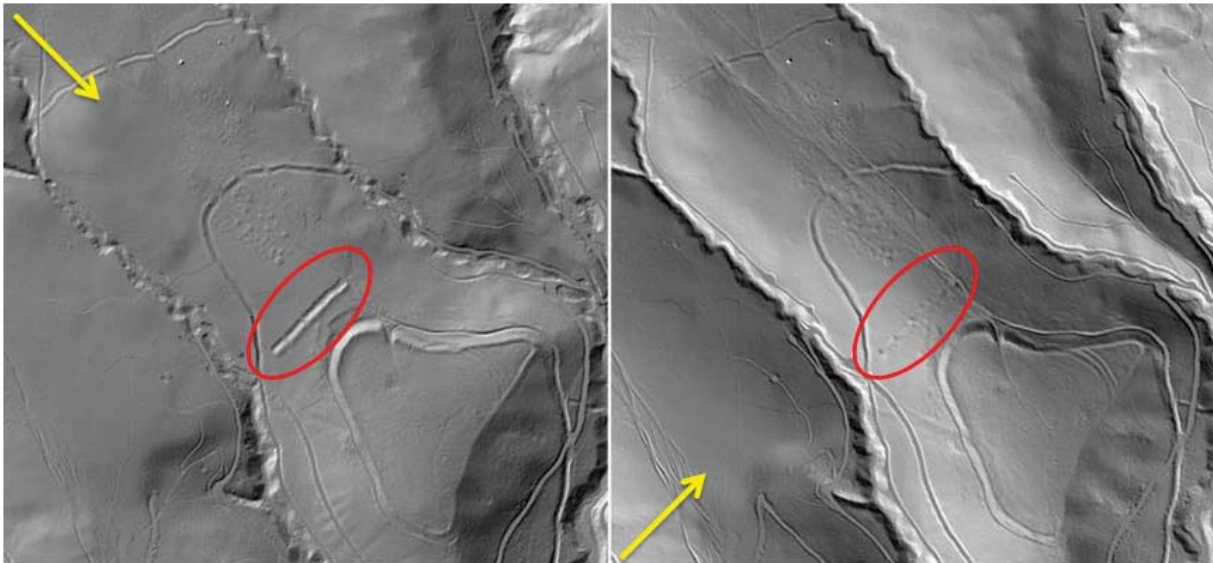


Abb. 2: Schattenproblematik beim schattierten Geländemodell als mögliche Fehlerquelle bei der manuellen Erfassung von Bodendenkmälern (DONEUS & BRIESE 2006: 103).

3 Methodik - Automatische Erfassung von Bodendenkmälern

Im Rahmen einer Masterarbeit am Geographischen Institut der Ruhr-Universität Bochum wurde in Kooperation mit der LWL-Archäologie eine Möglichkeit gesucht, die großen LiDAR-Datenmengen möglichst effizient auszuwerten und potentielle Bodendenkmäler weitestgehend automatisiert zu erfassen. Eine aktuelle Übersicht über ähnliche Projekte findet sich bei SEVARA et al. (2016). Die entwickelte Methode besteht aus drei Schritten, die jeweils vollautomatisch ablaufen. Die ersten beiden dienen der Datenvorbereitung und der dritte Schritt der eigentlichen Erfassung.

Als erstes entpackt eine *Batch*-Datei im Windows Explorer alle Archive mit den Rohdatensätzen, ändert die Dateierweiterungen und schreibt in jede Datei eine Kopfzeile mit Spaltenüberschriften. Dies funktioniert mit einer beliebigen Anzahl von Datensätzen, da mehrere Schleifen Verwendung finden und diese alle Dateien im gleichen Verzeichnis bearbeiten, die die Kriterien der jeweiligen Schleifen erfüllen (Abb. 3).

Dieser Schritt ist notwendig, damit das GIS-Tool die Koordinaten korrekt erkennt.

Als zweites berechnet ein eigens programmiertes *ArcGIS*-Tool die nötigen Geländemodelle und legt eine Ordnerhierarchie an, sodass die Daten (und später die Ergebnisse) nach den Untersuchungsgebieten sortiert vorliegen. Da sich Strukturen an Hängen mit *eCognition* nur sehr begrenzt in einem herkömmlichen DGM erfassen lassen, wird auf die *Difference Map* zurückgegriffen (HESSE 2010). Diese Visualisierung entsteht 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

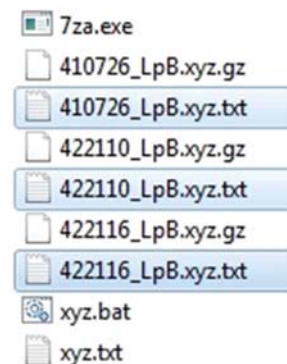


Abb. 3: Ergebnis der Batch-Datei (eigene Darstellung)

einem Tiefpassfilter geglättet, um es dann wieder vom originalen DGM zu subtrahieren. Die Differenz der beiden DGM wird dann visualisiert, wobei die Werte um null mäandrieren und die Amplitude von der Stärke des Tiefpassfilters abhängt (Abb. 4). Zu beachten ist, dass durch das Glätten v.a. an Geländekanten Pseudostrukturen entstehen, die die Klassifikation positiv und negativ beeinflussen können.

Im dritten Schritt findet schließlich die eigentliche Erfassung statt. *eCognition* arbeitet mit dem Konzept der objektorientierten Klassifikation (s.u.). Es bedient sich dabei der berechneten *Difference Map* und greift in Einzelfällen auf das herkömmliche DGM zurück. Am Ende der Erfassung liegen GIS-kompatible Verdachtsflächen vor, die der Interpretation bedürfen. Um diese zu erleichtern und zu beschleunigen, sind die Verdachtsflächen danach sortiert, wie sehr sie ihrem jeweiligen Idealtyp entsprechen. Auf diese Weise können die interessantesten Flächen zuerst begutachtet und uninteressante zunächst vernachlässigt werden. Der Workflow ist in Form eines *rulesets* gespeichert, das sich in neuen Gebieten mit wenig Aufwand anwenden lässt (Abb. 5). Somit entfällt das zeitaufwendige Erstellen der Prozesse, was schnelle Analysen neuer Gebiete ermöglicht.

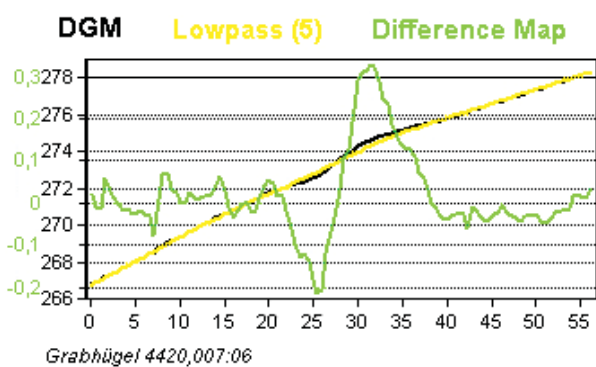


Abb. 4: Profil eines Grabhügels in DGM, geglättetem DGM und Difference Map (eigene Darstellung)

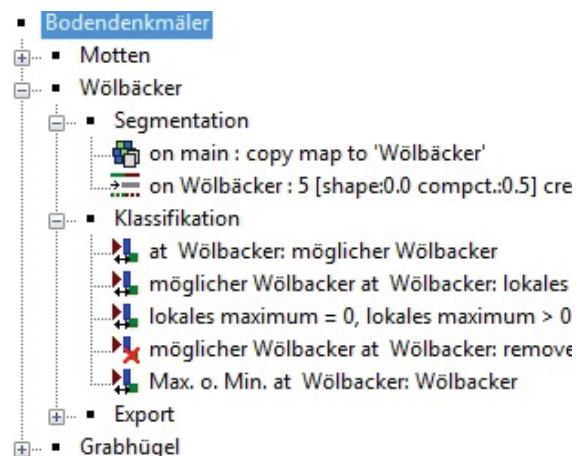


Abb. 5: Ruleset zur Suche nach Wölbäckern (eigene Darstellung)

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 - als Motten werden bestimmte Burgentypen bezeichnet, die

im Wesentlichen aus einem Turmhügel und einem umlaufenden Graben bestehen - und „lange, schmale, runde Objekte, die von höheren Bereichen umgeben sind und neben einem Turmhügel liegen“ die dazugehörigen Gräben. Passen die Eigenschaften eines Objektes zu denen einer Klasse, wird es ihr zugewiesen.

4 Ergebnisse

Das beschriebene Verfahren wurde anhand von Motten, Grabhügeln und Wölbäckern in Testgebieten in Nordwest- und Ostwestfalen entwickelt. Wölbäcker sind Überreste historischer Ackerfluren, die heute in Scharen parallel verlaufener Wölbungen erhalten sind. Zur Beurteilung der Qualität wurden Referenzdaten verwendet, wobei Treffer, die in diesen nicht vorhanden waren, als „falsche Treffer“ bezeichnet wurden – in Einzelfällen handelte es sich allerdings auch um bisher nicht gefundene Bodendenkmäler. Die Ergebnisse zeigen schließlich, dass alle drei Bodendenkmaltypen grundsätzlich erfassbar und Trefferquoten von 90-100%, je nach Testgebiet und Klasse, möglich sind.

Bei den Motten werden erfolgreich verschiedene Erhaltungszustände abgedeckt (Abb. 6). Der einzige falsche Treffer ist auf die verwendeten Parameter und/oder auf die *Difference Map* zurückzuführen und sieht außerdem einer Motte recht ähnlich. Durch die Verwendung der *Difference Map* treten darüber hinaus kleinere Probleme bei der Erfassung der Turmhügel auf, die durch die Ergänzung des herkömmlichen DGM umgangen werden.

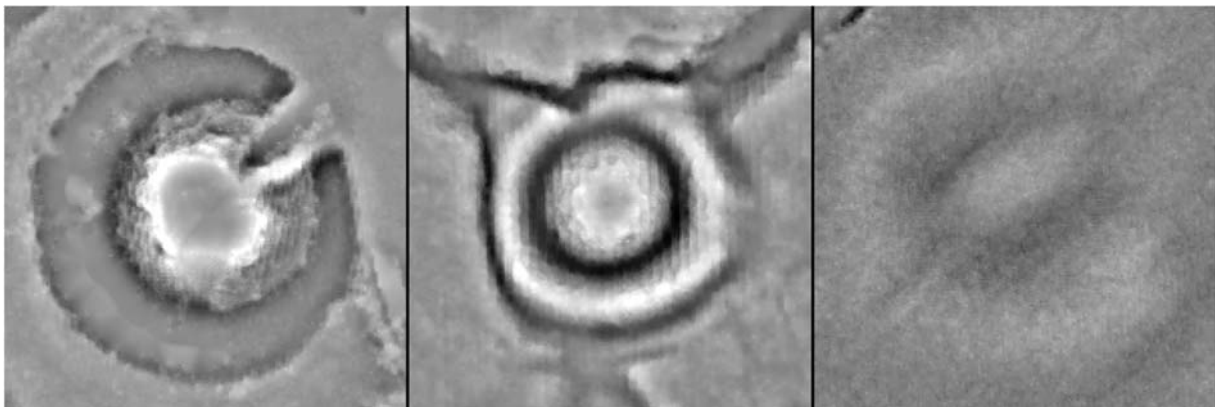


Abb. 6: Verschiedene Erhaltungszustände von Motten, die mit dem Verfahren gefunden werden können; eigene Darstellung

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. Nicht zuletzt ist auch auf die Nachteile der *Difference Map* zu verweisen. Trotz aller Schwierigkeiten wurden bisher unbekannte Grabhügel gefunden und in manchen Gebieten sehr gute Trefferquoten erzielt. (Abb. 7).

Bei den Wölbäckern ist das Ergebnis besser als erwartet. Gerade in flachem Gelände zeigt sich, dass die Klassifikation sehr gut funktioniert (Abb. 8). 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 Makroreliefs Strukturen entstehen, die in der *Difference Map* Wölbäckern sehr ähnlich sehen – die Verwendung des herkömmlichen DGM kommt im Gegensatz zu den Motten allerdings nicht in Frage. Die Qualität der Klassifikation hängt also von mehreren Faktoren ab, insbesondere dem Erosionsgrad der Strukturen, der Überprägung des Testgebietes durch den Menschen, den Klassifikationseinstellungen in *eCognition* und schließlich auch der Datenqualität.

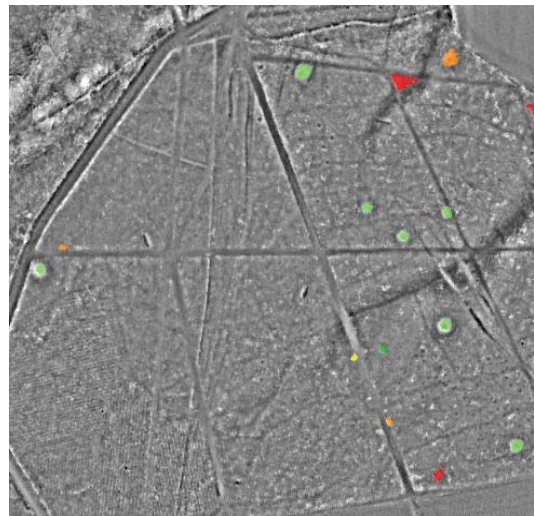


Abb. 7: Ausschnitt eines Testgebietes in Flaesheim nach abgeschlossener Klassifikation. In Grün die idealtypischen Verdachtsflächen, die für eine Interpretation am interessantesten sind (eigene Darstellung)

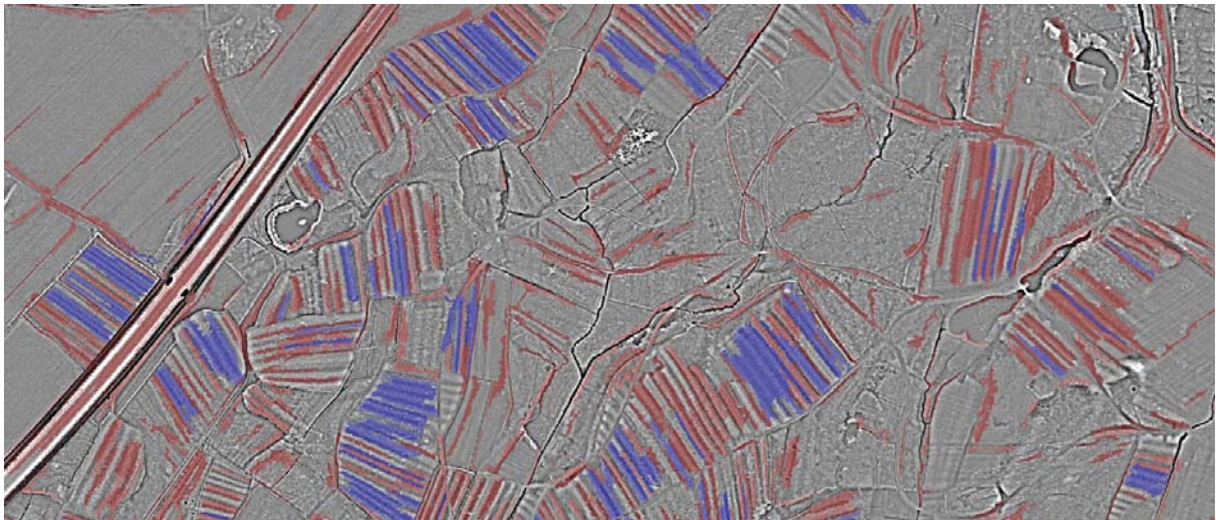


Abb. 8: Klassifikationsergebnis im Wildpark Dülmen. In Blau die erfassten Wölbackerstrukturen. In Rot weitere Strukturen, die zur Klassifikation der blauen Flächen benötigt werden, für die Interpretation aber nicht von großer Bedeutung sind (eigene Darstellung)

5 Ausblick

Bisher konnte lediglich ein erster Eindruck der Möglichkeiten des Verfahrens erlangt werden. Zukünftig könnte daher einerseits die Qualität der Klassifikation und andererseits der Grad der Automation erhöht werden, was die Methode für größere Untersuchungsgebiete interessant machte.

Die Klassifikation ist stark von der Datenqualität abhängig. Bisher wurden LiDAR-Daten aus den Jahren 2008-2010 verwendet. Inzwischen stehen jedoch neuere und genauere Daten zu Verfügung, die derzeit auf ihre Eignung geprüft werden.

Vor diesem Hintergrund 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. Möglicherweise lassen sich so Bodendenkmäler von natürlichen Strukturen unterscheiden, die im LiDAR-Scan identisch aussehen.

Der größte Qualitätsgewinn würde in der Datenvorbereitung im GIS erreicht werden, indem Flächen aus den Daten entfernt würden, die keine Bodendenkmäler mehr enthalten können – Siedlungs-, Verkehrs- und Abbauf Flächen kommen hier beispielsweise in Frage. Dies würde viele definitiv falsche Treffer bereits im Vorhinein vermeiden.

Da in allen Untersuchungsgebieten mehr oder weniger negative Einflüsse der verwendeten Visualisierungen beobachtet wurden, könnte die Verwendung anderer, aufwendigerer Visualisierungen die Ergebnisse zukünftig verbessern. Hier ist erstens das *Local Relief Model*, einer Weiterentwicklung der *Difference Map*, interessant, da es z.B. die Probleme bei der Erfassung von Turmhügeln lösen sollte. Außerdem könnte sich der *Sky-View Factor* als nützlich erweisen, da sich diese Visualisierung für Hänge eignet.

Auf Seiten von *eCognition* könnten die Trefferquoten mit strengeren, feineren und neuen Klassenbeschreibungen verbessert und verschiedene Erhaltungszustände durch mehrere, unabhängige Segmentationen abgedeckt werden. Neben den bisherigen Ansätzen gibt es noch weitere Möglichkeiten, Bodendenkmäler abzugrenzen, die bisher nicht behandelt werden konnten. Für eine Anwendung der ausgearbeiteten Methode in größerem Rahmen ist die Nutzung der Stapelverarbeitungsmodi von *ArcGIS* und *eCognition* vielversprechend und erforderlich, da so mehrere Datensätze auf einmal bearbeitet werden könnten. Dies wäre ein weiterer, großer Schritt hin zur vollständigen Automation.

6 Literaturverzeichnis

- DONEUS, M., 2013: Openness as Visualization Technique for Interpretative Mapping of Airborne Lidar Derived Digital Terrain Models. *Remote Sensing* **5**(12), 6427-6442.
- DONEUS, M. & BRIESE, C., 2006: Full-waveform airborne laser scanning as a tool for archaeological reconnaissance. From Space to Place. 2nd International Conference on Remote Sensing in Archaeology. Proceedings of the 2nd International Workshop, CNR, Rome, Italy, December 4-7, 2006, CAMPANA, S. & FORTE, M. (Hrsg.), Oxford, 99-105.
- HESSE, R., 2010: LiDAR-derived Local Relief Models – a new tool for archaeological prospection. *Archaeological Prospection* **17**, 62-72.
- 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.

Geiger-Mode – LiDAR-Technologien der nächsten Generation

THOMAS BAHR¹, NICOLAI HOLZER¹ & PHILIP SMITH²

Zusammenfassung: Die neuen Harris-Lösungen für Geiger-Mode Avalanche Photodiode (GmAPD) LiDAR umfassen Sensorik, Datenprozessierung und Analytik. Nach 15 Jahren Entwicklung für den Militärsektor bietet Harris diese Technologien erstmals im kommerziellen LiDAR-Markt an. Dieser Beitrag beschreibt, warum GmAPD die Effizienz von LiDAR-Akquisitionen steigert, wie der ITI-1000 Sensor eine kostengünstige und hochauflösende 3D-Topographieerfassung von über 1000 km²/h ermöglicht, und wie das zugehörige Breitband-Produktionssystem trotz erhöhter Präzision die Verarbeitungsdauer für GmAPD-Punktwolken reduziert. Zur Bewertung der GmAPD LiDAR-Daten, insbesondere für das „3D Elevation Program“ des Geologischen Dienstes der USA (USGS), werden zwei Fallstudien vorgestellt.

1 Einführung

Seit den frühen 1990er-Jahren erlebte der kommerzielle Markt für flugzeuggetragene LiDAR-Systeme ein beeindruckendes Wachstum. Mit der zunehmenden Verfügbarkeit von LiDAR-Punktwolken verbreitete sich auch ihre Nutzung in den unterschiedlichsten Branchen. Für die Expansion dieses Marktes und seine Entwicklung in neuen Anwendungsgebieten müssen jedoch die Kosten gesenkt werden, die mit der großflächigen und präzisen Erfassung von LiDAR-Daten mit hohen Punktdichten sowie deren Produktion verbundenen sind (CARY 2009).

Um dies zu erreichen, hat Harris neue Lösungen für die Sensorik, Datenprozessierung und Analytik von Geiger-Mode Avalanche Photodiode (GmAPD) LiDAR entwickelt. Verglichen mit linear betriebenen LiDAR-Sensoren sind GmAPD-gestützte Systeme wesentlich effizienter. GmAPD-Detektoren erfassen einzelne Photonen und sind darüber hinaus kleinvolumig. Daher lassen sie sich zu großen Feldern kombinieren und mit verfügbaren Lasern betreiben, die Pulsfrequenzen im Sub-Nanosekunden-Bereich aufweisen. Beides ermöglicht die Entwicklung eines GmAPD LiDAR-Systems zur 3D-Topographieerfassung mit einer Abdeckungsrate von z. B. 1000 km²/h bei einer Dichte von 8 Punkten/m². In Kombination mit fortschrittlichen Automatisierungstechniken für Breitband-Produktion und Datenanalyse bieten GmAPD-Sensoren neue Möglichkeiten für den kommerziellen Markt.

2 GmAPD-Detektorfelder

Im Vergleich zu linear betriebenen Systemen, die Photonen proportional in Stromsignale umwandeln, bestehen Geiger-Mode Detektoren aus einem Feld von Avalanche Photodioden (APDs), die mit hohem Bias betrieben werden. Damit werden große und einfach detektierbare Stromsignale erzeugt, die sich mit digitalen Zeitmesskreisen erfassen lassen, auch wenn nur

¹ Harris Corporation, Talhofstraße 32a, D-82205 Gilching,
E-Mail: [Thomas.Bahr, Nicolai.Holzer]@harris.com

² Harris Corporation, 1025 W Nasa Blvd, Melbourne, Florida 32919, USA, E-Mail: psmith29@harris.com

wenige Photonen einfallen. Aus folgenden Gründen sind GmAPD-Detektorfelder hervorragende LiDAR-Empfänger:

- **Geringer Jitter** – GmAPD-Felder integrieren einfache digitale CMOS Zeitmesskreise mit den einzelnen APDs. Bei der Messung der Laserpulslaufzeit reagieren die APDs innerhalb von 100 ps und ermöglichen so eine ausgezeichnete Präzision im Range-Bereich (ITZLER et al. 2010).
- **Hohe Detektions-Sensitivität und -Effizienz** – Derzeit verfügbare GmAPD Elemente weisen eine Photonendetektionseffizienz von bis zu 40 % auf. Dies bedeutet, dass ein reflektierter Puls aus lediglich 5 bis 10 Photonen bestehen muss, damit er mit hoher Wahrscheinlichkeit erfasst wird – eine Größenordnung weniger im Vergleich zu linearen Detektoren. Zusätzlich steigern Mikro-Linsen den effektiven Detektor-Füllfaktor auf ~ 75 %. Damit wird sichergestellt, dass jedes einfallende Photon effizient detektiert wird (ITZLER et al. 2010).
- **Kompakte Detektoren** – Da GmAPD-Detektorelemente keine analogen Verstärkungsschaltungen benötigen, können kompakte LiDAR-Detektorfelder mit hohen Abtastdichten produziert werden.

Dank dieser drei Eigenschaften können GmAPD-Sensoren schneller, höher und mit größeren Schwadbreiten geflogen werden – ohne Abstriche bei der vertikalen Genauigkeit.

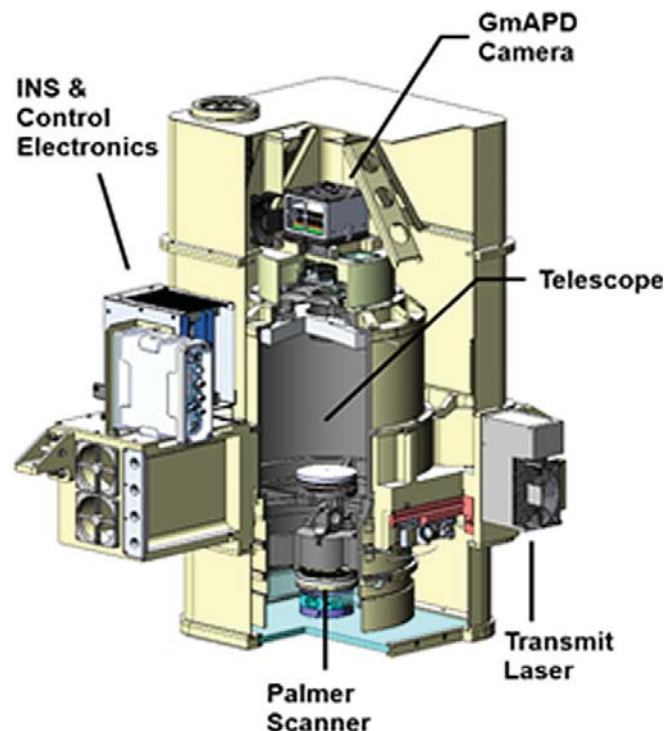


Abb. 1: CAD-Diagramm des Geiger-Mode LiDAR Sensors ITI-1000 (nach SMITH 2015)

3 ITI-1000 GmAPD LiDAR-Sensor

Der ITI-1000 GmAPD LiDAR-Sensor (Abb. 1 und Tab. 1) wurde für großflächige Datenakquisitionen mit großen Punktdichten optimiert. Das Konzept sieht den Betrieb in Flughöhen von 4 bis 10 km über Grund vor. Zum Einsatz kommen daher Starrflügelflugzeuge, die mit Druckkabinen ausgestattet sind. Die aufgenommenen Daten müssen mindestens die USGS v1.2 LiDAR-Qualitätsstufe 1 erfüllen, d. h. 8 Punkte/m² bei einer Flughöhe von 6 km (HEIDEMANN 2014). Der ITI-1000 nutzt dazu eine gleichachsige Sende- und Empfangseinheit, ausgestattet mit einem 270 mm Palmer Scanner mit 15° Erfassungshalbwinkel.

Tab. 1: Technische Eigenschaften des Geiger-Mode LiDAR Sensors ITI-1000 (nach SMITH 2015)

Flug	
Flughöhe (GND)	4-10 km (6 km nominell)
Fluggeschwindigkeit	280-560 km/h
Schwadbreite	2-5 km (3,2 km nominell)
Palmer Scanner	
Erfassungshalbwinkel	15°
Öffnungsdurchmesser	27 cm
Sendelaser	
Wellenlänge	1064 nm (Klasse IV)
Mittlere Leistung	20 W
Impulsbreite	550 ps
Impulsfolgefrequenz	50 kHz
NOHD / ENOHD	300 m / 2,2 km
GmAPD-Empfänger	
Felddimensionen	32 x 128
IFOV	35 µrad
PDE	30 %
Zeitliche Auflösung	250-500 ps
Abdeckungsrate (mit 50 % Überlappung)	
4 Punkte / m ²	1100 km ² /h
8 Punkte / m ²	1000 km ² /h
20 Punkte / m ²	700 km ² /h

Aufgrund der vorgesehenen Flughöhen sollte der Ausrichtungsversatz zwischen Sende- und Empfangseinheit für eine optimale Überlappung des Sendestrahls mit dem empfangenen Sichtfeld (FOV) elektronisch gesteuert werden können. Für den ITI-1000 wurde die dazu notwendige Technologie entwickelt und zum Patent angemeldet. Des Weiteren ist der ITI-1000 Sensor mit einer automatischen Verschlusssteuerung sowie einem Instrument zur Verfolgung und Kompensation von Mehrfachimpulsen ausgestattet – entscheidende Funktionen bezüglich der operationellen Einsatzflughöhen und Impulsfolgefrequenzen.

Bei der Datenakquisition mit GmAPD-Sensoren erfolgt die Digitalisierung der reflektierten Laserimpulse bereits auf Detektorebene. Daher muss ein anderer Aufnahmeansatz als bei derzeitig linear betriebenen Sensoren verfolgt werden:

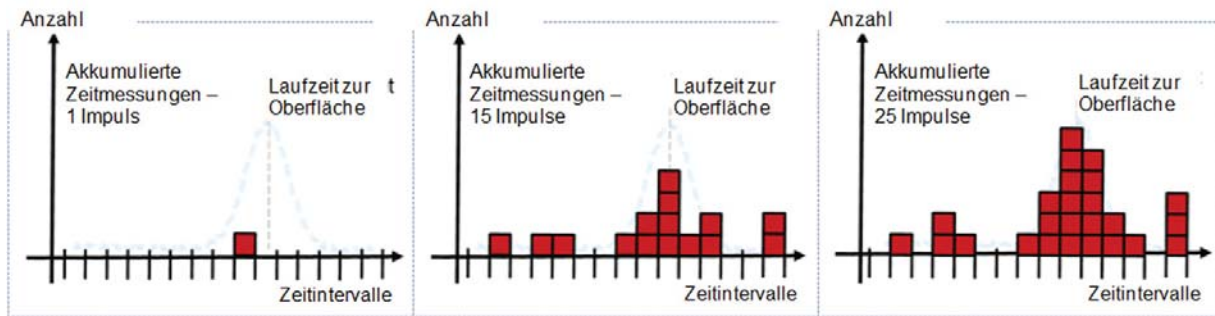


Abb. 2: GmAPD-Sensoren messen die Rangedistanz durch Anpassung der zu erwartenden Impulsform an die akkumulierten Zeitmessungen und Berechnung der gleichen Zeitdifferenz (SMITH 2015)

Auch GmAPD-basierte Systeme berechnen den Range durch eine Laufzeitmessung der Laserimpulse über eine vordefinierte Verschlusszeit. Allerdings werden aufgrund des Binärcharakters der Detektoren Mehrfachimpulse zur Erkennung von Oberflächen in den Daten benötigt. Denn nur durch die Aussendung multipler Impulse innerhalb des gleichen momentanen Sichtfeldes (IFOV) und deren anschließende Akkumulation ist es möglich, eine Oberfläche vom Rauschen zu unterscheiden (Abb. 2). Die Anzahl der emittierten Impulse wird entsprechend der vorgegebenen missionsabhängigen Auflösung und Punktdichte vorprogrammiert.

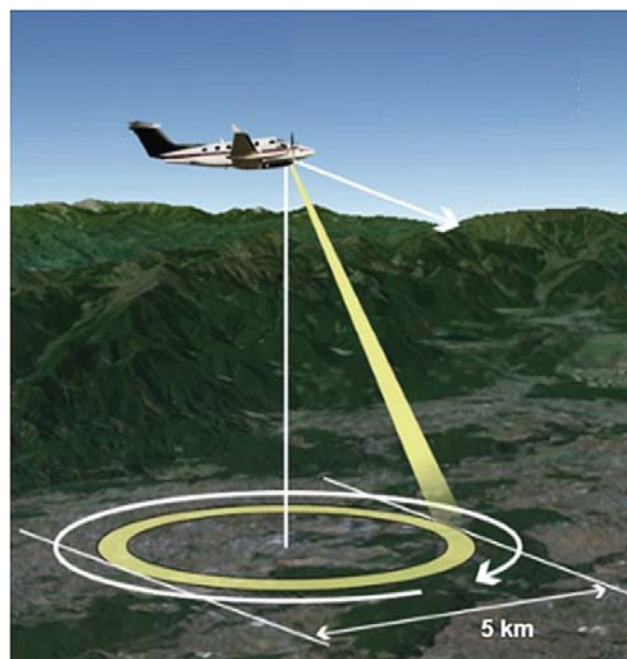


Abb. 3: Aufnahmeprinzip des Palmer Scanner im ITI-1000. Die vorderen und hinteren Scankreise werden während der Prozessierung als einzelne Schwaden behandelt (nach SMITH, 2015)

Für eine optimierte Datenakquisition mit dem ITI-1000 Sensor wurde ein spezielles Aufnahmeverfahren entwickelt (Abb. 3 und 4). Jeder Bodenpunkt wird mit der entsprechenden Impulsfolge in jedem Halbbogen abgetastet. Daraus ergeben sich separate vorwärts- und rückwärts-gerichtete Schwaden. Zusätzlich werden die Flugstreifen unter Berücksichtigung des Rollwinkels mit mindestens 50 % Überlappung ausgerichtet. So wird jeder Bodenpunkt aus vier

verschiedenen Richtungen beleuchtet. Die Daten dieser vier Bögen werden schließlich während der Prozessierung integriert. Dadurch reduzieren sich die Eigenschatten der aufgenommenen Objekte; zudem wird eine höhere Punktdichte erreicht.

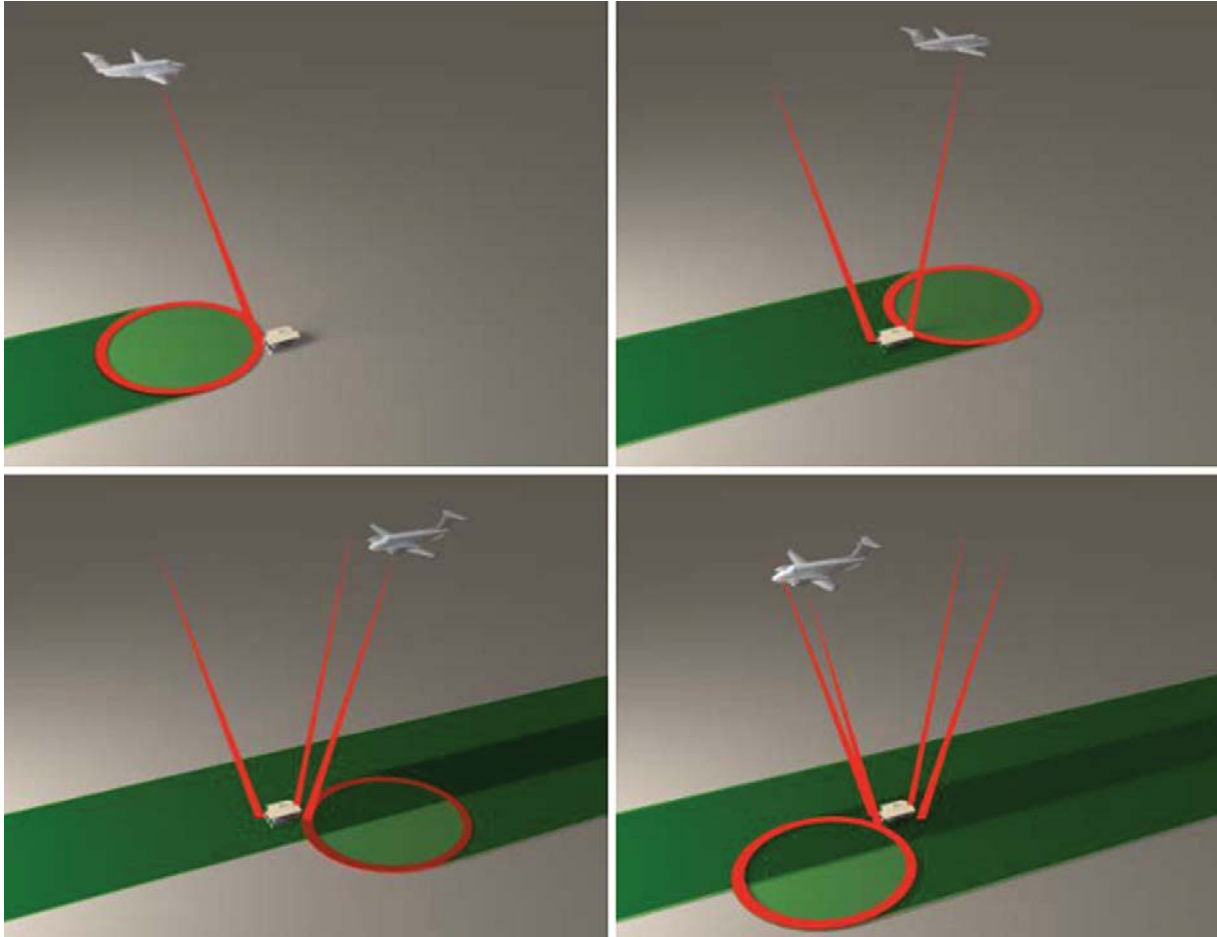


Abb. 4: Aufnahmeprinzip des Palmer Scanner im ITI-1000 mit 50 % Überlappung der Flugstreifen. Jeder erfasste Bodenpunkt wird über vier Kreise und unter vier verschiedenen Winkeln abgetastet (CLIFTON et al. 2015)

4 Großvolumige LiDAR-Datenverarbeitung

Für die Prozessierung großvolumiger LiDAR-Daten aus Akquisitionen des ITI-1000 GmAPD-Sensors zu qualitativ hochwertigen Endprodukten wurde ein automatisierter Arbeitsablauf aufgesetzt (Abb. 5). Die einzelnen Softwarekomponenten für z. B. Sensorkalibrierung, Rauschfilterung oder photogrammetrische Registrierung stammen aus dem Harris-Portfolio. Dabei soll in diesem Beitrag auf zwei wesentliche Verarbeitungsschritte eingegangen werden: die Rauschunterdrückung und die photogrammetrische Registrierung der Punktwolkedaten.

Obwohl analoges Verstärkerrauschen kein Problem darstellt, sind GmAPD-Detektorfelder aufgrund ihrer hochsensiblen Erfassungstechnik sehr empfindlich gegenüber Rauschen von anderen Quellen (Abb. 6). Hier dominiert – typisch für alle luftgestützten LiDAR-Systeme – der Anteil solarer Hintergrundsignale. Dieses Rauschen ist bei Tagaufnahmen etwa zwei

Größenordnungen höher als die Dunkel-Zählrate des Instruments (d. h. die Anzahl der Detektionen ohne Lichteinfall) und das Übersprechen zwischen den Detektoren. Das solare Hintergrundrauschen wird durch schmalbandige Interferenzfilter und eine optimierte Verschlusszeit reduziert.

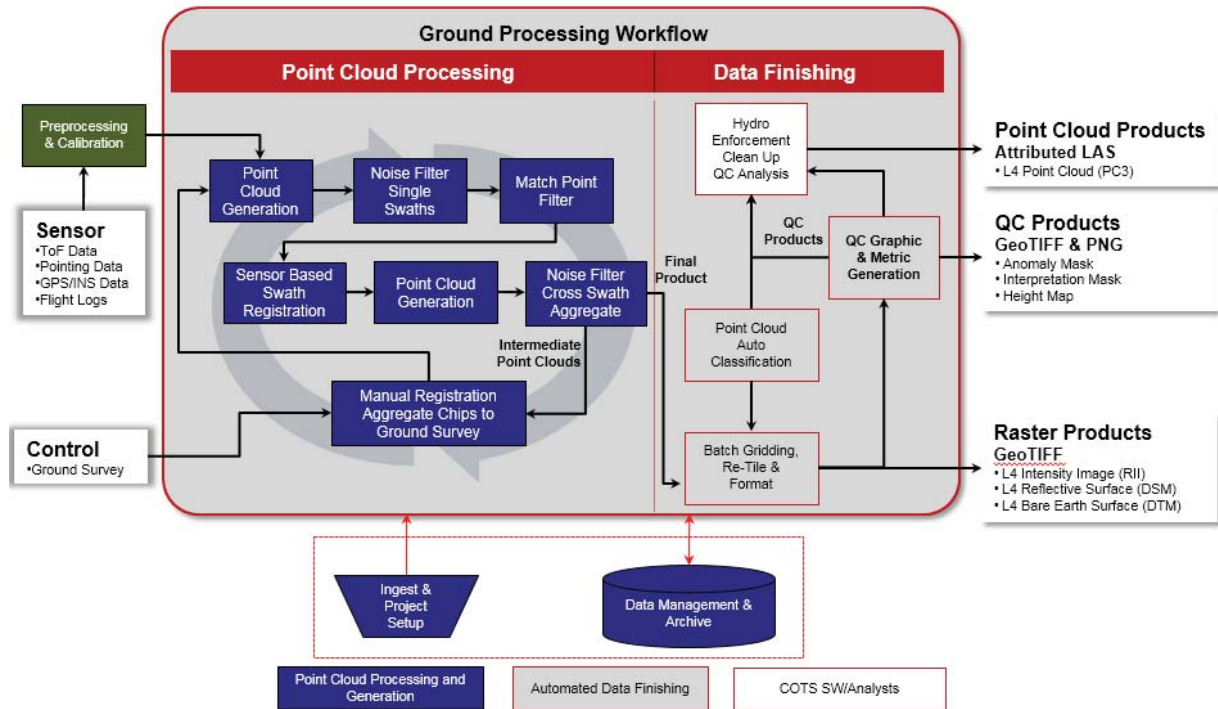


Abb. 5: Automatisierter Arbeitsablauf für die Prozessierung großvolumiger LiDAR-Daten aus Akquisitionen des ITI-1000 GmAPD-Sensors (SMITH 2015)

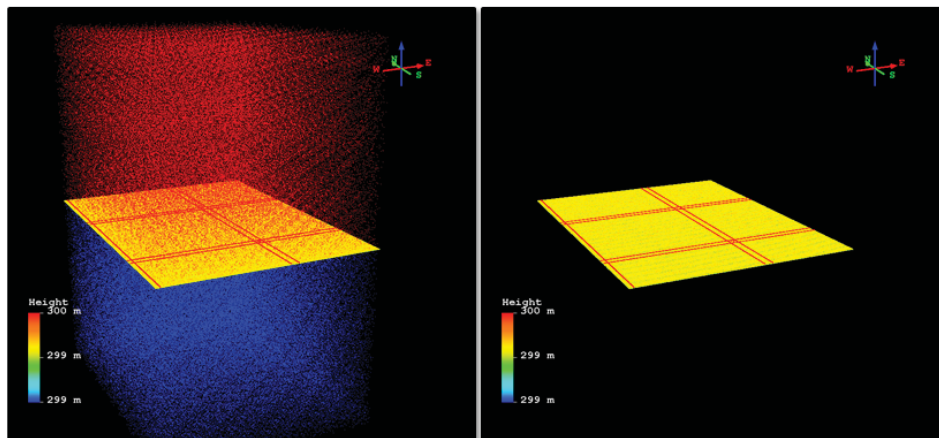


Abb. 6: Aufgrund der Empfindlichkeit von GmAPD-Detektoren weisen die daraus abgeleiteten Punktwolken erhöhtes Rauschen auf (links). Spezielle Filteralgorithmen ermöglichen die Rekonstruktion der Oberflächen (rechts) (SMITH 2015)

Geiger-Mode Sensoren benötigen hochgenaue Flugbahninformationen, da die Punktwolken über die Aggregation multipler Impulse generiert werden. Dazu wurde der ITI-1000 mit Modernster GPS-Technologie ausgestattet. Zusätzlich ist eine echte 3D-Registrierung der Einzelschwaden

(vgl. Kap. 3 und Abb. 3, 4) erforderlich. Hierfür wurde ein Verfahren entwickelt, das auf Basis des Sensormodells einen Bündelausgleich vornimmt (Abb. 7). Die verwendeten photogrammetrischen Registrierungstechniken justieren die Einzelschwaden mittels automatisch erzeugter Verknüpfungspunkte. Bei Bedarf lassen sich Bodenvermessungspunkte (GCPs) in diesen Prozess integrieren.

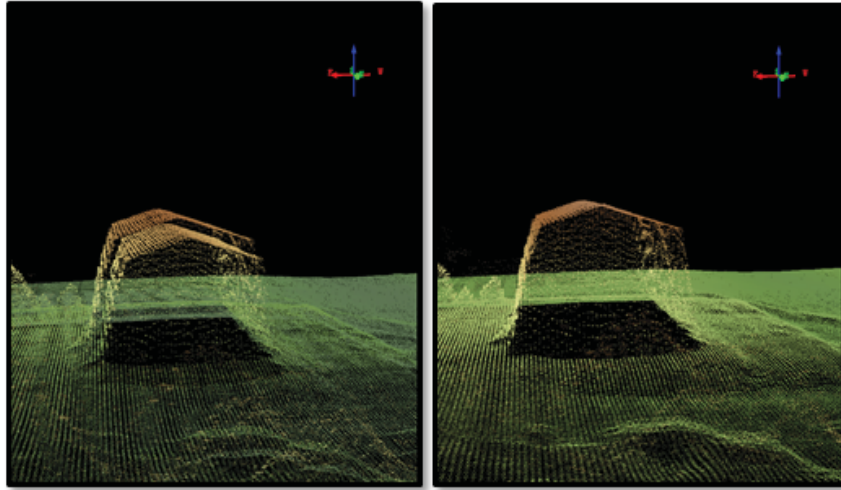


Abb. 7: (Links) Unregistrierte, aus zwei Flugstreifen integrierte GmAPD-Punktwolke. Die fehlerhafte Ausrichtung von Gebäude und Gelände ist vor allem auf Ungenauigkeiten der INS-Flugbahn zurückzuführen. (Rechts) Aggregierte Punktwolke nach vollständigem, sensorbasiertem 3D-Bündelausgleich (SMITH 2015)

5 Fallstudie zur Bewertung der GmAPD LiDAR-Daten

STOKER et al. (2016a) untersuchten LiDAR-Daten des Harris GmAPD-Sensors bezüglich ihrer Verwendbarkeit für das „3D Elevation Program“ (3DEP, USGS 2016) des Geologischen Dienstes der USA (USGS). Dafür wurde im September und Dezember 2015 im US-Bundesstaat Connecticut ein ca. 1300 km² großes Testgebiet mit dem ITI-1000 Sensor befliegen, bei einer Flughöhe von 7950 m und einer Punktdichte von 25 Punkten/m². Die Referenzdaten wurden in den Jahren 2014 und 2015 mit linear betriebenen LiDAR-Sensoren aufgenommen (z. B. Leica ALS70) und erfüllten die USGS v1.2 Qualitätsanforderungen Q2 (HEIDEMANN 2014).

Die von Harris bereitgestellten, photogrammetrisch kalibrierten LAS Punktwolken (vgl. Kap. 4) wurden von den Firmen Dewberry und Woolpert mit den Softwares Global Mapper, GeoCue, TerraScan, TerraModeler und Microstation ausgewertet.

Der Studie zufolge erfüllten die Geiger-Mode LiDAR-Daten problemlos die Anforderung von 6 cm relativer vertikaler Genauigkeit bei ebenem und leicht hügeligem Gelände („Smooth Surface Repeatability“). Nur bei stark reflektierenden Oberflächen wurden Abweichungen von ~15 cm gemessen (sogenannte „Range Walk“ Effekte), jedoch blieben die Oberflächenformen detailliert erhalten. Die Testergebnisse für die absolute vertikale Genauigkeit sind in Tab. 2 zusammengestellt. Hier entsprechen die Geiger-Mode LiDAR-Daten den Anforderungen der USGS v1.2 Qualitätsstufen Q1 und Q2. Lediglich in Waldgebieten wurde aufgrund des „Leaf-On“ Zustandes während der Datenakquisition der Schwellwert von 29,6 cm überschritten

Allerdings war die Durchdringung des Kronendaches bei den „Leaf-Off“ Daten sehr gut (Abb. 8): aus der Punktwolke wurde eine Bodendichte von 14 Punkten/m² abgeleitet. Die aus den Geiger-Mode LiDAR-Daten erzeugten Reflexionsbilder zeigten ähnliche Grauwerteigenschaften wie die Intensitätsbilder der Referenzdaten und erlaubten somit vergleichbare Interpretationen. Im Vergleich von Nacht- und Tagesakquisitionen konnten nur geringe Unterschiede in der Kronendurchdringung bzw. Kronenstruktur nachgewiesen werden. Auch der Rauschteil wurde durch das solare Hintergrundrauschen nicht wesentlich erhöht – eine Folge der sensorseitigen Rauschunterdrückung und der effizienten Filterung im Rahmen der Datenvorprozessierung (vgl. Kap. 4).

Tab. 2: Test der absoluten Genauigkeit des ITI-1000 (Flughöhe 7950 m) mit Einzelpunkten und digitalen Geländemodellen (nach STOKER et al. 2016a)

USGS Spezifikation (nach HEIDEMANN 2014)		
	Vertikale Genauigkeit unbewachsene Flächen (NVA)	Vertikale Genauigkeit bewachsene Flächen (VVA)
Konfidenzkoeffizient 95 %*	≤ 19,6 cm	≤ 29,4 cm
Test 1 – Einzelpunkte aus der Punktwolke		
ITI-1000	17,0 cm	25,6 cm
USGS QL2 Referenzdaten	12,3 cm	19,8 cm
Test 2 – Digitales Geländemodell		
ITI-1000	15,2 cm	92,0 cm
USGS QL2 Referenzdaten	14,6 cm	25,0 cm

* Minimale Genauigkeitsanforderungen an USGS v1.2 Qualitätsstufen Q1 und Q2.

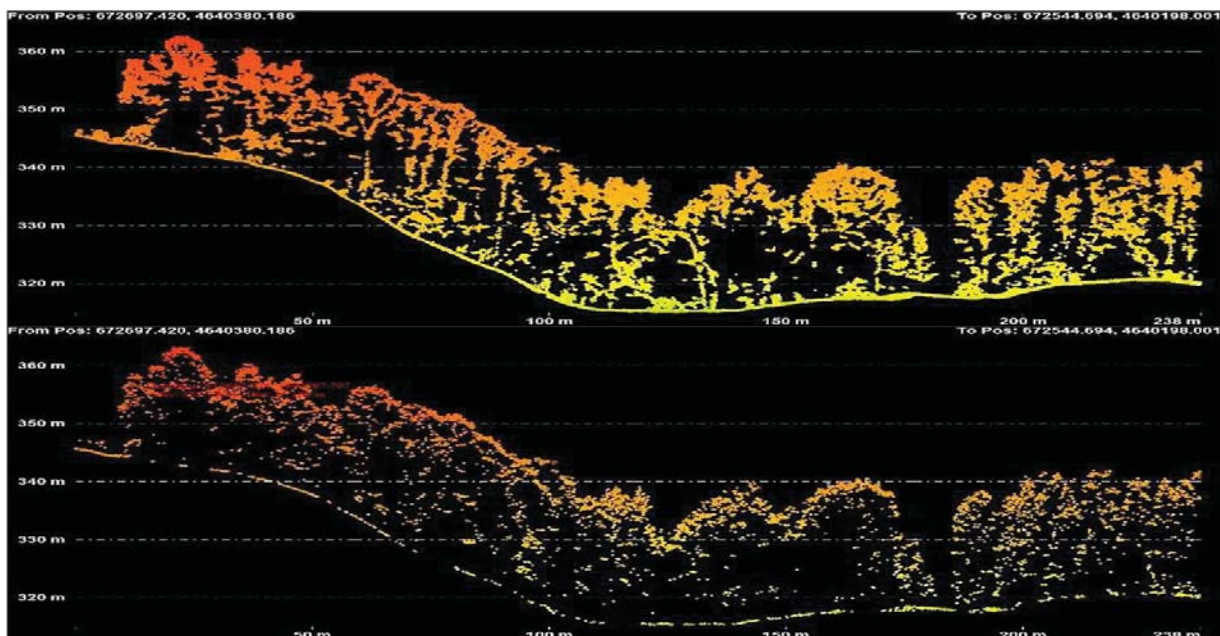


Abb. 8: Profil durch ein Waldgebiet (unbelaubt), aufgenommen mit Geiger-Mode LiDAR (oben, Flughöhe 7950 m) und linear betriebenen LiDAR (unten, Flughöhe 917 m) (STOKER et al. 2016b)

Nicht berücksichtigt wurde in der Studie von STOKER et al. (2016a) die neueste Sensorgeneration mit zahlreichen Verbesserungen, u. a. einem neuen holografischen Element, das eine präzisere Fokussierung und dadurch eine verbesserte Kronendurchdringung ermöglicht.

Auch die Untersuchungen durch FLOOD (2016) zeigen die Fähigkeit von Geiger-Mode LiDAR-Daten, lineare Konstruktionen wie z. B. Versorgungs- und Telefonleitungen auch bei 10-facher Flughöhe detailreich in einer Qualität abzubilden, die visuell den Linear-Mode Vergleichsdaten entspricht (Abb. 9). Dies erlaubt die Extraktion von z. B. Stromleitungen oder Gebäudeumrissen mit herkömmlichen LiDAR-Softwares.

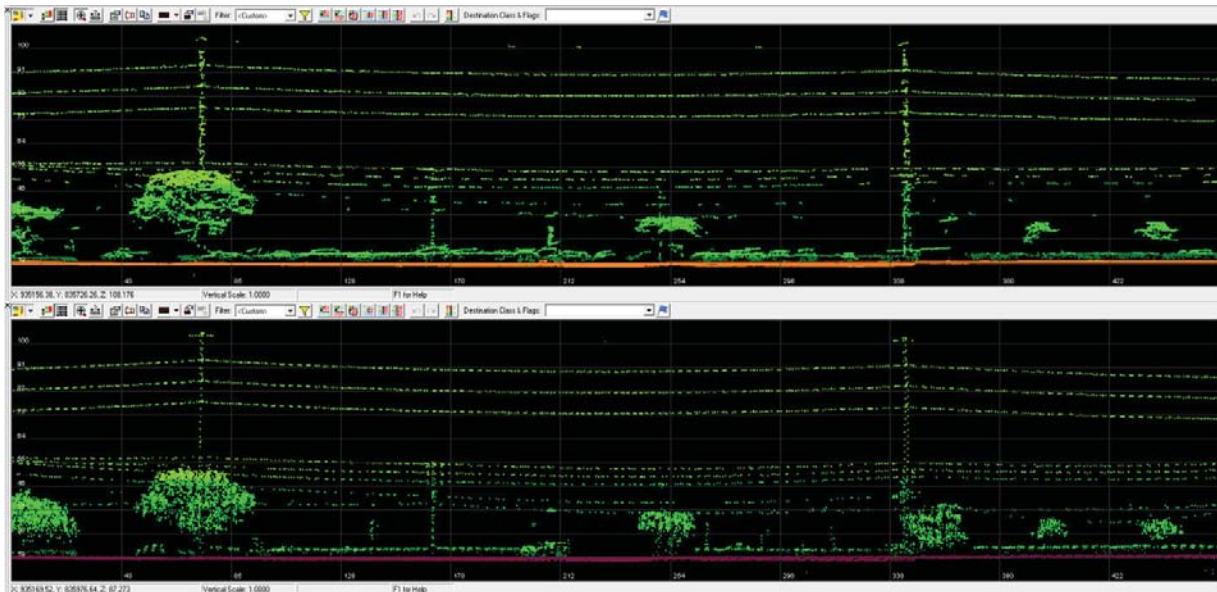


Abb. 9: Versorgungs- und Telefonleitungen in einem Wohngebiet, aufgenommen mit Geiger-Mode LiDAR (oben, Flughöhe 5180 m) und linear betriebenem LiDAR (unten, Flughöhe 518 m) (FLOOD 2016)

6 Fazit

Die neuen Harris-Lösungen für Geiger-Mode Avalanche Photodiode (GmAPD) LiDAR umfassen Sensorik, Datenprozessierung und Analytik.

Im Vergleich zu linear betriebenen LiDAR-Systemen können GmAPD-Sensoren wie der ITI-1000 schneller, höher und mit größeren Schwadbreiten geflogen werden – ohne Abstriche bei der vertikalen Genauigkeit. Dies bedeutet eine kostengünstige und hochauflösende 3D-Topographieerfassung. Das zugehörige Breitband-Produktionssystem reduziert trotz erhöhter Präzision die Verarbeitungsdauer für GmAPD-Punktwolken. Aus den resultierenden 3D-Punktwolken lassen sich dann mit einer Suite aus automatisierten und manuellen Werkzeugen hochwertige Endprodukte wie z. B. Oberflächen- und Geländemodelle generieren.

Nach 15 Jahren Entwicklung für den Militärssektor bietet Harris diese Technologien nun erstmals und als einziges Unternehmen im kommerziellen LiDAR-Markt an.

Geiger-Mode LiDAR-Daten können schon bald für das „3D Elevation Program“ des USGS genutzt werden, da die 2014 für linear betriebenes LiDAR definierten Qualitätsanforderungen

derzeit überarbeitet werden. Die aktualisierte Version wird dann Parameter enthalten, die sich aus neuen Technologien wie Geiger-Mode LiDAR ergeben.

7 Literaturverzeichnis

- CARY, T., 2009: LiDAR Market: Status and Growth Trends. International LiDAR Mapping Forum, Denver, CO.
- CLIFTON W. E., STEELE, B., NELSON, G., TRUSCOTT, A., ITZLER, M. & ENTWISTLE, M., 2015: Medium Altitude Airborne Geiger-Mode Mapping Lidar System. Proc. SPIE **965**, 946506-1 – 946506-8.
- FLOOD, M., 2016: Geiger-Mode LIDAR Workflow Review. <http://support.geocue.com/geiger-Mode-lidar-workflow-review/>
- HEIDEMANN, H. K., 2014: Lidar base specification. U.S. Geological Survey Techniques and Methods, 11-B4. doi: 10.3133/tm11B4
- ITZLER, M., ENTWISTLE, M., OWENS, M., PATEL, K., JIANG, X., SLOMKOWSKI, K. & FERRAR, J., 2010: Geiger-Mode Avalanche Photodiode Focal Plane Arrays for Three-Dimensional Imaging LADAR. Proc. SPIE **7808**, 78080C-1 – 78080C-14.
- SMITH, P., 2015: High Density, High Volume Point Cloud Production from Airborne Geiger Mode Lidar. International LiDAR Mapping Forum, Denver, CO.
- SMITH, P. W. & CLIFTON, W., 2012: An Overview of ALS Technology: Geiger-Mode LIDAR. RENSLOW, M. (Hrsg.), Manual of Airborne Topographic LIDAR, Bethesda, MD, ASPRS, 91-96.
- STOKER, J. M., ABDULLAH, Q. A., NAYEGHANDI, A. & WINEHOUSE, J., 2016a: Evaluation of Single Photon and Geiger Mode Lidar for the 3D Elevation Program. Remote Sensing **8**(9), 767.
- STOKER, J. M., ABDULLAH, Q. A. & NAYEGHANDI, A., 2016b: Geiger Mode and Single Photon Counting Evaluation. https://calval.cr.usgs.gov/wordpress/wp-content/uploads/GM-SPL-evaluation_JACIE2016_v3.pdf
- USGS, 2016: 3D Elevation Program (3DEP). <https://nationalmap.gov/3DEP/>

Urban Monitoring by 4D Change Detection Using Multi-temporal SAR Images

CHIA-HSIANG YANG¹ & UWE SOERGEL¹

Zusammenfassung: Die wachsende Bevölkerung und der Trend zur Urbanisierung haben viele bauliche Veränderungen in Städten zur Folge. Deren Erkennung und Überwachung ist essentiell für Zwecke wie kommunale Verwaltung, Stadtplanung oder Katasterfortführung. Konventionelle Vermessungen vor Ort sind zwar sehr genau, jedoch aufwendig und können nur punktuell durchgeführt werden. Im Gegensatz dazu ist mit Fernerkundung eine kostengünstige und großflächige Datengewinnung möglich. Moderne SAR-Satelliten sind in der Lage, hochaufgelöste Radarbilder in dichtem zeitlichen Raster zu erfassen. Aktives SAR hat zudem den Vorteil der Allwetterfähigkeit und der Nachtsicht, weshalb sich SAR insbesondere für auf Zeitreihen von Bildern basierende Überwachungsaufgaben eignet. Multi-temporale SAR-Bilder werden daher häufig zu Zwecken der Änderungsdetektion ausgewertet. Wir stellen ein neu entwickeltes Verfahren zur 4D-Änderungserkennung vor, womit gemeint ist, dass sowohl räumliche (3D) als auch zeitliche (1D) Änderungen erkannt werden können. Dieses Verfahren wird anhand von SAR-Bildstapeln demonstriert, die den Innenstadtbereich von Berlin abdecken. Insbesondere liegt das Augenmerk auf der Erkennung von Baumaßnahmen, in deren Zuge alte Gebäude abgerissen oder neue errichtet werden.

Abstract: The continuous rise in population and economic growth has led to frequent urban changes such as construction. Monitoring such changes is important for city management, urban planning, updating of cadastral maps, etc. In contrast to conventional field surveys, which are usually expensive and slow, remote sensing techniques are fast and cost-effective alternatives. Spaceborne synthetic aperture radar (SAR) provides radar images captured rapidly over vast areas at fine spatiotemporal resolution. In addition, the active microwave sensors are capable of day-and-night vision and independent of weather conditions. The advantages mentioned above make SAR suitable for monitoring tasks. Change detection approaches based on multi-temporal SAR images are widely used for urban monitoring. We developed a novel 4D change detection technique, which is capable of detecting spatial changes (3D) and occurrence times (1D). In this study, we apply our technique to a built-up area in the centre of Berlin, Germany. As a result, the disappearing and emerging structures along with their occurrence times are successfully detected. We have demonstrated that these spatiotemporal results are able to provide detailed and comprehensive information for urban monitoring.

1 Introduction

Human activities, such as population growth, economic globalization, urban extension, and natural disasters have led to frequent urban changes. Monitoring such changes is important for city management, urban planning, updating of cadastral map, environmental monitoring, disaster assessment, etc. (GAMBA 2013; MARIN et al. 2015). In contrast to conventional field surveys, which are usually expensive and slow, remote sensing techniques are fast and cost-effective

¹ Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, E-Mail: [chia-hsiang.yang, uwe.soergel]@ifp.uni-stuttgart.de

alternatives. Spaceborne synthetic aperture radar (SAR) provides radar images captured rapidly over vast areas at fine spatiotemporal resolution. In addition, the active microwave sensors are capable of day-and-night vision and independent of weather conditions. The advantages mentioned above make SAR suitable for monitoring tasks.

Time series analysis based on SAR images is widely used for urban monitoring. Among them, persistent scatterer interferometry (PSI) (CROSETTO et al. 2016; FERRETTI et al. 2000, 2001, 2011; HOOPER et al. 2004; KAMPES 2006) detects and analyses PS points, which are characterized by strong, stable, and coherent radar signals throughout a SAR image stack and can be regarded as substructures in human settlements. Attributes of PS points, including line-of-sight velocities (mm/year level), topography heights, geographic positions, etc., can be derived and used for monitoring of structural deformation and 3D modelling. A prerequisite of forming PS points is that their signals must maintain temporal coherence. For example, to avoid coherence loss, buildings covered with PS-like substructures must be steady and free of any big changes during an entire acquisition period of SAR images. In contrast, if the substructures disappear or emerge arbitrarily due to construction, the corresponding temporary PS points are discarded at the initial screening of temporally stable scatterers in a standard PSI processing.

Certain approaches (BRCIC & ADAM 2013; FERRETTI et al. 2003; NOVALI et al. 2004; YANG et al. 2016) are dedicated to retrieve temporary PS points, which exist in a portion of time-series SAR images due to big changes. Among them, our previous work (YANG et al. 2016) proposed "4D Change Detection Based on PSI" (4DCDPSI) to recognize two types of temporary PS points subject to big changes, which are called disappearing big change (DBC) and emerging big change (EBC) points, along with their occurrence times. This technique has been validated by the simulated and real data tests.

In this paper, we apply our 4D change detection technique to the centre of Berlin, Germany to explore the applications of monitoring construction progress, business districts, sports playgrounds, traffic infrastructures, and single buildings. We first introduce PSI and 4DCDPSI in Sections 2 and 3, respectively. Section 4 demonstrates the spatiotemporal changes over the entire study area, followed by the in-depth discussions on five zoom-in areas with respect to different applications. Finally, the conclusions are summarized in Section 5.

2 Persistent Scatterer Interferometry

A time series of N complex SAR images, which are acquired from the same orbit and cover a common extent, is required as input data. Among the series, slave images are precisely co-registered to a master image, which is selected under small baseline constraint (BERARDINO et al. 2002; LANARI et al. 2004). Then, $N-1$ interferograms between the master and all of the slave images are computed. The interferometric phases of each pixel are used to estimate its temporal coherence, line-of-sight velocity, and relative topography height via a Periodogram process (FERRETTI et al. 2001). A temporal coherence serves as a measure of phase stability throughout the SAR image stack. Finally, pixels with high temporal coherences are selected as PS points. However, DBC and EBC points, if any, are just discarded as they suffer coherence loss during the entire SAR image sequence. To retrieve such big change information, we resort to the 4D change detection technique described in the next section.

3 4D Change Detection Based on Persistent Scatterer Interferometry

We first illustrate the change detection scheme subject to a single break date that big changes occur before or after. Complete, front, and back SAR image sets are defined from an image sequence for use in this scheme. The complete set consists of all of the images in the sequence. The front and back sets comprise the images taken before and after a specified break date, respectively. Our aim is to find PS points that exist in the front set but disappear in the back set and vice versa. The PS points that suddenly disappear are termed DBC points and those emerging in the back set are called EBC points.

The flowchart of single-break-date change detection (Fig. 1) is composed of the persistence, disappearance, and emergence scenarios, in which the complete, front, and back sets are mainly involved, respectively. These three image sets are processed by a standard PSI procedure to generate three temporal coherence images. We suppose that the temporal coherence of a DBC or EBC point in the front or back set is higher than that in the complete set, which is reduced due to the big change. Based on this assumption, the change indices of each pixel x in the disappearance (CI^D) and emergence (CI^E) scenarios are calculated by

$$CI^D(x) = \gamma_T^F(x) - \gamma_T^C(x) \quad (1)$$

$$CI^E(x) = \gamma_T^B(x) - \gamma_T^C(x) \quad (2)$$

where γ_T^C , γ_T^F , and γ_T^B denote the temporal coherences in the complete, front, and back sets. A pixel is more likely to be a DBC or EBC point when CI^D and CI^E is closer to 1, respectively. A change index distribution over DBC or EBC points is modelled to be a right-tailed probability function towards 1. Then, a statistical-based thresholding is applied to the change indices of the pixels to extract DBC and EBC points. The extracted points are jointly analysed with the PS points, which are selected in the persistence scenario, to reject two types of outliers. First, PS points are discarded if they coincide with the other two point labels. Second, a DBC point must not be an EBC point as well and vice versa. The remaining pixels without any label are regarded as undefined (U) points. Finally, the PS, DBC, and EBC points are combined into a change detection result. However, the accurate times of big changes are lacking as they are only known to disappear and emerge after and before the break date, respectively.

To detect accurate big change times, a set of single-break-date results are jointly analysed in a multi-break-date change detection (Fig. 2). For each pixel, two sequences, i.e., change indices and initial point labels (PS, DBC, EBC, or U), have been determined thus far. Some of the initial labels may be erroneous due to processing uncertainty in each single-break-date change detection. The majority vote is then applied to each label sequence to determine the pixel' label, e.g., a pixel is labelled PS if most of its initial labels are PS. Nevertheless, false labels are still unavoidable but can be removed by an outlier filtering. Three outlier types are described below along with their removal strategies using sliding window operation.

- Homogeneous points are expected to form a single object. For example, PS points are unlikely to appear on a demolished apartment full of DBC points. For this reason, PS, DBC, or BC points, which are in the majority in a window, are retained; the other inconsistent points are then deleted.

- An isolated PS, DBC, or EBC point in a window is removed considering that its reliability cannot be inspected by comparing with neighbours.
- A PS point is removed if its velocity is too large or quite different from the velocities of the neighbouring PS points in a window.

Once the PS, DBC, and EBC points are confirmed, the remaining points are considered to be U points. Then, the change date of each DBC or EBC point is detected from the time-series break dates based on the temporal variation in its change index sequence. The concept is to detect the turning point of a change index sequence, which corresponds to a disappearance or emergence date. In the end, the PS, DBC, and EBC points along with the change dates are combined to illustrate the spatiotemporal changes.

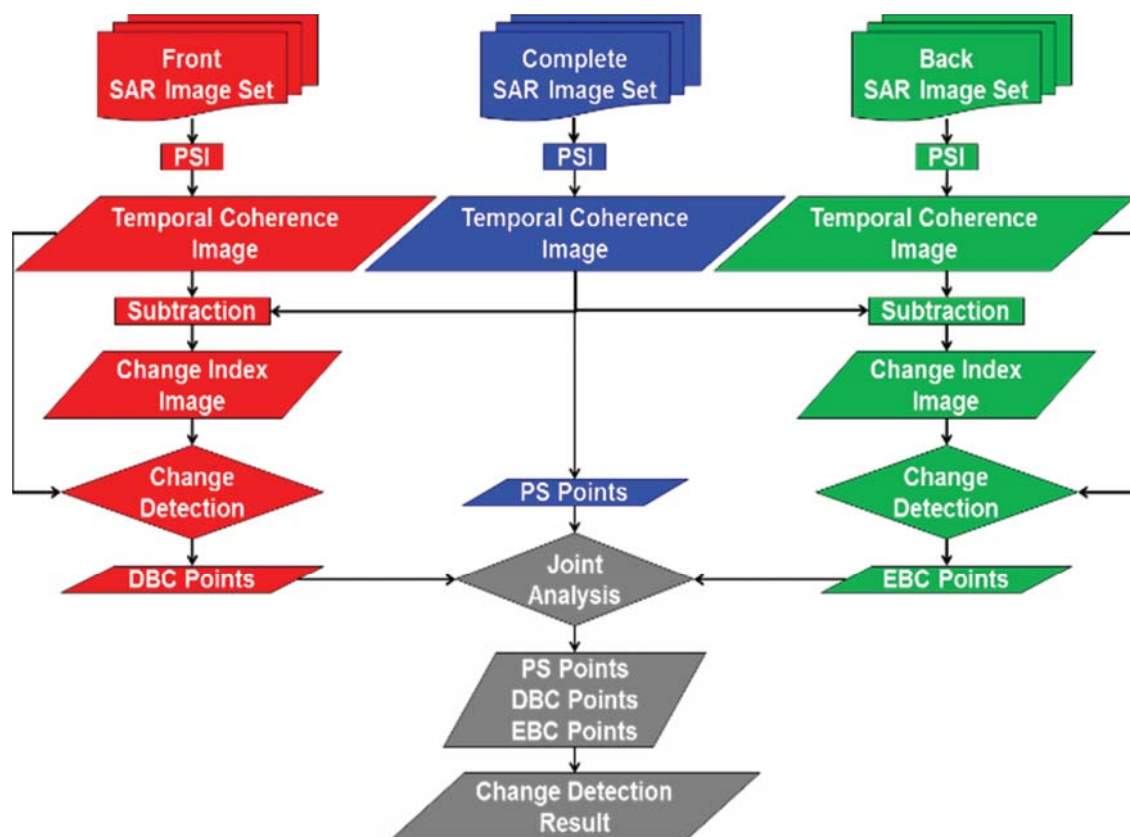


Fig. 1: Flowchart of single-break-date change detection. Persistence (blue), disappearance (red), and emergence (green) scenarios are dedicated to extracting PS, DBC, and EBC points, respectively (YANG et al. 2016)

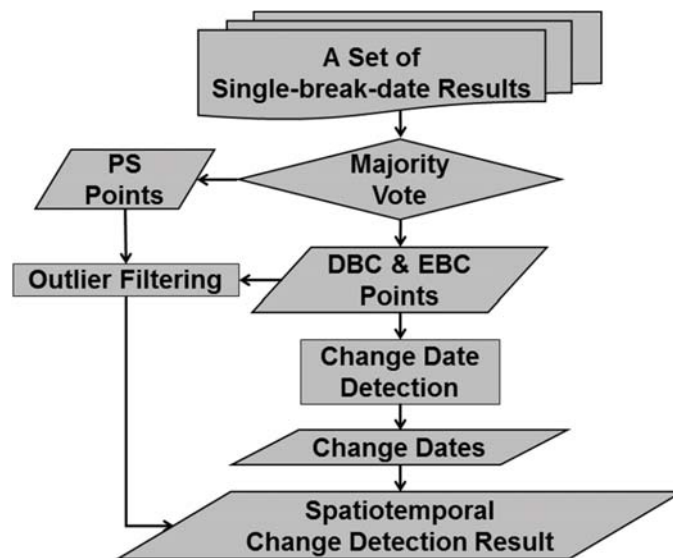


Fig. 2: Flowchart of multi-break-date change detection. (YANG et al. 2016)

4 Real Data Test

In this test, we adopted forty TerraSAR-X (TSX) images acquired in High Resolution Spotlight mode from October 27, 2010 to September 4, 2014. Thirteen break dates are set during 2013, i.e., we conducted thirteen single-break-date instances, which were jointly processed by multi-break-date change detection. All of the images are precisely co-registered and resampled into 5000×5000 with ground resolution of 1 m, which is able to represent detailed substructures. The study area (Fig. 3) covering the city centre in Berlin, Germany shows many bright clusters of strong signals on structures that appear to be potential PS points. Two Google Earth (GE) images (ground truth) taken on September 12, 2010 and September 5, 2014 reveal that many building constructions occurred within the image acquisition period and are thus good examples for comparison with our test result. The spatiotemporal changes (Fig. 4) show where and when the structures disappeared and emerged. We focus our following discussions on the patches 1 to 5 for different applications of urban monitoring.

We first explore the construction events around Berlin Central Station (Fig. 5). The office complex of Federal Ministry of the Interior (area 1) had been constructed in the second half of 2013. A series of construction events is present in area 2. The quad-square structures were removed at the early stage. The upper-left hotel had been erected over time by 2013; another hotel and two office buildings cannot be detected because the constructions were still in progress. Certain new surface substructures in areas 3, 4, and 7 are able to be detected by our method; in contrast, these substructures are hardly identified from the GE images. Another construction event, which is also difficult to be recognized from the GE images, is illustrated in area 5 where the bridge was renovated during a couple of early months in 2013. Area 6 displays two new office buildings that were constructed at different schedules. The right building had been completed earlier, giving rise to a building-shaped pattern formed by clustered EBC points. In contrast, the construction progress on the left building was slower as only sparse EBC points appeared in late 2013.



Fig. 3: Mean TSX image over study area. Patches 1 to 5 are used for in-depth analysis

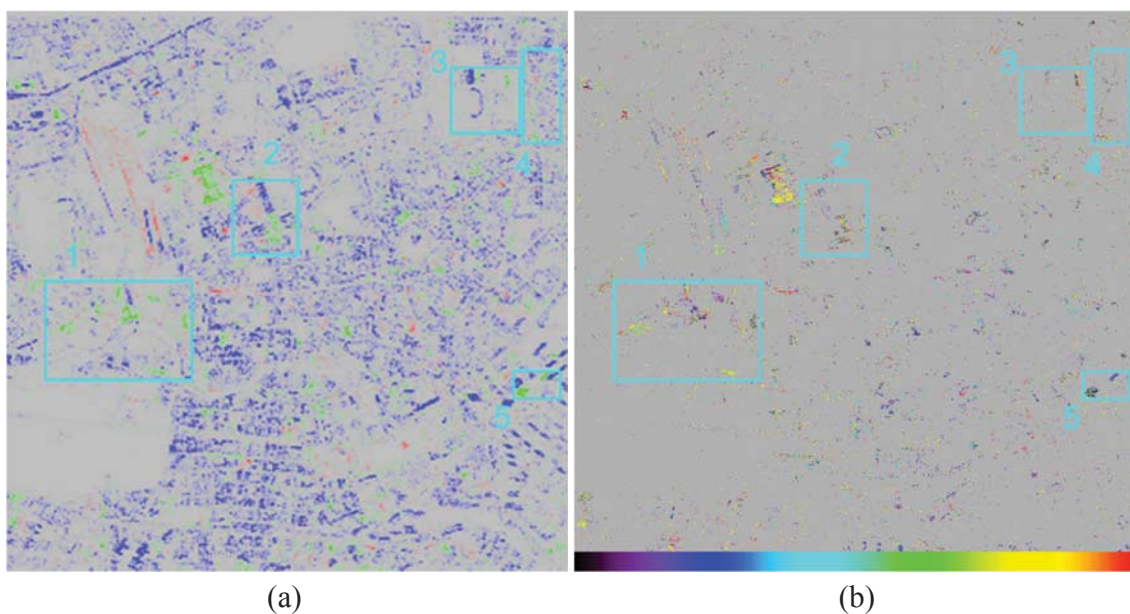


Fig. 4: Spatiotemporal change detection result. Patches 1 to 5 are used for in-depth analysis. (a) Steady, disappearing, and emerging structures represented by PS (blue, 41277/km²), DBC (red, 2200/km²), and EBC (green, 7180/km²) points. (b) Disappearance and emergence dates: black to red, earliest to latest in 2013.

The second example (Fig. 6) is about monitoring a business district, in which building changes are usually frequent and need cost-effective surveillance schemes. In the early 2013, the buildings in areas 1 and 2 were demolished; and the main structures of the new buildings appeared in areas 3 to 7 and their constructions continued to the end of 2013. Since the second half of 2013, certain substructures had been added to the German Railway's office complex in area 8. These additions seem vague in the GE images but are clearly revealed in the spatiotemporal change detection result.

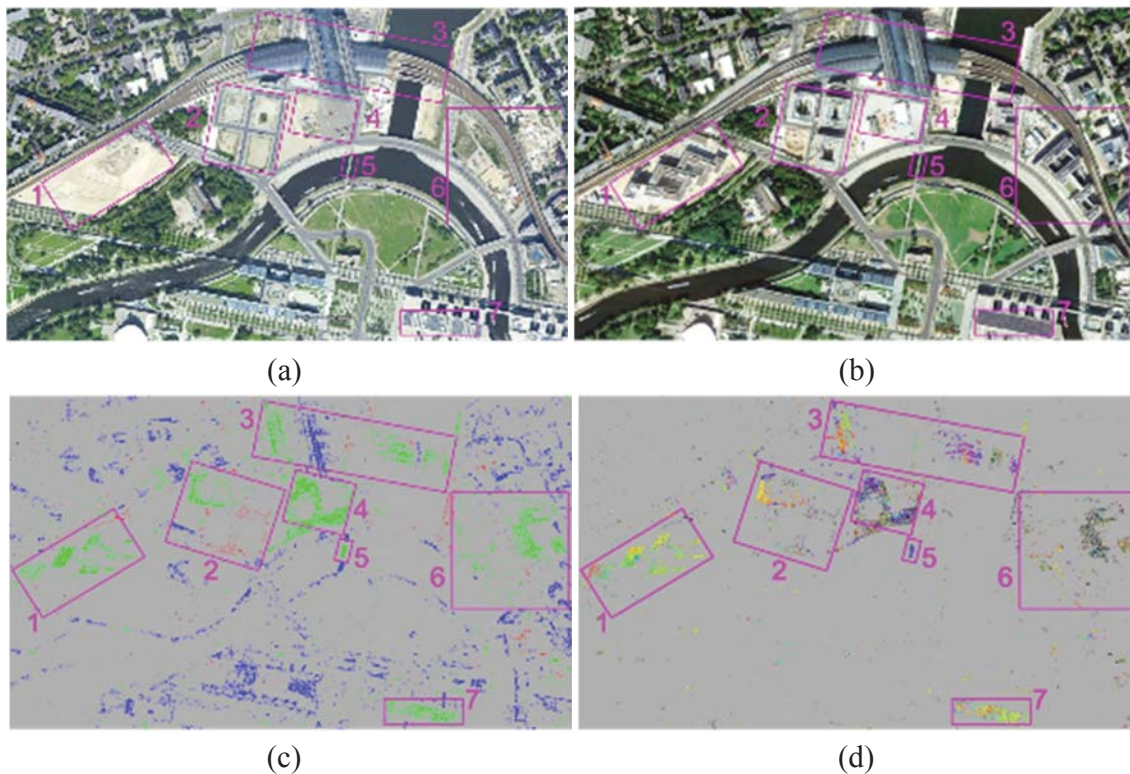


Fig. 5: Construction monitoring in patch 1 (Fig. 3) around Berlin Central Station. Areas 1 to 7 are used for in-depth analysis. GE images were acquired on (a) September 12, 2010 and (b) September 5, 2014. (c) and (d): spatiotemporal change detection result in patch 1 (Fig. 4).

Monitoring sports facilities is demanded for citizen safety. In case of structural damages, DBC points might be found at certain times. In Olympiastadion Berlin (Fig. 7), two detected change events include a structural renovation (area 1) on the upper-left arena and an erection of a new building (area 2) beside a sports playground. Most importantly, the upper-left arena and lower-left stadium seem steady without structural damages as the intensive PS points are found on them rather than DBC points. If necessary, the line-of-sight velocities of the PS points can be utilized for investigation on structural deformation.

Traffic infrastructure monitoring is useful for transportation management especially in busy cities and extending human settlements. For example, our technique detects a new elevated metro line across a couple of blocks that was under construction in 2013 (Fig. 8). The main structure had been accomplished in the early dates, followed by some partial substructures.

The last example demonstrates construction monitoring of single high-rise buildings (Fig. 9). The left building's main structure (area 1) had been erected in early 2013 and the remaining substructures were later complemented over time. In area 2, certain new storeys were built upon an existing building from low to high level in sequence along the magenta arrow (Fig. 9(d)).

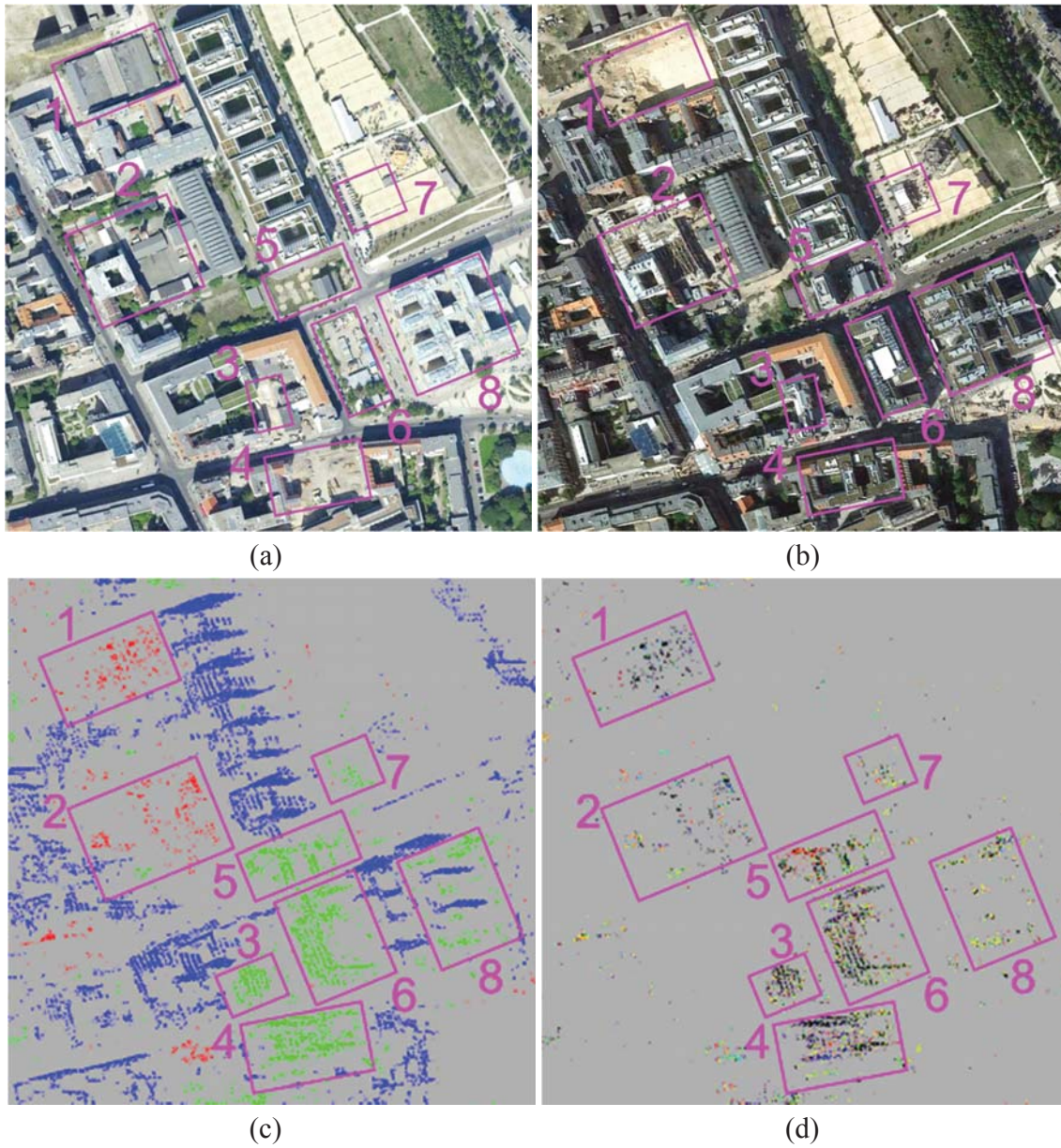


Fig. 6: Business district monitoring in patch 2 (Fig. 3). Areas 1 to 8 are used for in-depth analysis. GE images were acquired on (a) September 12, 2010 and (b) September 5, 2014. (c) and (d): spatiotemporal change detection result in patch 2 (Fig. 4).

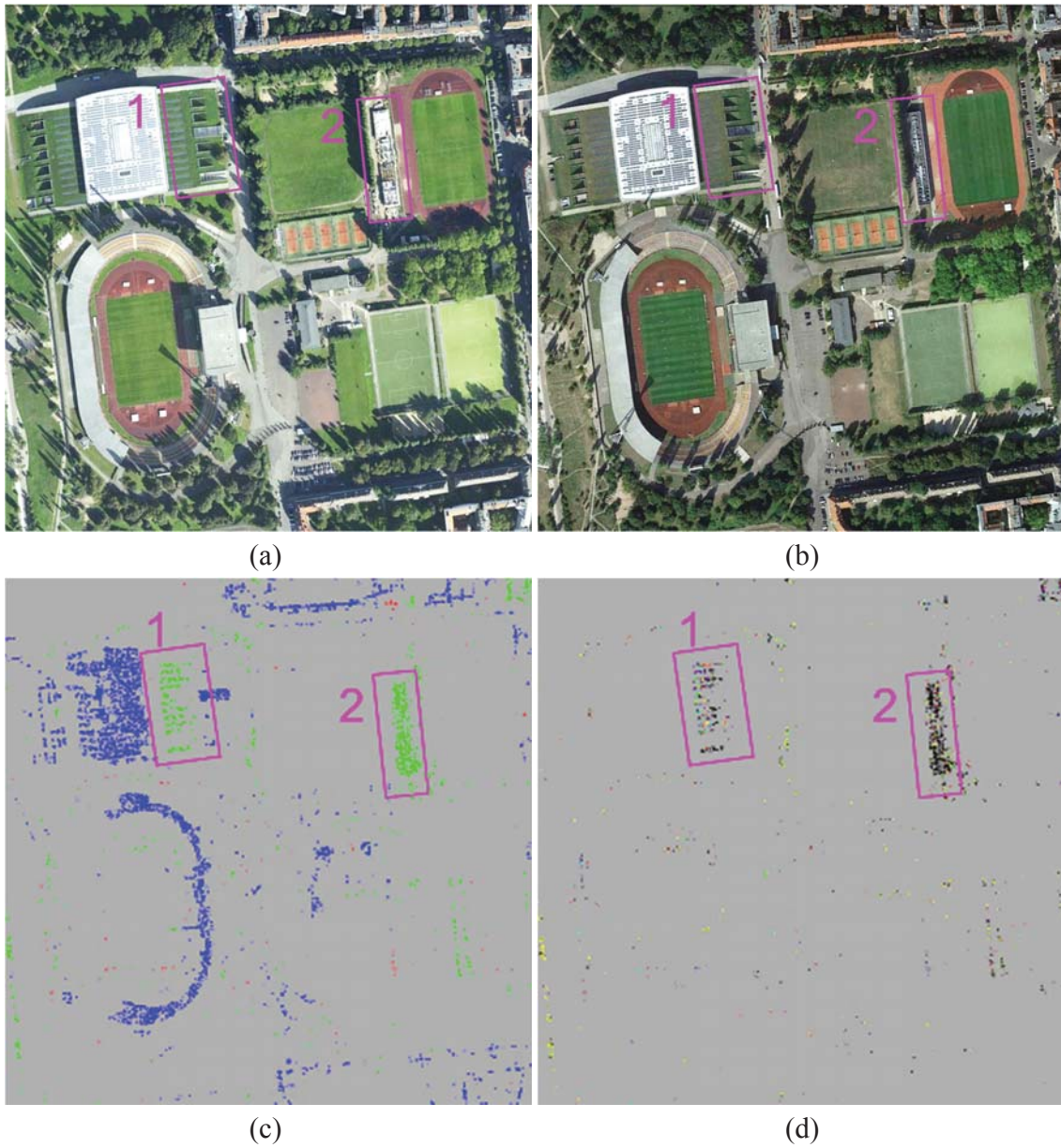


Fig. 7: Sports facility monitoring in patch 3 (Fig. 3). Areas 1 to 2 are used for in-depth analysis. GE images were acquired on (a) September 12, 2010 and (b) September 5, 2014. (c) and (d): spatiotemporal change detection result in patch 3 (Fig. 4).

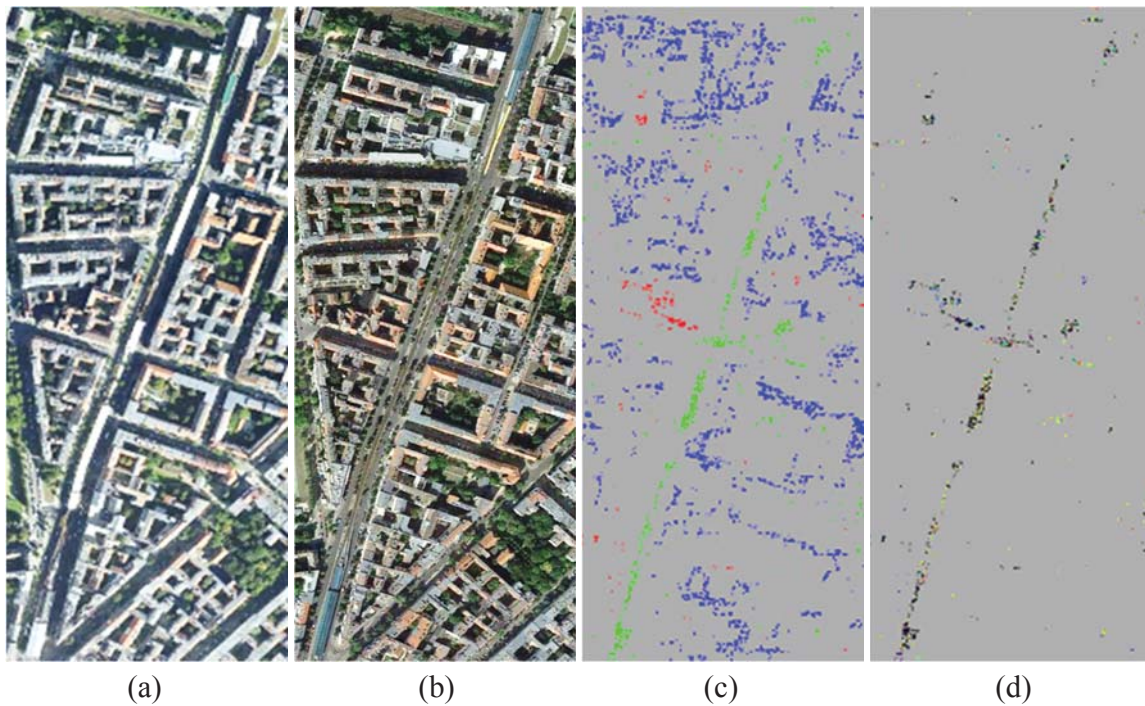


Fig. 8: Traffic infrastructure monitoring in patch 4 (Fig. 3). GE images were acquired on (a) September 12, 2010 and (b) September 5, 2014. (c) and (d): spatiotemporal change detection result in patch 4 (Fig. 4).

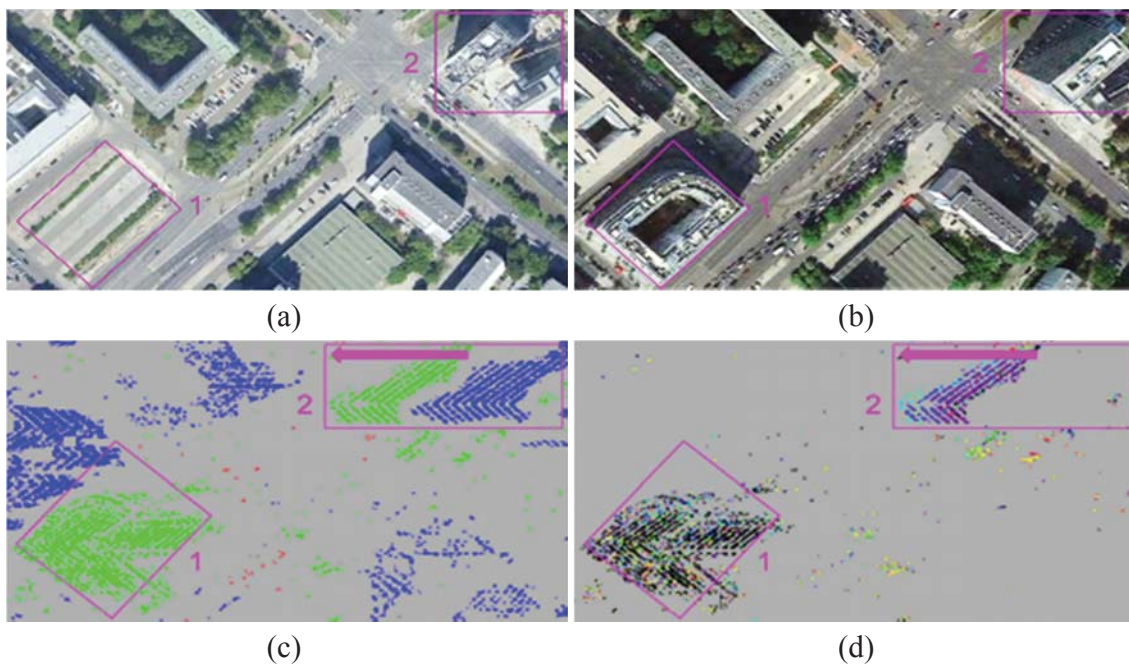


Fig. 9: Construction monitoring of single high-rise buildings in patch 5 (Fig. 3). Areas 1 to 2 are used for in-depth analysis. GE images were acquired on (a) September 12, 2010 and (b) September 5, 2014. (c) and (d): spatiotemporal change detection result in patch 5 (Fig. 4).

5 Conclusions

A novel time series analysis, 4DCDPSI, has been proven capable of detecting spatial changes (3D) and occurrence times (1D). In this study, we explore the feasibility and applicability of 4DCDPSI for urban monitoring. The aims of the five case studies are to monitor construction progress, business districts, sports facilities, traffic infrastructures, and single buildings. As to construction monitoring, three main construction types, i.e., demolition, erection, and renovation, can be distinctly recognized along with change times that substructures are added or removed. Such spatiotemporal change information is able to be derived from a business district characterized by frequent and intensive building changes. Our method can also provide detailed construction progress on single buildings. A typical example in our case studies detects some new storeys that were built upon an existing high-rise building from low to high level in sequence. In addition to building construction, we also demonstrate that construction monitoring of traffic infrastructures is also feasible by bringing an example of a newly-built elevated metro line. Last but not least, fortunately, we only find dense PS points instead of DBC points, which are regarded as structural damages, on the sports facilities. In summary, we have demonstrated that our technique can provide detailed and comprehensive information for urban monitoring.

6 Literatures

- BERARDINO, P., FORNARO, G., LANARI, R. & SANSOSTI, E., 2002: A New Algorithm for Surface Deformation Monitoring Based on Small Baseline Differential SAR Interferogram. *IEEE Transactions on Geoscience and Remote Sensing* **40**(11), 2375-2382.
- BRCIC, R. & ADAM, N., 2013: Detecting Changes in Persistent Scatterers. *Proceedings of IEEE IGARSS*, Melbourne, Australia, 117-120.
- CROSETTO, M., MONSERRAT, O., CUEVAS-GONZÁLEZ, M., DEVANTHÉRY, N. & CRIPPA, B., 2016: Persistent Scatterer Interferometry: A Review. *ISPRS Journal of Photogrammetry and Remote Sensing* **115**, 78-89.
- FERRETTI, A., PRATI, C. & ROCCA, F., 2000: Nonlinear Subsidence Rate Estimation Using Permanent Scatterers in Differential SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing* **38**(5), 2202-2212.
- FERRETTI, A., PRATI, C. & ROCCA, F., 2001: Permanent Scatterers in SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing* **39**(1), 8-20.
- FERRETTI, A., COLESANTI, C., PERISSIN, D., PRATI, C. & ROCCA, F., 2003: Evaluating The Effect of The Observation Time on The Distribution of SAR Permanent Scatterers. *Proceedings of FRINGE*, Frascati, Italy.
- FERRETTI, A., FUMAGALLI, A., NOVALI, A., PRATI, C., ROCCA, F. & RUCCI, A., 2011: A New Algorithm for Processing Interferometric Data-stacks: SqueeSAR. *IEEE Transactions on Geoscience and Remote Sensing* **49**(9), 3460-3470.
- GAMBA, P., 2013: Human Settlements: A Global Challenge for EO Data Processing and Interpretation. *Proceedings of the IEEE* **101**(3), 570-581.
- HOOPER, A., ZEBKER, H., SEGALL, P. & KAMPES, B., 2004: A New Method for Measuring Deformation on Volcanoes and Other Natural Terrains Using InSAR Persistent Scatterers. *Geophysical Research Letters* **31**(23), 1-5.

- KAMPES, B. M., 2006: Radar Interferometry: Persistent Scatterer Technique. Springer.
- LANARI, R., MORA, O., MANUNTA, M., MALLORQUÍ, J. J., BERARDINO, P. & SANSOSTI, E., 2004: A Small Baseline Approach for Investigating Deformations on Full Resolution Differential SAR Interferograms. *IEEE Transactions on Geoscience and Remote Sensing* **42**(7), 1377-1386.
- MARIN, C., BOVOLO, F. & BRUZZONE, L., 2015: Building Change Detection in Multitemporal Very High Resolution SAR Images. *IEEE Transactions on Geoscience and Remote Sensing* **53**(5), 2664-2682.
- Novali, F., Basilico, M., Ferretti, A., Prati, C. & Rocca, F., 2004: Advances in Permanent Scatterer Analysis: Semi and Temporary PS. Proceedings of EUSAR, Berlin, Germany.
- YANG, C., KENDUIYWO, B. K. & SOERGEL, U., 2016: 4D Change Detection Based on Persistent Scatterer Interferometry. Proceedings of ICPR&PRRS, Cancun, Mexico.

Automatische Detektion von Bombenkratern in Kriegsluftbildern mittels markierter Punktprozesse

CHRISTIAN KRUSE¹, HAJNALKA NEUBERGER², FRANZ ROTTENSTEINER¹, THORSTEN HOBERG³,
MARCEL ZIEMS³, JULIA HUTH³ & CHRISTIAN HEIPKE¹

Zusammenfassung: In diesem Beitrag wird eine Methode zur automatischen Detektion von Bombenkratern in Kriegsluftbildern aus dem 2. Weltkrieg vorgeschlagen. Hierbei wird das Objektmodell für die Bombenkrater durch Ellipsen repräsentiert. Mithilfe eines probabilistischen Ansatzes, basierend auf markierten Punktprozessen, wird die wahrscheinlichste Konfiguration von Objekten innerhalb der Szene bestimmt. Durch das Hinzufügen und Entfernen neuer Objekte zur bzw. von der aktuellen Konfiguration sowie Modifikationen der Ellipsenparameter werden zufällig neue Objektkonfigurationen erzeugt. Anschließend wird jede Konfiguration anhand einer Energiefunktion bewertet. Dabei werden hohe Gradientenmagnituden entlang des Ellipsenrandes favorisiert und überlappende Ellipsen bestraft. Das globale Optimum der Energie, welches die Übereinstimmung mit einem vordefinierten Modell beschreibt, liefert das Reversible Jump Markov Chain Monte Carlo Sampling in Kombination mit Simulated Annealing. Unsere Ergebnisse zeigen das generelle Potential der Methode zur automatischen Detektion von Bombenkratern in einem heterogenen Bildbestand auf.

1 Einleitung

Das Bundesland Niedersachsen war Ziel zahlreicher Luftangriffe der alliierten Streitkräfte während des zweiten Weltkriegs. Die Nachwirkungen dieser Angriffe sind bis heute präsent. Zahlreiche Blindgänger lagern noch immer im Erdreich. Insbesondere bei Erdarbeiten ist die Gefahr einer Detonation der Blindgänger gegeben, auch sind Vorfälle ohne äußere Einwirkung bekannt.

Eine zentrale Aufgabe des Kampfmittelbeseitigungsdienstes Niedersachsens ist die Auswertung von Kriegsluftbildern hinsichtlich Kriegsbelastungen, insbesondere der Identifikation von Bombenblindgängerverdachtspunkten. Hierzu werden Stereobildpaare einzeln ausgewertet. Dies bringt, trotz der Einschränkung auf besonders gefährdete oder anderweitig relevante Flächen, einen hohen Bearbeitungsaufwand mit sich. Zu einer verbesserten Prozesssteuerung und Dokumentation wird aktuell das Kampfmittelinformationssystem Niedersachsen (KISNi) aufgebaut, welches künftig zusätzliche Analysemöglichkeiten bieten wird. Für viele Fragestellungen ist es ausreichend, über flächendeckende Informationen hinsichtlich des grundsätzlichen Vorkommens von Kriegsbelastungen in Form von „Belastungskarten“ zu verfügen. Belastungskarten können erzeugt werden, wenn mit hoher Sicherheit ein großer

¹ Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Straße 1, D-30167 Hannover, E-Mail: [kruse, rottensteiner, heipke]@ipi.uni-hannover.de

² Budapest University of Technology and Economics, Department of Photogrammetry and Geoinformatics, Műegyetem rkp. 3, H-1111 Budapest, E-Mail: neuberger.hajnalka@epito.bme.hu

³ Landesamt für GeoInformation und Landesvermessung Niedersachsen, Podbielskistraße 331, D-30659 Hannover, E-Mail: [thorsten.hoberg, marcel.ziems, julia.huth]@lgin.niedersachsen.de

Prozentsatz sämtlicher Belastungen detektiert werden kann. Für die kosteneffiziente Erstellung einer solchen Belastungskarte ist eine automatische Erkennung von Hinweisen auf Kriegsbelastung, vor allem von Bombenkratern (Abb. 1), in Kriegsluftbildern unerlässlich.

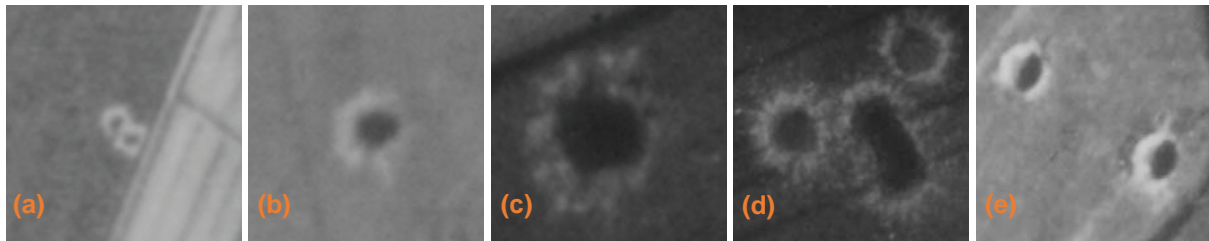


Abb. 1: Bombenkrater prägen sich in den Kriegsluftbildern aufgrund von (a-c) verschiedenen Maßstäben und Größen sowie (d, e) unterschiedlichem Aussehen verschiedenartig aus.

In der Bildanalyse werden für derartige Aufgaben immer häufiger probabilistische Modelle genutzt, bei welchen Vorwissen in Form von Wahrscheinlichkeiten integriert wird. Bekannte Ansätze stellen Markov Random Fields (GEMAN & GEMAN 1984) sowie Conditional Random Fields (KUMAR & HEBERT 2006) dar. Durch die Modellierung von Kontext kann Wissen über die Objekte im Bild integriert werden. Beispielsweise lassen sich auf diese Weise ähnliche Klassen für Pixel in einer lokalen Nachbarschaft begünstigen (LI 2009). Jedoch ist es schwierig, globalere Bedingungen der Objekte, wie beispielsweise hinsichtlich der Form, zu integrieren. Hier setzen markierte Punktprozesse (DALEY & VERE-JONES 2003; DESCOMBES & ZERUBIA 2002) an, welche ein starkes Objektmodell nutzen. Das Sampling liefert die global optimale Konfiguration an Objekten eines bestimmten Typs.

Die Integration von Modellwissen kann auf verschiedene Arten erfolgen. In der Regel werden einfache geometrische Primitive, die über eine geringe Anzahl an Parametern beschrieben werden können, zur Repräsentation der Objekte genutzt. Ein vielfach in der Literatur verwendetes Objekt sind Rechtecke, mit deren Hilfe z.B. Gebäude in der Szene extrahiert werden können. CHAI et al. (2012) nutzen hierfür einen kombinierten Ansatz aus Markov Random Fields und markierten Punktprozessen auf Basis von Satellitenbildern. Während auf der High-Level-Ebene Gebäude durch Rechtecke repräsentiert und in das Verfahren der markierten Punktprozesse integriert werden, wird auf der Low-Level-Ebene ein Markov Random Field genutzt, um das Aussehen der Gebäude mittels Farbhistogrammen zu modellieren. Gleichermaßen modellieren TOURNAIRE et al. (2010) und BREDIF et al. (2013) Objekte als Rechtecke im Zusammenhang mit markierten Punktprozessen. Die Datengrundlage bilden digitale Oberflächenmodelle, innerhalb derer die Rechtecke zufällig verteilt werden. Damit ein Rechteck in die Objektkonfiguration aufgenommen wird, müssen hohe Gradientenmagnituden der Höhen am Rechteckrand vorliegen. Ferner wird in diesem Modell gefordert, dass sich Rechtecke nicht überlappen dürfen. Anders als bei CHAI et al. (2012) sind starke Überlappungen nicht unmöglich, sondern unwahrscheinlich. Ziel dieser Formulierung ist es, starke Überlagerungen von Objekten zu vermeiden, wohingegen geringfügige Überlappungen wahrscheinlicher sind. Darüber hinaus werden Rechtecke verwendet, um Fassaden an Gebäuden auf Grundlage von rektifizierten Bildern zu interpretieren (WENZEL & FÖRSTNER 2016) oder Fahrzeuggruppen aus Laserscandaten zu extrahieren (BÖRCS & BENEDEK 2015). Bei ersterem Ansatz werden im Gegensatz zu den bisher genannten Verfahren die Statistiken von typischen

Konfigurationen der Fassadenobjekte (Fenster, Eingänge) mittels Trainingsdaten gelernt. Neben dieser Form von Objekten werden z.B. Zylinder zur Personendetektion in 3D (UTASI & BENEDEK 2011) und Verteilungsfunktionen zur Modellierung von Wellenformen (MALLET et al. 2010) genutzt. Ferner erlauben Segmente Linien-Netzwerke, wie Straßen oder Blutgefäße, zu extrahieren (LACOSTE et al. 2005; SUN et al. 2007). Darüber hinaus werden markierte Punktprozesse zur Detektion von Flamingos (DESCAMPS et al. 2008), Saatgutprodukten (DUBOSCLARD et al. 2014) und Planetenkratern (TROGLIO et al. 2010) herangezogen. In allen Fällen sind die Objekte durch Ellipsen modelliert. LAFARGE et al. (2010) verwenden verschiedenste Objektmodelle, welche eine große Flexibilität hinsichtlich verschiedenster Anwendungen ermöglichen. So werden z.B. Kreise, Rechtecke und Linien genutzt, um Baumkronen, Grundrisse von Gebäuden sowie Straßennetzwerke aus Bildern zu extrahieren. Den Erhalt netzwerkartiger Strukturen, wie beispielsweise von Straßennetzwerken, erlauben Verfahrensweisen auf Basis von Graphen, da jeder Knoten entweder zu einem Kreuzungs- oder Endpunkt gehört. SCHMIDT et al. (2016) nutzen diese graphenbasierte Repräsentation um Flussnetzwerke in digitalen Geländemodellen zu finden.

Die aufgeführten Beiträge zeigen das Potential stochastischer Modelle basierend auf markierten Punktprozessen in verschiedensten Anwendungsfeldern der Bildanalyse auf. Neben der Bestimmung der global optimalen Objektkonfiguration profitieren sie von ihrer Flexibilität hinsichtlich der Integration von Wissen über die Objekte und deren Beziehung zu anderen Objekten. Darüber hinaus ist die Anzahl an Objekten in der Szene während des Samplings variabel. Wir wollen diese Vorteile nutzen, um Bombenkrater in Kriegsluftbildern zu detektieren. Das hierbei genutzte Modell sind Ellipsen, deren Konfiguration während des Samplings optimiert wird. Zur Bewertung der Objektkonfiguration werden hohe Magnituden der Grauwertgradienten am Ellipsenrand favorisiert und die Überlappung von Ellipsen bestraft. Im Folgenden werden zunächst die mathematischen Grundlagen der stochastischen Optimierung mithilfe von markierten Punktprozessen beschrieben (Abschnitt 2). In Abschnitt 3 stellen wir das Modell vor, welches zur Detektion der Bombenkrater genutzt werden soll. Daraufhin werden Experimente und Ergebnisse auf Grundlage von Kriegsluftbildern gezeigt (Abschnitt 4). Abschließend werden in Abschnitt 5 eine Zusammenfassung und ein Ausblick für künftige Arbeiten gegeben.

2 Mathematische Grundlagen

2.1 Markierte Punktprozesse

Punktprozesse (für Details sei auf DALEY & VERE-JONES (2003) und DESCOMBES (2013) verwiesen) gehören zu den stochastischen Prozessen und ermöglichen die mathematische Beschreibung von Zufallsgeschehen. Anders als bei einem deterministischen Ansatz erfolgt die Modellierung in Form von Wahrscheinlichkeiten. Es existieren verschiedene Arten von Punktprozessen, wobei die Realisierung des Punktprozesses aus isolierten Punkten in einem mathematischen Raum besteht. Der Parameterraum weist hierbei häufig einen zeitlichen oder räumlichen Bezug auf. In letzterem Fall wird von einem räumlichen Punktprozess gesprochen, welcher die Grundlage der Verfahren zur Extraktion von Objekten in Fernerkundungsdaten bildet und somit auch bei uns Anwendung findet.

Punktprozesse haben die Aufgabe, aus oftmals großen Datenmengen die wahrscheinlichste Konfiguration von Objekten eines bestimmten Typs zu finden, welche ein beobachtetes Phänomen am besten charakterisiert. Sie bieten die Möglichkeit, eine Szene mittels einer ungeordneten Menge an Punkten innerhalb einer begrenzten Region $F \subset R^k$ zu beschreiben (bei uns ist $k = 2$). Dabei ist ein Punkt über seine Position $p_i = (x_i, y_i)$ charakterisiert. Da häufig Punkte allein eine Szene nur schwer beschreiben können, wurden Ansätze entwickelt, die jeden Punkt um zusätzliche Parameter m_i erweitern. Auf diese Weise kann der Punkt mit einem geometrischen Objekt $u_i = (p_i, m_i)$ in Verbindung gebracht werden: das Objekt u_i wird durch die Position p_i fixiert (oft der Massenmittelpunkt) und die Merkmale m_i , welche sich von Punkt zu Punkt verschieden ausprägen können, enthalten weitere Informationen zum Objekt. Durch das Hinzufügen von Objektparametern zu jedem Punkt entsteht ein markierter Punktprozess, welcher als ein stochastisches Modell von Konfigurationen einer unbekannt Anzahl von Objekten des Typs $u_i = (p_i, m_i)$ in F aufgefasst werden kann.

Das Modell des homogenen Poisson-Punktprozesses, welcher als grundlegender Punktprozess aufgefasst werden kann, nimmt eine rein zufällige Verteilung von Objekten im Raum an. Ferner stehen die Objekte in keiner Beziehung zueinander. Die Wahrscheinlichkeit $P_\lambda(n)$ für die Anzahl der Objekte n folgt einer diskreten Poisson-Verteilung

$$P_\lambda(n) = \frac{\lambda(F)^n}{n!} \cdot e^{-\lambda(F)}. \quad (1)$$

Der Parameter λ beschreibt dabei die erwartete Anzahl an Objekten innerhalb von F und wird häufig als Intensitätsparameter bezeichnet. Die Wahrscheinlichkeit für die Positionen der Objekte folgt einer Gleichverteilung und die Objekte sind stochastisch unabhängig in F verteilt. In der Praxis trifft die Annahme einer völligen Zufälligkeit häufig nicht zu, da Abhängigkeiten zwischen den Objekten bestehen. Es werden komplexere Modelle gefordert, mit welchen die Qualität der Objektkonfiguration gemessen werden kann. Hierfür kann eine Wahrscheinlichkeitsdichte $h(\cdot)$ des markierten Punktprozesses formuliert werden. Erreicht werden kann dies mit Bezug auf einen Referenz-Punktprozess, welcher üblicherweise als Poisson-Punktprozess definiert ist.

In unserem Ansatz wird $h(\cdot)$ durch eine Gibbs-Energie $U(\cdot)$ in der Form $h \propto \exp - U(\cdot)$ ausgedrückt. Die Gibbs-Energie setzt sich aus zwei Teilen, der Daten- und der Priorenergie, zusammen. Der relative Einfluss kann durch einen Parameter $\beta \in [0, 1]$ modelliert werden, um eine den Daten angepasste Gewichtung vornehmen zu können

$$U(\cdot) = \beta \cdot U_D(\cdot) + (1 - \beta) \cdot U_P(\cdot). \quad (2)$$

Hierbei entspricht $U_D(\cdot)$ der Datenenergie, welche die Kongruenz der Objektkonfiguration mit den Eingangsdaten vergleicht. Interaktionen zwischen den Objekten werden innerhalb von $U_P(\cdot)$ berücksichtigt. Die optimale Objektkonfiguration $u^* = \{u_1, \dots, u_n\}$ kann bestimmt werden, indem die Wahrscheinlichkeitsdichte $h(\cdot)$ maximiert wird, also $u^* = \arg \max h(\cdot)$. Diese Vorgehensweise entspricht ebenfalls der Konfiguration, welche die Gibbs-Energie $U(\cdot)$ minimiert, das heißt $u^* = \arg \min U(\cdot)$. Nun ist die Wahrscheinlichkeitsdichte $h(\cdot)$ in der Regel multi-modal und darüber hinaus in einem Konfigurationsraum definiert, welcher eine variable Dimension besitzt. Aus diesem Grund wird zum Schätzen des globalen Minimums

häufig ein Reversible Jump Markov Chain Monte Carlo Sampler in Kombination mit Simulated Annealing verwendet, um eine Approximation von u^* zu finden.

2.2 Reversible Jump Markov Chain Monte Carlo Verfahren

Mithilfe von Markov-Ketten (z.B. ANDRIEU et al. 2003) können zufällige Zustandsänderungen eines Systems modelliert werden. Der Poisson-Prozess (Abschn. 2.1) stellt ein Beispiel einer Markov-Kette dar, in welchem die Anzahl der Objekte der Poisson-Verteilung (1) folgt. Die Ketten besitzen die Markov-Eigenschaft, welche besagt, dass die Zukunft eines stochastischen Prozesses unabhängig von der Vergangenheit bei gegebenem gegenwärtigen Zustand ist. Das heißt, in einer Sequenz von Zuständen besitzt jede Stichprobe X_t eine Wahrscheinlichkeitsverteilung, die nur vom vorherigen Wert X_{t-1} abhängt.

Markov Chain Monte Carlo Methoden (METROPOLIS et al. 1953; HASTINGS 1970) können nicht mit Zustandsräumen verschiedener Dimensionen umgehen. Hier setzen die Reversible Jump Markov Chain Monte Carlo (RJMCMC) Verfahren an, welche die Modellierung von Szenen mit einer unbekanntem Anzahl an Objekten sowie Änderungen der Dimension der Parameter zwischen zwei Sampling-Schritten erlauben. Diese Erweiterung wurde erstmals von GREEN (1995) vorgeschlagen und wird durch die Definition einer Menge an Bewegungen bzw. Änderungen, den sogenannten reversiblen Sprüngen (*reversible jumps*), erreicht. Reversibilität bedeutet, dass es möglich ist, jede Änderung rückgängig zu machen, sodass man immer zu einem früheren Zustand zurückkehren kann. In jeder Iteration t schlägt der Sampler eine Änderung der aktuellen Objektkonfiguration aus jener vordefinierten Menge an Bewegungen (*jumps*) vor. Dabei existiert zu jeder Art der Änderung eine Dichtefunktion Q_m , welche auch als Kernel (Kern) bezeichnet wird. Dieser Prozess führt von einer Objektkonfiguration X_t zu einer neuen Konfiguration X_{t+1} gemäß einer Wahrscheinlichkeit $Q_m(X_t \rightarrow X_{t+1})$. Die Startkonfiguration X_0 kann aufgrund der Eigenschaften der Markov-Ketten (TIERNEY 1996) zufällig gewählt werden. Die Änderung bzw. neue Konfiguration wird mit einer gewissen Akzeptanzwahrscheinlichkeit α , die von der Energievariation zwischen X_t und X_{t+1} abhängt, angenommen

$$\alpha = \min\left(1, \frac{Q_m(X_{t+1} \rightarrow X_t)}{Q_m(X_t \rightarrow X_{t+1})} \cdot \exp\left(-\frac{U(X_{t+1}) - U(X_t)}{T_t}\right)\right). \quad (3)$$

In (3) beschreibt das Kernel-Verhältnis $Q_m(X_{t+1} \rightarrow X_t)/Q_m(X_t \rightarrow X_{t+1})$ das Verhältnis der Wahrscheinlichkeiten für die Änderung der Konfiguration von X_{t+1} zu X_t und umgekehrt. $U(X_{t+1})$ und $U(X_t)$ sind die Gibbs-Energie (2) der neuen bzw. aktuellen Objektkonfiguration und T_t bezeichnet die Temperatur des Simulated Annealing der Iteration t , welche in Abschnitt 2.3 beschrieben wird.

2.3 Simulated Annealing

Um das Optimum der Energie zu finden, wird das RJMCMC Verfahren mit Simulated Annealing kombiniert. Der Algorithmus bildet den Abkühlungsprozess von Stoffen in der Festkörperphysik nach und lässt sich auf das heuristische Optimierungsverfahren des Simulated Annealing übertragen (METROPOLIS et al. 1953; KIRKPATRICK et al. 1983). Die Temperatur entspricht einer Wahrscheinlichkeit, mit der sich ein Zwischenergebnis der Optimierung auch verschlechtern darf. Wesentlicher Vorteil hiervon ist, dass ein lokales Optimum wieder verlassen werden kann.

T_t strebt mit zunehmender Iteration gegen Null, während t gegen unendlich strebt. Zu Beginn des Algorithmus, wenn die Temperatur hoch ist, ist das Verfahren nicht selektiv. Bei abnehmender Temperatur wird der Prozess selektiver, das heißt Konfigurationen mit einer hohen Wahrscheinlichkeitsdichte werden bevorzugt. Der Abkühlungsprozess kann über eine logarithmische Abnahme realisiert werden, welche theoretisch Konvergenz zum globalen Optimum für jede beliebige Anfangskonfiguration X_0 gewährleistet. Aufgrund der hohen Berechnungszeiten wird in der Praxis üblicherweise ein schnelleres geometrisches Schema zur Abkühlung verwendet. Dieses liefert ein approximatives Resultat, welches nahe an der optimalen Lösung liegt (SALAMON et al. 2002; VAN LAARHOVEN & AARTS 1987).

3 Methodik

In dieser Arbeit werden Bombenkrater als Ellipsen modelliert. Mithilfe eines markierten Punktprozesses soll die optimale Konfiguration von solchen Objekten in der Szene gefunden werden. Dabei wird eine globale Energiefunktion minimiert, deren Optimierung iterativ durch das RJMCMC Sampling in Kombination mit Simulated Annealing erfolgt (Abschn. 2.2, 2.3). Während des Optimierungsprozesses verändert sich die Objektkonfiguration kontinuierlich und wird basierend auf der Übereinstimmung mit einem vordefinierten Modell evaluiert. Wir favorisieren in diesem Modell hohe Gradientenmagnituden entlang des Ellipsenrandes und bestrafen sich überlappende Ellipsen.

Im ersten Schritt wird das Objektmodell formuliert, welches die Bombenkrater im Bild repräsentiert (Abschn. 3.1). In Abschnitt 3.2 werden die möglichen Änderungen der Objektkonfiguration eingeführt. Nachfolgend wird die Energiefunktion definiert, die es während des Samplings zu minimieren gilt (Abschn. 3.3). Abschließend wird in Abschnitt 3.4 eine Methode zur Einschränkung des Suchbereiches im Bild beschrieben, die dem Verfahren der markierten Punktprozesse vorangestellt wird.

3.1 Objektmodell

Das für die Detektion der Bombenkrater verwendete Modell besteht aus einem markierten Punktprozess von Ellipsen. Wie beispielsweise bei PERRIN et al. (2006) wird jede Ellipse über ihre zweidimensionale Position $P(x, y)$ und drei Merkmale $M(a, b, \theta)$ beschrieben. Das resultierende 5-Tupel (x, y, a, b, θ) nimmt im Zustandsraum Z die Werte

$$Z = P \times M = [0, X] \times [0, Y] \times [a_m, a_M] \times [b_m, b_M] \times [0, \pi[\quad (4)$$

an. Hierbei sind (x, y) die Koordinaten des Ellipsenzentrums, X und Y repräsentieren die Breite und Höhe des Bildes. Der minimale und maximale Wert der großen bzw. kleinen Halbachse wird durch $[a_m, a_M]$ bzw. $[b_m, b_M]$ begrenzt. Die Orientierung der Ellipsen wird über den Winkel $\theta \in [0, \pi[$ beschrieben, welche im Uhrzeigersinn relativ zur positiven x-Achse definiert ist (Abb. 2). Wir fordern $a \geq b$.

3.2 Änderungen der Objektkonfiguration

Um eine optimale Objektkonfiguration zu finden, wird die Konfiguration aus Ellipsen generiert und den Daten iterativ durch ein RJMCMC Sampling in Kombination mit dem Simulated

Annealing angepasst. Die Methode stellt vier Änderungsarten der Konfiguration bereit, wobei die Modifikationen mithilfe zugehöriger Kernel ermöglicht werden. Zum einen erlauben *Geburts- und Todes-Kernel*, Q_B bzw. Q_D , mit zugehöriger Vorschlagswahrscheinlichkeit $p_{Q_B} = p_{Q_D}$, das Hinzufügen bzw. Entfernen eines Objektes zur bzw. von der aktuellen Objektkonfiguration. Für den Fall einer Geburt (Abb. 3a) werden die Position, Orientierung und zwei Halbachsen einer neuen Ellipse gemäß (4) generiert. Für jeden Parameter wird eine Gleichverteilung angenommen. Beim Todesevent wird eine zufällig gewählte Ellipse aus der bestehenden Objektkonfiguration entfernt (Abb. 3b). Das Kernel-Verhältnis aus Gleichung (3) berücksichtigt die Wahrscheinlichkeit für die Änderung der Konfiguration von X_t zu X_{t+1} und umgekehrt. Wie SCHMIDT et al. (2016) modellieren wir das Kernel-Verhältnis des Geburtsereignisses durch

$$\frac{Q_{BD}(X_{t+1} \rightarrow X_t)}{Q_{DB}(X_t \rightarrow X_{t+1})} = \frac{p_D}{p_B} \cdot \frac{\lambda}{n}, \quad (5)$$

wobei p_B und p_D der Wahrscheinlichkeit für die Auswahl eines Geburts- bzw. Todesevents entsprechen. Die aktuelle Anzahl an Ellipsen innerhalb der Objektkonfiguration wird durch n repräsentiert und der Poisson-Parameter λ beschreibt die erwartete Anzahl an Objekten im Bild. Beim Todesereignis entspricht das Kernel-Verhältnis der inversen Geburtsrate.

Darüber hinaus können die Parameter eines Objektes einer Konfiguration modifiziert werden. Mithilfe des *Translations-Kernels* Q_T wird eine zufällig gewählte Ellipse von ihrer aktuellen Position verschoben. Hierfür wird zufällig ein Verschiebungsvektor in der lokalen Nachbarschaft auf Basis eines vorgegebenen Intervalls generiert (Abb. 4a). Die drei Merkmale der Ellipse lassen sich durch den *Merkmalvariations-Kernel* Q_M verändern. Hierfür wählen wir zufällig eine Ellipse aus der aktuellen Konfiguration aus, von welcher die kleine und große Halbachse sowie die Orientierung innerhalb vordefinierter Schwellwerte modifiziert werden (Abb. 4b). Diesen beiden Änderungsarten sind die Vorschlagswahrscheinlichkeiten p_{Q_T} bzw. p_{Q_M} zugehörig. Sowohl für die Translation als auch für die Merkmalsvariation wird das Kernel-Verhältnis auf 1 gesetzt

$$\frac{Q_T(X_{t+1} \rightarrow X_t)}{Q_T(X_t \rightarrow X_{t+1})} = \frac{Q_M(X_{t+1} \rightarrow X_t)}{Q_M(X_t \rightarrow X_{t+1})} = 1. \quad (6)$$

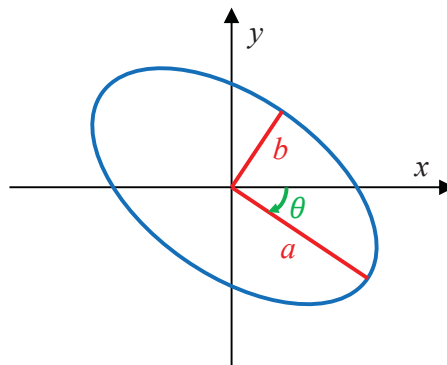


Abb. 2: Unser Objektmodell ist eine Ellipse, welche über ihre Position (x, y) und drei Merkmale (a, b, θ) beschrieben wird.

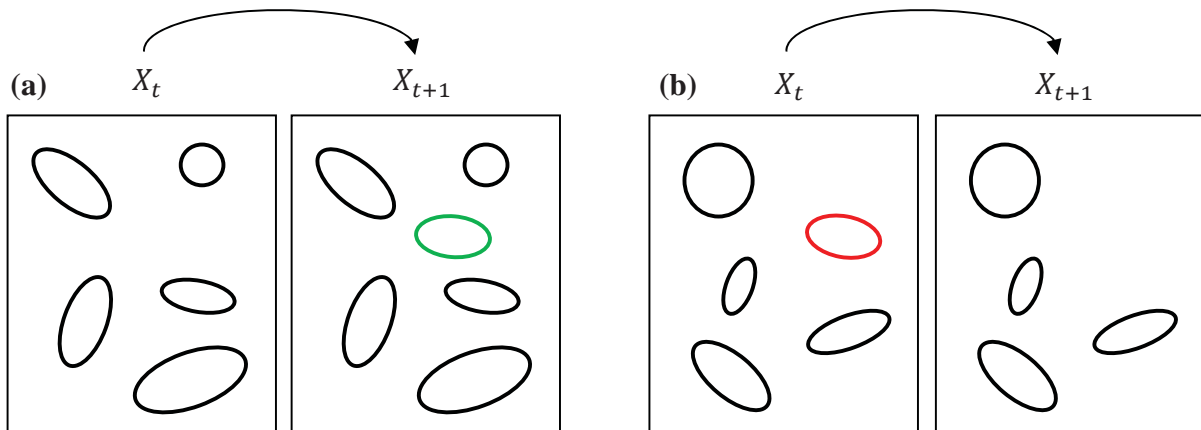


Abb. 3: Funktionsweise zweier Kernel, welche das Hinzufügen und Entfernen von Ellipsen zur bzw. von der Objektkonfiguration verdeutlichen.

- (a) *Geburts-Kernel*: Die neue Objektkonfiguration X_{t+1} wird um eine Ellipse (grün) erweitert.
- (b) *Todes-Kernel*: Aus der aktuellen Objektkonfiguration X_t wird eine Ellipse (rot) entfernt.

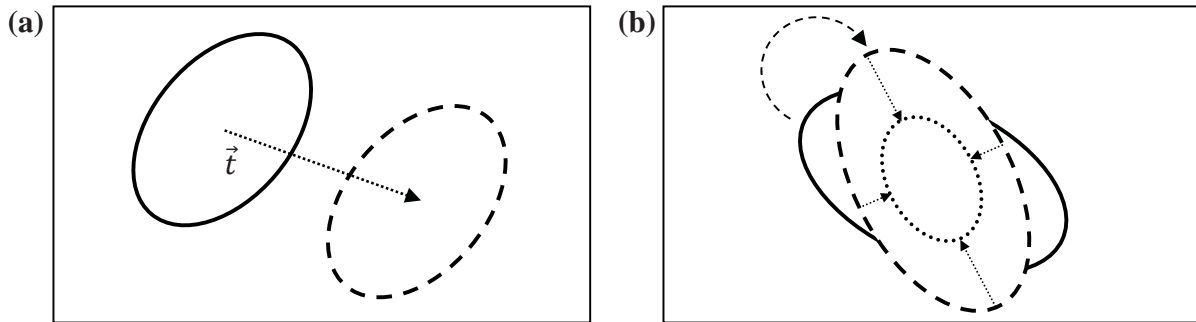


Abb. 4: Funktionsweise zweier Kernel, welche Änderungen an den Parametern des Objektes ermöglichen.

- (a) *Translations-Kernel*: Die Position der Ellipse wird um den Verschiebungsvektor \vec{t} verschoben.
- (b) *Merkmalsvariations-Kernel*: Die Merkmale der Ellipse, also die Orientierung und die Halbachsen, werden verändert.

3.3 Energiefunktion

In jedem Iterationsschritt wird die aktuelle mit der neuen Objektkonfiguration auf Basis der Gibbs-Energie (2) verglichen. Diese beschreibt die Übereinstimmung der Konfiguration mit einem vordefinierten Modell für die Bombenkrater. Ziel des Samplings ist es, diejenige Konfiguration zu finden, für welche die Gibbs-Energie minimal ist. In unserem Fall setzt sich die Gibbs-Energie aus zwei Teilen zusammen, der Daten- und der Priorenergie.

3.3.1 Datenenergie

Die Datenenergie $U_D(X_t)$ aus Gleichung (2) beschreibt die Übereinstimmung der Objektkonfiguration mit den Eingangsdaten. Bombenkrater sind durch überwiegend lokal dunklere Grauwerte im Vergleich zur Umgebung gekennzeichnet. Maßgeblich hierfür ist der durch die Sonne hervorgerufene Schattenwurf, dessen Form sich innerhalb der Krater hauptsächlich ellipsenförmig ausprägt (Abb. 1). Eine neu erzeugte oder veränderte Ellipse führt

zu einer Verminderung der Datenenergie, wenn entlang der Ränder der Ellipse vergleichsweise höhere Gradientenmagnituden auftreten, das Modell der Ellipse also besser zu dem erwarteten Schattenwurf passt. Wir bestimmen die Gradienten entlang des Ellipsenrandes ähnlich SCHMIDT et al. (2016) durch

$$U_D(X_t) = \sum_{e_j \in X_t} \left(c - \sum_{l_j=1}^{64} \frac{1}{n_j} \sum_{p_j=1}^{n_j} \nabla_{IMG_{p_j}}^b \right). \quad (7)$$

In Gleichung (7) ist $\nabla_{IMG_{p_j}}^b$ die Komponente des Grauwertgradienten am Randpixel p_j in Richtung des Normalenvektors der Ellipse e_j (Abb. 5). Zur Berechnung der Summe der gleichgewichteten Gradienten entlang des Ellipsenrandes werden n_j Pixel p_j herangezogen. Der Rand der Ellipse wird durch ein Polygon mit einer konstanten Anzahl von zurzeit 64 Punkten, welche durch Linien $l_j \in \{1, 2, \dots, 64\}$ miteinander verbunden sind, approximiert. In zukünftigen Untersuchungen könnte z.B. die Anzahl der Punkte, die den Ellipsenrand repräsentieren, abhängig von der Bildmaßstabszahl und von den in der Szene zu erwartenden Bombenkratergrößen automatisch gewählt werden. Ferner könnten Ellipsen, insbesondere solche, bei denen die große Halbachse deutlich kürzer als die kleine Halbachse ist, in Bereichen hoher Krümmung durch vergleichsweise mehr Punkte approximiert werden. Eine Konstante $c \geq 0$ sorgt dafür, dass die Energie nur minimiert wird, falls die Summe in Gleichung (7) oberhalb eines vorgegebenen Wertes liegt.

3.3.2 Priorenergie

Mithilfe der Priorenergie können bestimmte Konfigurationen auf Basis von Vorwissen favorisiert werden. Um die Ansammlung von vielen Objekten in Regionen mit niedriger Datenenergie zu vermeiden, werden Konfigurationen bestraft, in denen sich Ellipsen überlappen. Da sich Bombenkrater in der Regel nicht oder nur in geringem Maße überlappen, lassen wir in der Formulierung des Priorterms eine gewisse Überlappung zu, eine starke Überlappung wird jedoch vermieden. Zur Berechnung der Energie werden wie bei PERRIN et al. (2005) alle möglichen Kombinationen von sich überlappenden Ellipsenpaaren e_i, e_j herangezogen und jeweils die Überlappungsfläche A_{ij} der Ellipsen e_i und e_j sowie die zugehörigen relativen Überlappungsflächen $A_{ij}/A_i(e_i)$ und $A_{ij}/A_j(e_j)$ berechnet (Abb. 6). Für die Priorenergie folgt

$$U_P(X_t) = f \cdot \sum_{e_i, e_j} \max\left(\frac{A_{ij}}{A_i(e_i)}, \frac{A_{ij}}{A_j(e_j)}\right), \quad (8)$$

wobei sich in (8) die Stärke der Bestrafung mithilfe eines konstanten Faktors f gewichten lässt.

3.4 Einschränkung des Suchbereiches

Dem Verfahren der markierten Punktprozesse wird eine Methode zur Einschränkung des Suchbereiches im Bild vorangestellt. Aufgrund der Bildgrößen von ca. 5000 x 5000 Pixeln und mehr ist dieser Schritt sinnvoll, um den Rechenaufwand beim Sampling zu reduzieren. So nutzen beispielsweise ZHANG et al. (2014), ebenfalls im Zusammenhang mit markierten Punktprozessen, ein Baumkronendachmodell zur Suchraumeinschränkung. Innerhalb dieser Arbeit wird vorab ein Blob-Detektor angewandt (MALLICK 2015). Ein Blob ist eine Gruppe

zusammenhängender Pixel, die eine oder mehrere gemeinsame Eigenschaften (z.B. ähnliche Grauwerte) teilen. Ziel der Blob-Detektion ist es, diese Pixelregionen im Bild zu finden und zu markieren.

Im ersten Schritt der Blob-Detektion erfolgt eine Schwellwertbildung, bei welcher das Bild in mehrere Binärbilder konvertiert wird. Beginnend bei einem minimalen Schwellwert B_{S_min} wird dieser mittels eines Parameters für die Schrittweite B_{S_step} erhöht, sodass sich für den ersten Schwellwert B_{S_min} , für den zweiten $B_{S_min} + B_{S_step}$, für den dritten $B_{S_min} + 2 \cdot B_{S_step}$, usw. ergibt, bis ein maximaler Schwellwert B_{S_max} erreicht ist. Anschließend werden Zusammenhangskomponenten aus jedem Binärbild extrahiert und deren Zentren berechnet. Die Zentren werden in Abhängigkeit ihrer Koordinaten gruppiert, dabei bilden benachbarte Zentren eine Gruppe bzw. einen Blob. In diesem Zusammenhang stellt das Verfahren einen weiteren Parameter B_D bereit, welcher Blobs verschmelzen lässt, deren Distanz kleiner als B_D ist. Abschließend werden ausgehend von den Gruppen die finalen Zentren und Radien der Blobs bestimmt. Das Verfahren stellt zusätzlich zahlreiche Filtermöglichkeiten bereit. So lassen sich Blobs nach ihrer Rundheit B_R , Konvexität B_K und dem Trägheitsverhältnis B_T filtern. Für jede Größe stehen zwei Parameter, die untere und obere Grenze, definiert im Intervall $[0, 1]$, zur Verfügung. Hierbei stehen Werte nahe Eins für eine kreisähnliche Form. Ebenfalls ist es möglich, nur Blobs einer gewissen Pixelanzahl B_P zu erkennen, die sich zwischen einem minimalen und maximalen Wert bewegt. Ferner erlaubt das Verfahren, ausschließlich dunkle, helle oder beide Arten von Blobs zu detektieren.

Unser Ziel ist es, mithilfe des Blob-Detektors möglichst alle Bombenkrater zu erkennen, wenn auch um den Preis zahlreicher Fehldetektionen. Basierend auf diesen Ergebnissen können im Sampling-Prozess neue Objekte nur noch an denjenigen Positionen zufällig generiert werden, an denen bereits ein Blob detektiert wurde. Im Falle einer Geburt wird die Ellipse zunächst mit den Parametern des zugehörigen Blobs initialisiert. Somit werden potentielle Geburtsorte in Gleichung (4) nicht mehr durch die Breite und Höhe des Bildes beschränkt, sondern sind nur noch an Blob-Orten möglich. Die Ausgangskonfiguration besteht aus einer leeren Menge von Objekten.

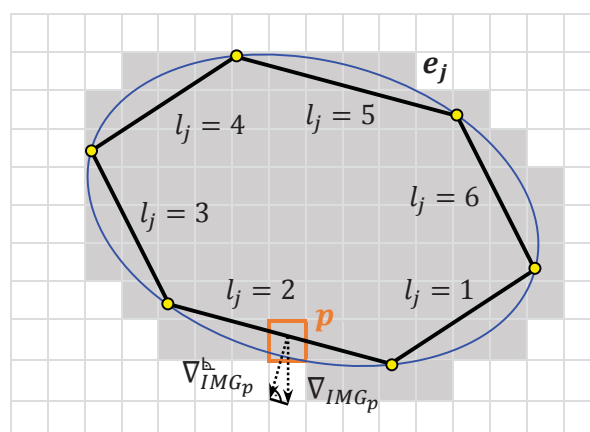


Abb. 5: Zur Berechnung des Datenterms berücksichtigen wir für jedes Pixel p des Ellipsenrandes (hier durch 6 Punkte approximiert, d.h. $l_j \in \{1, 2, \dots, 6\}$) die Summe der Komponenten der Gradienten $\nabla_{IMG_{p_j}}^b$ in Richtung des Normalenvektors. Graue Pixel kennzeichnen die aufgrund des Schattenwurfs zu erwartende dunkle Fläche innerhalb eines Bombenkraters im Bild.

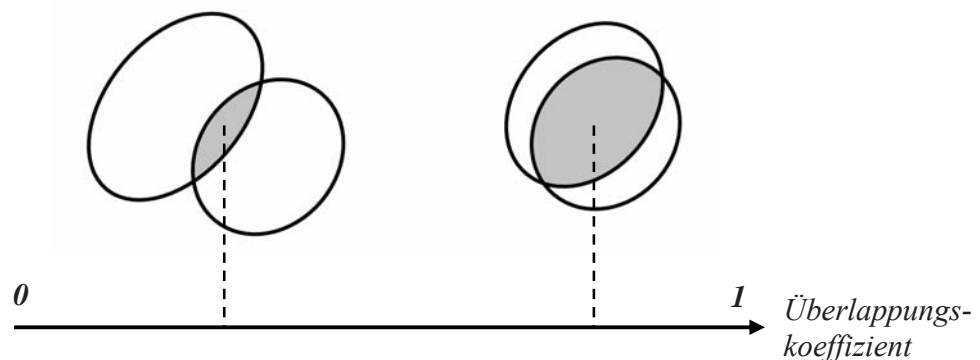


Abb. 6: Der Priorterm bestraft überlappende Objektkonfigurationen abhängig von ihrer relativen Überlappungsfläche. Je größer die gemeinsame Fläche, desto höhere Werte nimmt die Priorenergie an.

4 Experimente

4.1 Testdaten

Die Analysen unserer Methode erfolgen auf Grundlage von mit 1200 dpi gescannten panchromatischen Kriegsluftbildern mit 8-bit radiometrischer Auflösung aus dem zweiten Weltkrieg, welche vom Landesamt für Geoinformation und Landesvermessung Niedersachsen (LGLN) bereitgestellt wurden. Die Bilder verfügen über Bildmaßstabszahlen kleiner 15.000, sodass sich eine Bodenauflösung von ca. 0,3 m und besser ergibt. Zu Testzwecken werden drei Ausschnitte von Luftbildern aus Lemwerder, einer Gemeinde bei Bremen (Datensatz 1 und 2) sowie der Stadt Hannover (Datensatz 3) herangezogen. Innerhalb dieser Bildausschnitte herrschen unterschiedliche Beleuchtungssituationen und Bildinhalte vor. Der Schwerpunkt der Untersuchungen bzw. Entwicklungen liegt auf ländlichen Bereichen, weil in dicht bebauten Gebieten Bombenkrater durch den Schutt der zerstörten Gebäude verdeckt werden.

4.2 Parametereinstellungen

Die Parameterwahl wurde für unsere Experimente empirisch durchgeführt. Wird nicht gesondert darauf hingewiesen, sind die Parameter für die drei Bildausschnitte identisch gesetzt. Zunächst werden die Parameter des Blob-Detektors (Abschn. 3.4) auf $B_{S,min} = 10$, $B_{S,max} = 245$ und $B_{S,step} = B_D = 5$ gesetzt. Durch diese Wahl von $B_{S,min}$ werden Bildfehler, die sich in der Regel sehr dunkel im Bild ausprägen, nicht detektiert. Da Bombenkrater bzw. deren Schattenwürfe im Bild im Allgemeinen durch dunklere Grauwerte als die in ihrer Umgebung gekennzeichnet sind, soll das Verfahren nur dunkle Blobs erkennen. Vor dem Hintergrund, möglichst alle Bombenkrater im Bild zu detektieren, werden die Parameter zur Filterung mit $B_R \in \{0.1, 1\}$, $B_K \in \{0.4, 1\}$ und $B_T \in \{0.1, 1\}$ in der Form gesetzt, dass die Krater in ihrer Erscheinung durchaus von einem Kreis abweichen können. Auch wenn dies viele Fehldetektionen zur Folge hat, haben Experimente gezeigt, dass eine beschränkendere Parameterwahl die Detektion vieler Bombenkrater im Vorhinein ausschließen würde. Gleichmaßen ist es durch die Wahl von B_P im Intervall $[64, 4900]$ möglich, Bombenkrater einer sehr unterschiedlichen Pixelanzahl und

damit unterschiedlicher Größe zu detektieren. Dies ist notwendig, da die Größe der Bombenkrater in Abhängigkeit von der Bildmaßstabszahl stark variiert.

Im Zusammenhang mit den markierten Punktprozessen gewichten wir die Daten- und Priorenergie gleichermaßen, sodass der Parameter aus Gleichung (2) auf $\beta = 0.5$ gesetzt wird. Weiterhin wird beim Simulated Annealing ein geometrisches Schema zur Abkühlung genutzt, indem die Temperatur T_t aus Gleichung (3) mithilfe eines Faktors f_T in der Form $T_t = T_0 \cdot f_T^t$ verringert wird. Hierbei setzen wir die Starttemperatur $T_0 = 100$ und $f_T = 0.999994$. Die untere und obere Grenze der beiden Ellipsenhalbachsen (4) ergibt sich aus dem minimal und maximal im Bild vorhandenen Blob-Radius $B_r \in \{B_{r_{min}}, B_{r_{max}}\}$ zu $B_{r_{min}} = a_m = b_m$ bzw. $B_{r_{max}} = a_M = b_M$. Die Vorschlagswahrscheinlichkeiten der Kernel aus Abschnitt 3.2 werden auf $p_{Q_B} = p_{Q_D} = 0.05$ und $p_{Q_T} = p_{Q_M} = 0.45$ gesetzt. Die Wahrscheinlichkeit für die Auswahl eines Geburtsevents ($p_B = 0.05$) liegt deutlich unterhalb der für die Auswahl eines Todesevents mit $p_D = 0.95$ (Gleichung 5), da viele der bei der Blob-Detektion erzeugten potentiellen Krater in Wirklichkeit keinem Bombenkrater entsprechen. Auf diese Weise lassen sich die Berechnungen beschleunigen. Der Poisson-Parameter λ aus den Gleichungen (1) und (5) hängt von der erwarteten Anzahl an Objekten im Bild ab und wird für die Datensätze 1 und 2 auf $\lambda = 20$, für Datensatz 3 auf $\lambda = 40$ gesetzt. Ferner setzen wir den Parameter c für die ersten beiden Datensätze auf $c = 100$, für Datensatz 3 auf $c = 200$ (Gleichung 7), da sich die Grauwertgradienten unterschiedlich stark in den Bildern ausprägen. In Anlehnung an die Formulierungen aus Abschnitt 3.3.2 wählen wir den Faktor $f = 1000$ (Gleichung 8), da hierdurch Überlappungen von Objekten in geringem Maße möglich bleiben.

4.3 Ergebnisse

Die optimale Objektkonfiguration wird durch das Hinzufügen bzw. Entfernen von Ellipsen oder deren Parametervariation iterativ erzeugt (Abb. 7c-e). Ausgehend von den Ergebnissen des Blob-Detektors (Abb. 7b, rote Kreise) sind im Sampling-Prozess nur noch Geburten an Positionen möglich, an denen bereits ein Blob detektiert wurde. Untersuchungen haben gezeigt, dass die Möglichkeit von Geburten außerhalb dieser Blob-Orte bei ähnlichen Ergebnissen zu höheren Laufzeiten führt. Neben einigen falschen Detektionen (Abb. 7e, rote Ellipsen) werden bis auf zwei Bombenkrater (Abb. 7e, blaue Pfeile) alle gefunden.

Abbildung 8 zeigt die Ergebnisse für die Datensätze 2 und 3. Auffällig für den dritten Datensatz ist, dass trotz der zahlreichen falschen Detektionen des Blob-Detektors (Abb. 8e) viele dieser durch das Verfahren der markierten Punktprozesse eliminiert werden (Abb. 8f). Um die Ergebnisse zu evaluieren, wurden manuell flächenhafte Referenzdaten für die Bombenkrater erzeugt. Die Evaluierung erfolgt dabei gesondert für die Ergebnisse des Blob-Detektors und die der markierten Punktprozesse. Liegt ein Blob- bzw. Ellipsenmittelpunkt innerhalb der Referenzfläche eines Bombenkraters, so wird der Blob bzw. die Ellipse als korrekt definiert. Die Korrektheit ist somit der prozentuale Anteil an Blob- bzw. Ellipsenmittelpunkten, die innerhalb der Referenzfläche liegen. Die Vollständigkeit der Objektkonfiguration ist der prozentuale Anteil an Bombenkratern, der durch die jeweilige Methode gefunden wurde. Die Ergebnisse in Tabelle 1 zeigen, dass mithilfe des Blob-Detektors eine mittlere Vollständigkeit von 93 % erreicht werden kann, jedoch liegt die Korrektheit mit ca. 34 % deutlich unterhalb des Wertes für die Vollständigkeit. Das hier vorgestellte Verfahren auf Basis von markierten Punktprozessen

erhöht die Korrektheit auf ca. 62 % unter geringer Einbuße der Vollständigkeit um in etwa fünf Prozentpunkte, wobei für Datensatz 3 die Steigerung der Korrektheit um 40 % am höchsten ist. Durch die Elimination vieler falscher Detektionen können im Hinblick auf die Erstellung einer Belastungskarte Bereiche entstehen, die als nicht durch Kriegseinwirkung gefährdet einzustufen sind (Abb. 8c, f, grüne Pfeile).

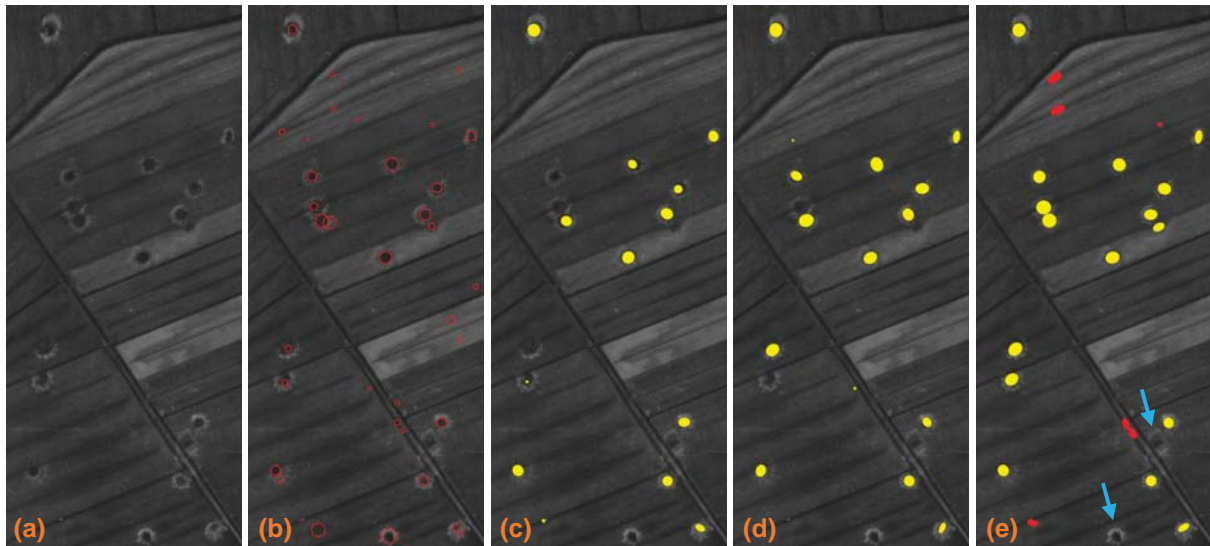


Abb. 7: (a) Datensatz 1 mit Bombenkratern, (b) Ergebnisse der Blob-Detektion (rote Kreise) und (c-e) die auf den Blob-Ergebnissen aufbauenden Resultate des Verfahrens der markierten Punktprozesse (Ellipsen) nach $1 \cdot 10^5$, $3 \cdot 10^5$ und $10 \cdot 10^5$ Iterationen, wobei korrekte Detektionen in Gelb und falsche Detektionen in Rot dargestellt sind.

Tab. 1: Zusammenstellung der Evaluierungsergebnisse

	Blob-Detektor		Markierte Punktprozesse	
	Vollständigkeit	Korrektheit	Vollständigkeit	Korrektheit
Datensatz 1	94,4%	47,2%	88,8%	72,7%
Datensatz 2	89,5%	38,6%	84,2%	55,2%
Datensatz 3	95,1%	16,9%	90,2%	56,9%
Mittelwert	93,0%	34,2%	87,7%	61,6%

Abschließend möchten wir kurz eine Schwäche unseres aktuellen Verfahrens aufzeigen. Objekte, die sich im Bild ähnlich wie Bombenkrater ausprägen, führen zu vielen Fehldetektionen. Hierfür sind insbesondere Schattenwürfe von Häusern (Abb. 9) oder auch Bäumen (Abb. 8d-f, rechts) verantwortlich. Im Ausblick (Abschn. 5) werden Ideen diskutiert, wie diesen Einschränkungen entgegengewirkt werden kann.

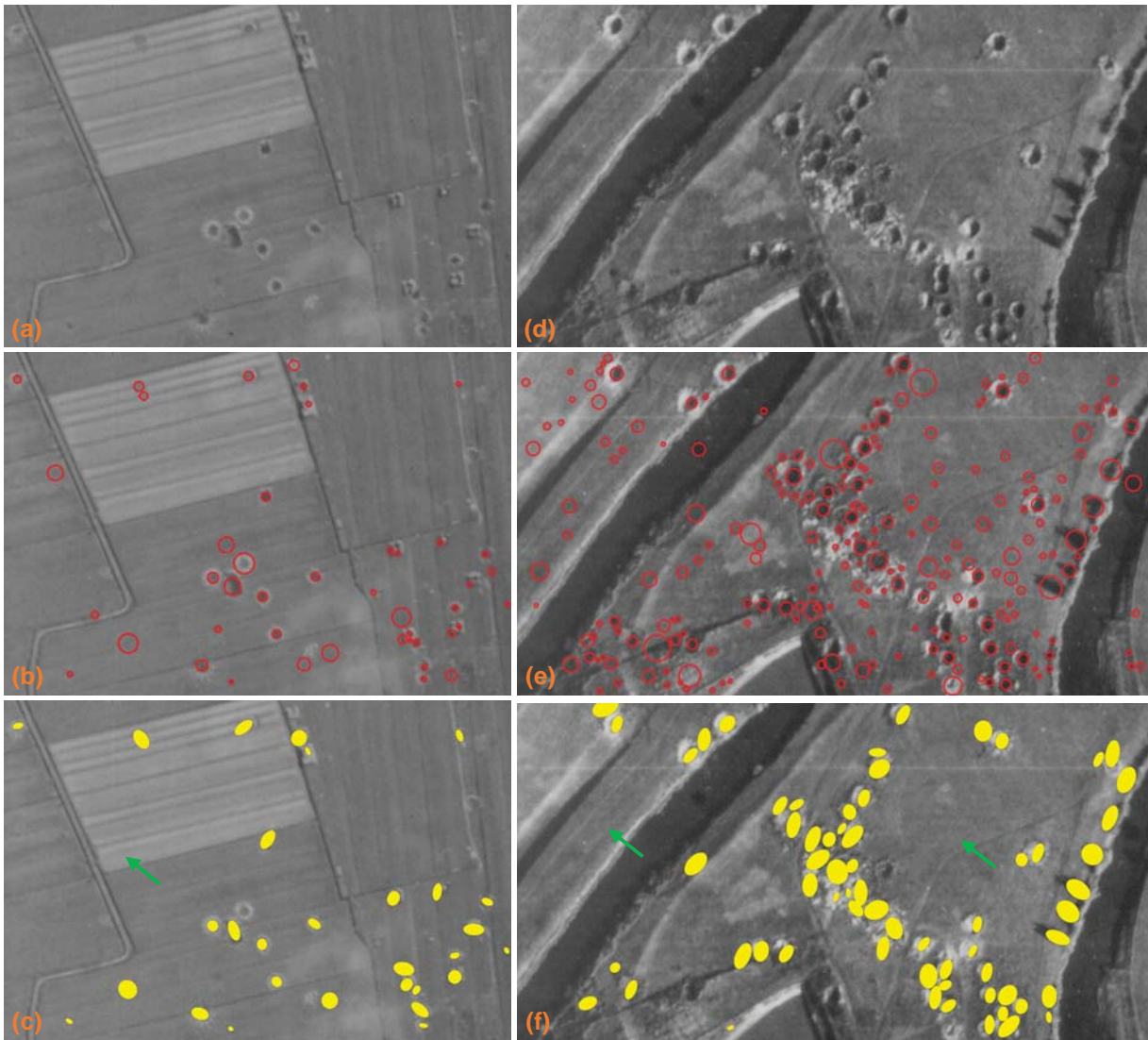


Abb. 8: (a, d) Datensätze 2 und 3 mit Bombenkratern, (b, e) Ergebnisse der Blob-Detektion (rote Kreise) und (c, f) die auf den Blob-Ergebnissen aufbauenden Resultate des Verfahrens der markierten Punktprozesse, wobei Ellipsen in Gelb dargestellt sind.

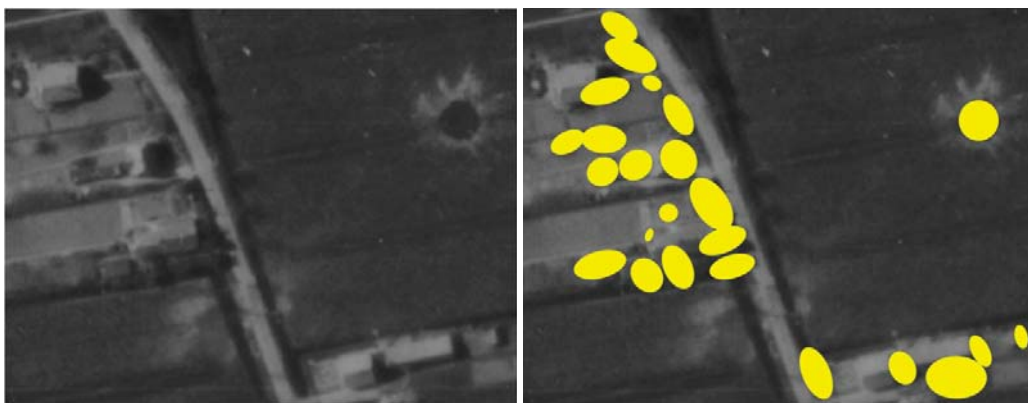


Abb. 9: Schattenbereiche (hier von Häusern verursacht) führen zu zahlreichen Fehldetektionen

5 Schlussfolgerungen und Ausblick

Wie die Ergebnisse zeigen, kann ein großer Teil der im Bild vorhandenen Bombenkrater durch den vorgestellten Ansatz gefunden werden, jedoch treten auch – abhängig vom Bildinhalt – Fehldetektion auf. Diese Ergebnisse zeigen das Potential, aber auch die derzeitigen Grenzen der Methode auf. Probleme entstehen vor allem durch Objekte, die sich im Bild ähnlich wie Bombenkrater ausprägen (z.B. Schatten von Häusern oder Bäumen). Um dem entgegenzuwirken, soll in Zukunft der Datenterm des Modells dahingehend erweitert werden, dass es auch einen Kranz an helleren Pixeln um den Ellipsenrand herum fordert (z.B. DESCAMPS et al. 2011). Dieses Phänomen ist in den Bildern aufgrund des bei der Detonation verursachten Auswurfs zu beobachten. In Gebieten mit stereoskopischer Abdeckung durch Bilder kann auch Höheninformation integriert werden, wobei die Qualität der Luftbilder einen limitierenden Faktor darstellen könnte. Ferner soll der Sonnenstand ebenso berücksichtigt werden wie der Umstand, dass Bombenkrater selten einzeln vorkommen. Daher sollte die Erweiterung des Priorterms um ein Modell für Kratergruppen die Ergebnisse weiter verbessern.

6 Danksagung

Die Autoren danken dem Landesamt für Geoinformation und Landesvermessung Niedersachsen (LGLN) und dem als Dezernat in die Regionaldirektion Hameln-Hannover des LGLN integrierten Kampfmittelbeseitigungsdienstes (KBD) herzlich für die Bereitstellung der Daten und die finanzielle Unterstützung des Projekts.

7 Literaturverzeichnis

- ANDRIEU, C., DE FREITAS, N., DOUCET, A. & JORDAN, M. I., 2003: An Introduction to MCMC for Machine Learning. *Machine Learning* **50**, 5-43.
- BORCS, A. & BENEDEK, C., 2015: Extraction of Vehicle Groups in Airborne Lidar Point Clouds With Two-Level Point Processes. *IEEE Transactions on Geoscience and Remote Sensing (GRS)* **53**(3), 1475-1489.
- BRÉDIF, M., TOURNAIRE, O., VALLET, B. & CHAMPION, N., 2013: Extracting polygonal building footprints from digital surface models: A fully-automatic global optimization framework. *ISPRS Journal of Photogrammetry and Remote Sensing* **77**, 57-65.
- CHAI, D., FÖRSTNER, W. & YING YANG, M., 2012: Combine Markov random fields and marked point processes to extract building from remotely sensed images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **I-3**, 365-370.
- DALEY, D. J. & VERE-JONES, D., 2003: An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods (2. Aufl.). *Probability and its Applications*. New York: Springer.
- DESCAMPS, S., DESCOMBES, X., BECHET, A. & ZERUBIA, J., 2008: Automatic Flamingo detection using a multiple birth and death process. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1113-1116.

- DESCAMPS, S., BÉCHET, A., DESCOMBES, X., ARNAUD, A. & ZERUBIA, J., 2011: An automatic counter for aerial images of aggregations of large birds. *Bird Study* **58**(3), 302-308.
- DESCOMBES, X. & ZERUBIA, J., 2002: Marked Point Process in Image Analysis. *IEEE Signal Processing Magazine* **19**(5), 77-84.
- DESCOMBES, X., 2013: Marked Point Processes for Object Detection. *Stochastic Geometry for Image Analysis*, Descombes, X. (Hrsg.), John Wiley & Sons, Inc., Hoboken: Wiley, 11-27.
- DUBOSCLARD, P., LARNIER, S., KONIK, H., HERBULOT, A. & DEVY, M., 2014: Automatic Method for Visual Grading of Seed Food Products. *International Conference on Image Analysis and Recognition (ICIAR)* **1**, 485-495.
- GEMAN, S. & GEMAN, D., 1984: Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* **6**(6), 721-741.
- GILKS, W. R., RICHARDSON, S. & SPIEGELHALTER, D., 1996: *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC Interdisciplinary Statistics: Taylor & Francis.
- GREEN, P. J., 1995: Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**(4), 711-732.
- HASTINGS, W. K., 1970: Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**(1), 97-109.
- KIRKPATRICK, S., GELATT, C. D. & VECCHI, M. P., 1983: Optimization by Simulated Annealing. *Science* **220**(4598), 671-680.
- KUMAR, S. & HEBERT, M., 2006: Discriminative Random Fields. *International Journal of Computer Vision* **68**(2), 179-201.
- LACOSTE, C., DESCOMBES, X. & ZERUBIA, J., 2005: Point processes for unsupervised line network extraction in remote sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* **27** (10), 1568-1579.
- LAFARGE, F., GIMEL'FARB, G. & DESCOMBES, X., 2010: Geometric feature extraction by a multimarked point process. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* **32**(9), 1597-1609.
- LI, S. Z., 2009: *Markov Random Field Modeling in Image Analysis* (3. Aufl.). London: Springer.
- MALLET, C., LAFARGE, F., ROUX, M., SOERGEL, U., BRETAR, F. & HEIPKE, C., 2010: A Marked Point Process for Modeling Lidar Waveforms. *IEEE Transactions on Image Processing (IP)* **19**(12), 3204-3221.
- MALLICK, S., 2015: Blob Detection Using OpenCV (Python, C++). <https://www.learnopencv.com/blob-detection-using-opencv-python-c/>, letzter Zugriff Dezember 2016.
- METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H. & TELLER, E., 1953: Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* **21**(6), 1087-1092.
- PERRIN, G., DESCOMBES, X. & ZERUBIA, J., 2005: Adaptive Simulated Annealing for Energy Minimization Problem in a Marked Point Process Application. *Energy Minimization Methods in Computer Vision and Pattern Recognition*, RANGARAJAN, A., VEMURI, B. & YUILLE, A. (Hrsg.), Berlin Heidelberg: Springer, 3-17.

- PERRIN, G., DESCOMBES, X. & ZERUBIA, J., 2006: 2D and 3D Vegetation Resource Parameters Assessment using Marked Point Processes. International Conference on Pattern Recognition (ICPR), 1-4.
- SALAMON, P., SIBANI, P. & FROST, R., 2002: Facts, Conjectures, and Improvements for Simulated Annealing. Society for Industrial and Applied Mathematics. Philadelphia, USA.
- SCHMIDT, A., KRUSE, C., ROTTENSTEINER, F., SOERGEL, U. & HEIPKE, C., 2016: Network detection in raster data using marked point processes. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XLI-B3**, 701-708.
- SUN, K., SANG, N., & ZHANG, T., 2007: Marked Point Process for Vascular Tree Extraction on Angiogram. Energy Minimization Methods in Computer Vision and Pattern Recognition, YUILLE, A., ZHU, S.-C., CREMERS, D. & WANG, Y. (Hrsg.), Berlin Heidelberg: Springer, 467-478.
- TOURNAIRE, O., BRÉDIF, M., BOLDO, D. & DURUPT, M., 2010: An efficient stochastic approach for building footprint extraction from digital elevation models. ISPRS Journal of Photogrammetry and Remote Sensing **65**(4), 317-327.
- TROGLIO, G., BENEDIKTSSON, J., MOSER, G. & SERPICO, S., 2010: Crater detection based on marked point processes. Proceedings of the International Geoscience and Remote Sensing Symposium, Honolulu, 1378-1381.
- UTASI, A. & BENEDEK, C., 2011: A 3-D marked point process model for multi-view people detection. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3385-3392.
- VAN LAARHOVEN, P. J. & AARTS, E. H., 1987: Simulated Annealing: Theory and Applications. Ellis Horwood Series in Mathematics and Its Applications. Niederlande: Springer.
- WENZEL, S. & FÖRSTNER, W., 2016: Facade interpretation using a marked point process. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **III-3**, 363-370.
- ZHANG, J., SOHN, G. & BRÉDIF, M., 2014: A hybrid framework for single tree detection from airborne laser scanning data: A case study in temperate mature coniferous forests in Ontario, Canada. ISPRS Journal of Photogrammetry and Remote Sensing **98**, 44-57.

Konzept zur Gefährdungserkennung im städtischen Verkehrsraum durch Personendetektion in MLS-Punktwolken

BJÖRN BORGMANN^{1,2}, MARCUS HEBEL¹, MICHAEL ARENS¹ & UWE STILLA²

Zusammenfassung: Die Detektion von personenbezogenen Gefährdungen im städtischen Verkehrsraum stellt eine Aufgabe dar, deren automatische Erfüllung durch ein technisches System z.B. im Hinblick auf die Entwicklung von Fahrerassistenzsystemen wünschenswert ist. In diesem Beitrag wird zunächst das Konzept einer Methode vorgestellt, welche diese Aufgabe unter Verwendung von 3D-LiDAR-Sensoren lösen kann. In der Methode wird die Aufgabe in mehrere Teilschritte zerlegt, zu denen eine Vorverarbeitung, die Detektion und das Tracking von Personen sowie eine abschließende Situationsbewertung gehören. Die zweite Hälfte des Beitrags konzentriert sich auf den Teilschritt der Detektion von Personen und auf die dafür notwendigen Schritte der Datenvorverarbeitung. Die Datenvorverarbeitung generiert dabei eine Reihe von Punktwolkensegmenten, bei denen es sich um Personen handeln könnte. Die Personendetektion erfolgt dann durch Klassifizierung dieser Segmente auf Basis eines „Implicit Shape Models“. In ersten Untersuchungen wurde eine Genauigkeit von ca. 66 % bei einem Recall von ca. 76 % erreicht.

1 Einleitung

Besonders im städtischen Verkehrsraum bestehen Gefährdungen für Personen, die z.B. durch Fahrzeuge verursacht werden. Solche Gefährdungen sind für einen Fahrer nur schwer umfassend zu erkennen, da die Vielzahl an Strukturen in städtischen Umgebungen wie z.B. Fahrzeuge, Stadtmöbel, Verkehrsschilder, Ampeln und andere Personen es ihm erschweren, alle Personen in seinem Umfeld rechtzeitig wahrzunehmen und zu erkennen. Oft fehlt ihm ein vollständiges Bild der Lage, weswegen seine Fähigkeit, auf eine mögliche Gefahrensituation zu reagieren, eingeschränkt ist. An dieser Stelle sind technische Systeme wünschenswert, die z.B. in Form eines Assistenzsystems dabei unterstützen, Gefährdungssituationen im Zusammenhang mit Personen im eigenen Umfeld rechtzeitig zu erkennen. Neben dem Einsatz derartiger Systeme in Fahrzeugen ist auch die stationäre Verwendung z.B. bei der Nutzung schwerer Maschinen oder für besondere Ereignisse bzw. Überwachungsaufgaben hilfreich.

Systeme zur automatischen Gefährdungserkennung können mit unterschiedlichem Funktionsumfang ausgestattet sein. Es ist möglich, dass sie lediglich bei der Wahrnehmung des Umfelds unterstützen und z.B. Personen, die ein Mensch evtl. noch nicht selbst wahrgenommen hat, hervorheben bzw. auf diese aufmerksam machen. Wünschenswert ist, dass sie darüber hinaus auch Rückschlüsse aus dem erkannten Verhalten von Personen ziehen und dadurch mögliche Gefahrensituationen selbständig detektieren.

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

Bei der Realisierung eines solchen technischen Systems stellt die jederzeitige Erfassung des Umfeldes durch das technische System eine notwendige Voraussetzung dar. Hierfür eignen sich unterschiedliche Arten von Sensoren. Neben Kameras im sichtbaren oder infraroten Spektralbereich gehören dazu auch verschiedenartige Sensoren zur direkten 3D-Erfassung einer Szene. Zu diesen zählen LiDAR-Systeme. Diese sind unter Kenntnis ihrer Position und Ausrichtung in der Lage 3D-Koordinaten von Oberflächen zu messen, welche durch Punktwolken repräsentiert werden können. Solche Systeme sind in der Lage, die Geometrie der beobachteten Oberflächen direkt und unabhängig von den Lichtverhältnissen zu erfassen.

Der vorliegende Beitrag stellt das Konzept einer Methode vor, durch die basierend auf 3D-Punktwolken wie sie von LiDAR-Systemen aufgenommen werden, oben genannte Gefährdungssituationen automatisch erkannt werden sollen. Hierfür sind verschiedene Teilaufgaben zu erfüllen. In diesem Beitrag wird neben dem Konzept der gesamten Methode auf eine dieser zu erfüllenden Teilaufgaben, nämlich die Detektion und Erkennung von Personen, detailliert eingegangen.

2 Aufgabenstellung

Aus der bereits beschriebenen Zielsetzung, personenbezogene Gefährdungen im städtischen Verkehrsraum zu detektieren, lassen sich verschiedene Aufgaben ableiten, die zu lösen sind. Bei dem Design sowie der Umsetzung der gesamten Methode muss beachtet werden, dass die dafür eingesetzten Verfahren schritthaltend mit der Datenaufnahme arbeiten können, da die Zielsetzung vorsieht ein System zu schaffen, welches die Gefährdungen im Live-Betrieb detektiert. Hieraus ergibt sich ein Kriterium für die maximale Laufzeit bzw. den Aufwand der verwendeten Verfahren.

Zunächst ist es erforderlich, Personen in den aufgenommenen Sensordaten zu detektieren. Hierbei gibt es die Schwierigkeit, dass aufgrund der möglichst vollständigen (bis zu 360°) und zeitlich schnell aufeinanderfolgend wiederholten Umgebungserfassung die jeweils aktuell verfügbare Datendichte (d.h. die Anzahl der aufgenommenen 3D-Punkte bezogen auf den aufgenommenen Raum) vor allem in größeren Entfernungen gering ist. Personen werden daher oft nur grob erfasst. Das verwendete Detektionsverfahren sollte daher auch bei geringen Datendichten robust funktionieren. Abb. 1 verdeutlicht diese Problematik am Beispiel von Personen, die mit einem Velodyne HDL-64E LiDAR-Sensor in einer Entfernung von ca. 6 m bzw. 12 m aufgenommen wurden. Derartige LiDAR-Sensoren werden zurzeit z.B. im Bereich des autonomen Fahrens verwendet.

Die reine Position einer Person liefert allerdings nur sehr begrenzte Informationen darüber, was die Person gerade tut, wohin sie sich bewegt oder was sie vorhat. Diese Informationen sind jedoch für eine Gefährdungserkennung nützlich und sollten soweit möglich ebenfalls ermittelt werden. Ein weiteres zu lösendes Teilproblem besteht daher darin, bereits detektierte Personen im Zeitverlauf zu verfolgen um Informationen über ihr Bewegungsverhalten zu erhalten. Ebenfalls interessant sind die Körperposen detektierter Personen, da diese Rückschlüsse darüber erlauben was die Personen gerade tun. So haben zwei Personen, die z.B. am Straßenrand stehen und ein Gespräch führen, meist eine andere Pose als eine Person, die aktuell im Begriff ist auf die Fahrbahn zu treten. Es ist daher wünschenswert, Merkmale über die Körperpose wie z.B. die Position der Extremitäten oder auch die Blickrichtung zu ermitteln.

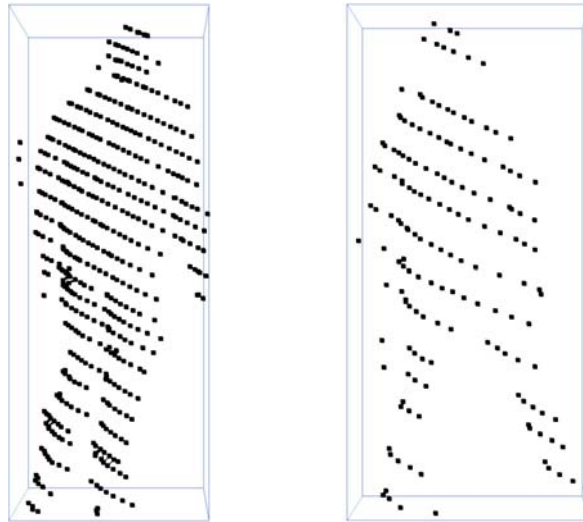


Abb. 1: Beispiele für Personen aufgenommen mit einem Velodyne HDL-64E LiDAR in ca. 6 m (links) und ca. 12 m (rechts) Abstand.

Der letzte Schritt einer automatischen Gefährdungserkennung besteht darin, die gesammelten Informationen auf semantischer Ebene zu einer Situationsbewertung zusammenzuführen und so den eigentlichen Output des Systems zu generieren.

3 Verwandte Arbeiten

In diesem Abschnitt werden verschiedene relevante Arbeiten vorgestellt, die sich bereits mit den im vorherigen Abschnitt angeschnittenen Themen beschäftigen und somit den gegenwärtigen Stand der Technik beschreiben. Insbesondere im Bereich des autonomen Fahrens sowie der Fahrerassistenzsysteme entstanden in den letzten Jahren viele Arbeiten, die sich mit ähnlichen Fragestellungen beschäftigen. So stellen beispielsweise KELLER et al. (2011) ein Fahrerassistenzsystem vor, welches der Fußgängersicherheit dient, indem es bevorstehende Kollisionen mit diesen detektiert und dann selbständig Brems- bzw. Ausweichmanöver durchführt. In ihrer Arbeit verwenden sie ein Stereokamerasystem und fusionieren eine Fußgängerdetektion und eine Detektion bewegter Objekte miteinander. Sie nutzen die so gesammelten Daten um die Kennzahlen TTB (engl. *Time to Break*) und TTS (engl. *Time to Steer*) zu ermitteln, welche die Basis für ihre Situationsanalyse darstellen. Eine umfassende aber bereits etwas ältere Übersicht über verschiedene Aspekte und Ansätze im Hinblick auf Fahrerassistenzsysteme zur Erhöhung der Fußgängersicherheit wurde von GANDHI & TRIVEDI (2007) vorgestellt.

Zur Detektion von Personen in 3D-Daten gibt es unterschiedliche Verfahren. Dazu gehören verschiedene Ansätze, die auf *Support Vector Machines* (SVMs) als Klassifikator zurückgreifen. Bei diesen wird in einer Trainingsphase im Merkmalsraum eine Hyperebene ermittelt, welche die unterschiedlichen Klassen voneinander trennt. NAVARRO-SERMENT et al. (2010) detektieren Personen in LiDAR-Punktwolken und verwenden dafür zwei hintereinander angewendete SVMs. Die erste erhält als Eingabe eine Reihe geometrischer Merkmale, die durch Projektion der Punkte eines

Detektionskandidaten auf zwei verschiedene zweidimensionale Ebenen ermittelt werden. Die beiden Ebenen werden dabei aus der Kombination von jeweils zwei Eigenvektoren der Punktwolke des Detektionskandidaten gebildet. Die zweite SVM erhält als Eingabe das Ergebnis der ersten sowie eine Reihe von Bewegungsmerkmalen, die in einem zuvor erfolgten Segmentations- und Tracking-Schritt ermittelt wurden.

Random Decision Forests sind eine Klassifikationsmethode bei der in einem Trainingsschritt mehrere Entscheidungsbäume mit einem bestimmten Zufallselement unabhängig voneinander trainiert werden. Sie sind so in der Lage dem Problem der Überanpassung (engl. Overfitting) zu begegnen. Sie sind geeignet in einzelnen Tiefenbildern sowohl Personen zu detektieren als auch dazu bestimmte Körperteile zu erkennen (SHOTTON et al. 2011 und SHOTTON et al. 2013). Sie klassifizieren hierfür jedes Pixel des Tiefenbildes dahingehend, zu welchem Körperteil es gehört und sind dabei ausreichend schnell für eine schritthaltende Verarbeitung, erfordern jedoch ein umfangreiches Training und ausreichend hoch aufgelöste Tiefenbilder.

Eine weitere Gruppe von Klassifikationsverfahren stellen *Bag of Words* Methoden dar. Bei diesen werden zunächst Merkmale ermittelt und diese einem „Wort“ (visuelles. bzw. geometrisches Wort) in einem „Wörterbuch“ zugeordnet. Die zugeordneten Wörter wiederum entscheiden in einem Abstimmprozess darüber, welcher Klasse ein Objekt zuzuordnen ist. Hierfür erfolgt zunächst ein Training, bei dem die Wörter sowie die Stimmen der Wörter für Objektklassen trainiert werden. *Implicit Shape Models* (ISM) erweitern klassische *Bag of Words* Methoden dahingehend, dass sie auch berücksichtigen wo an einem Objekt ein bestimmtes Merkmal zu finden ist, wodurch die Korrektheit der Klassifikation verbessert wird. Sie tun dies indem für die Wörter bzw. die den Wörtern zugeordneten Stimmen z.B. hinterlegt ist, wo sich von der Position des Wortes aus gesehen der Objektmittelpunkt befindet. Im Abstimmprozess wird dann ermittelt, an welcher Position die Stimmen für einen Objektmittelpunkt einer bestimmten Klasse konvergieren. Ursprünglich wurden sie für die Verwendung im Bereich der Objekterkennung in Bildern entwickelt (LEIBE et al. 2008, JÜNGLING & ARENS 2011) werden jedoch mittlerweile auch auf 3D-Punktwolken angewendet (KNOPP et al. 2010 und VELIZHEV et al. 2012). Ein an *Implicit Shape Models* angelehntes Verfahren wurde von SPINELLO et al. (2010) vorgestellt. Sie betrachten die Daten zunächst in mehreren flachen horizontal übereinander angeordneten Ebenen und bestimmen auf diesen Segmente und verschiedene Merkmale. Hierbei verwenden sie Methoden, die aus der Verarbeitung von Daten aus 2D-LiDAR-Sensoren (LiDAR-Sensoren mit nur einer Messebene) bekannt sind. Anschließend stimmen die extrahierten Merkmale für Objektpositionen im 3D-Raum ab.

4 Konzeptionelles Design der Methode

In diesem Abschnitt wird ein Gesamtüberblick über unsere Methode zum Erkennen von Gefährdungen im städtischen Verkehrsraum gegeben und die einzelnen Komponenten werden erläutert. Die Methode ist in Abb. 2 schematisch dargestellt. Sie erhält als Eingabe dreidimensionale Punktwolken, die bereits untereinander registriert sind, sodass alle ein gemeinsames Koordinatensystem verwenden. Eine solche Registrierung ist z.B. bei einem stationären System gegeben oder kann durch die Verwendung eines SLAM-Verfahrens (engl. *Simultaneous Localization and Mapping*) bzw. messtechnisch durch den Einsatz eines inertialen Navigationssystems (INS) gekoppelt mit

GNSS-Empfängern erzeugt werden (direkte Georeferenzierung). In diesem Beitrag wird davon ausgegangen, dass eine Registrierung dieser Art bereits vorhanden ist.

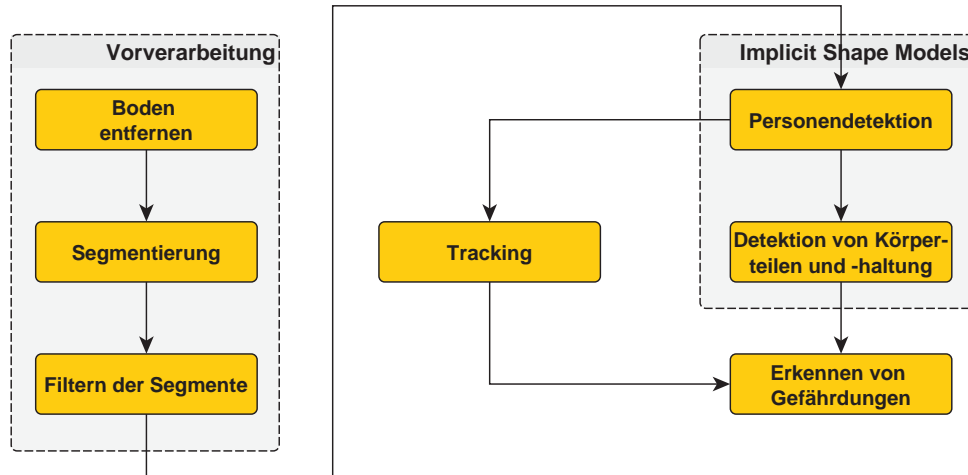


Abb. 2: Überblick über unsere Methodik.

Unsere bisherigen Arbeiten konzentrierten sich auf den Bereich der Vorverarbeitung und der Detektion von Personen sowie in einer späteren Ausbaustufe auf die Detektion von Körperteilen und der Körperhaltung der detektierten Personen. Auf diese beiden Bereiche wird in den beiden folgenden Abschnitten noch näher eingegangen, sie werden daher in diesem Unterabschnitt nur kurz zusammengefasst.

Für die Punktwolken erfolgt zunächst eine Vorverarbeitung, deren Ziel es ist, den Datenumfang mit Methoden kurzer Verarbeitungszeit möglichst stark zu reduzieren und in eine Form zu bringen, die eine parallele Verarbeitung in den nachfolgenden Schritten begünstigt. Für die Detektion von Personen, Körperteilen und der Körperhaltung verwenden wir einen Ansatz, der auf dem Konzept der *Implicit Shape Models* (ISMs) beruht. Diesen möchten wir für die Erkennung von Körperteilen in 3D-Punktwolken erweitern.

Das Tracking dient primär dazu, Informationen über das Verhalten der erfassten Personen im Zeitverlauf zu sammeln. Es wird in späteren Ausbaustufen jedoch auch möglich sein, es zusätzlich zu verwenden, um bereits detektierte Personen im Zeitverlauf nicht in jeder neu erfassten Punktwolke erneut detektieren zu müssen. Denkbar ist z.B., Segmente innerhalb eines Tracks solange nur mit niedriger Priorität bzw. nur in jeder x -ten Punktwolke erneut mit dem Personendetektor zu verarbeiten, wie der Konfidenzwert des Trackings über einem bestimmten Schwellwert liegt. Hierdurch kann eine Laufzeitersparnis erzielt werden, die eine schritthaltende Verarbeitung begünstigt. Für das Tracking ist derzeit die Verwendung eines Kalman-Filter geplant.

Die eigentliche Gefährdungserkennung erhält als Eingabedaten eine Vielzahl sehr unterschiedlicher Informationen. Neben den ermittelten Informationen über die erfassten Personen können noch weitere Daten z.B. über die Eigenbewegung des Sensorträgers in diese einfließen. Hierfür ist daher ein Verfahren notwendig, welches in der Lage ist, basierend auf sehr heterogenen Eingabedaten

zu arbeiten. Denkbar ist hier z.B. die Verwendung eines Verfahrens, welches auf Entscheidungsbäumen basiert. Diese Untersuchungen stehen noch am Anfang und werden Gegenstand unserer zukünftigen Arbeiten sein.

5 Vorverarbeitung

Die einzelnen Schritte der Vorverarbeitung sind in Abb. 2 dargestellt und in Abb. 3 sind ihre jeweiligen Ausgaben zu sehen. Die Vorverarbeitungskette besteht aus einem Entfernen der Bodenfläche, einer Segmentierung mithilfe eines Regionenwachstumsverfahrens (engl. *Region Growing*) und dem Filtern der daraus resultierenden Segmente anhand einiger einfacher Merkmale und Kriterien. Die einzelnen Schritte werden im Folgenden näher erläutert.

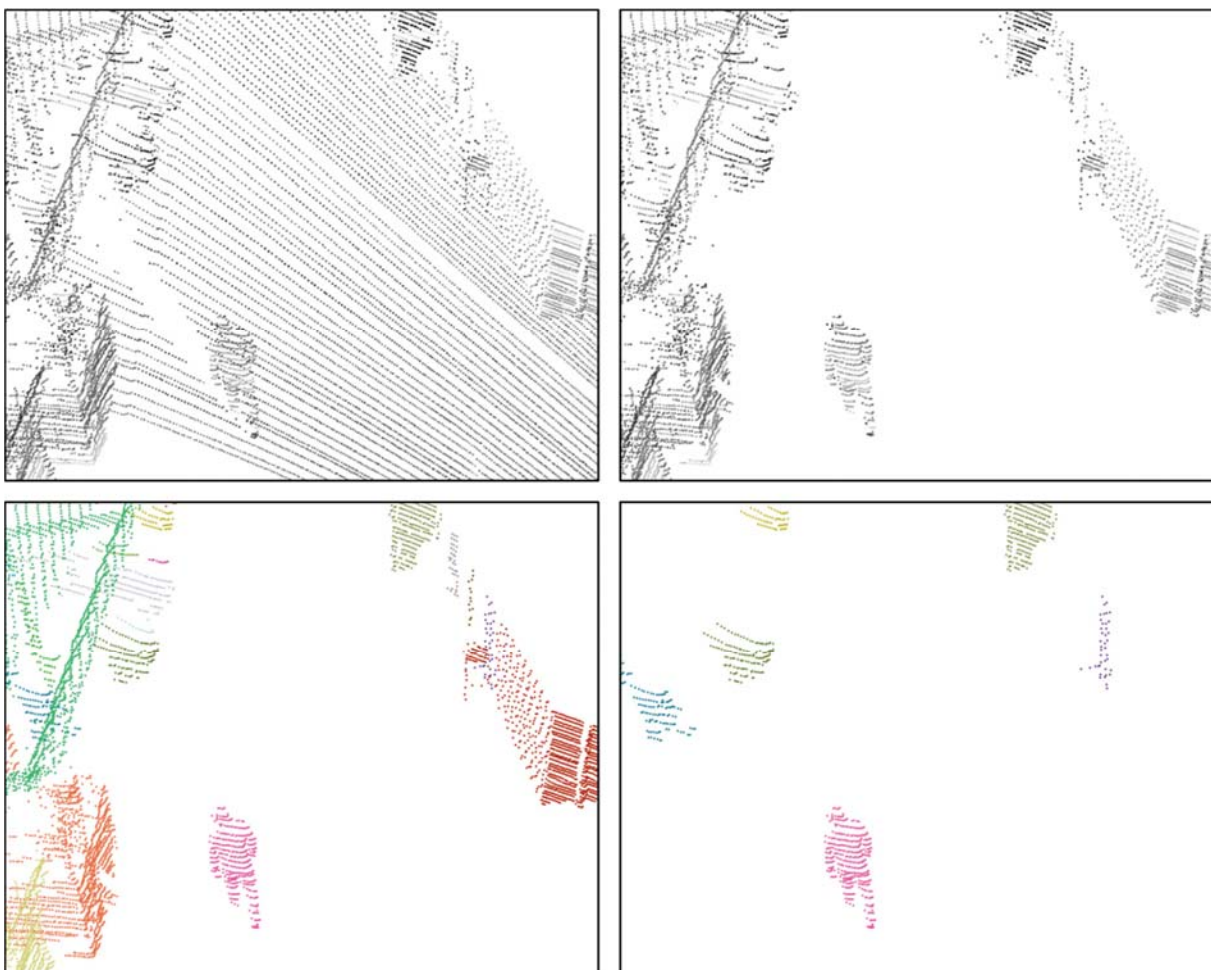


Abb. 3: Verarbeitungsschritte der Vorverarbeitung. Oben links: Eingabedaten. Oben rechts: Bodenebene entfernt. Unten links: Ergebnis der Segmentierung (Segmente sind farbcodiert). Unten rechts: Ergebnis der Filterung (Segmente sind farbcodiert).

5.1 Ermitteln und Entfernen der Bodenebene

Durch das Entfernen der Bodenfläche erfolgt bereits eine deutliche Reduzierung des Datenumfanges. Außerdem erleichtert sie die nachfolgende Segmentierung, da ansonsten während eines Regionenwachstums viele zu unterscheidende Segmente über die Bodenebene miteinander zu einem einzigen Segment verbunden wären. Es müssten dann beim Segmentieren weitere Kriterien wie z.B. die Richtung der Normalen mitberücksichtigt werden, die hierfür jedoch zuvor erst rechenaufwendig bestimmt werden müssten. Für die Entfernung der Bodenebene erzeugen wir zunächst ein niedrig aufgelöstes Bodenraster und bestimmen dann für jeden Punkt seinen Abstand zu der aus diesem Raster gebildeten Bodenebene. Punkte deren Abstand einen bestimmten Grenzwert unterschreiten werden als Teil des Bodens angesehen und entfernt.

Zur Bestimmung des Bodenrasters gehen wir davon aus, dass das Koordinatensystem unserer Eingabedaten über eine Höhenachse verfügt (üblicherweise die z -Achse). Außerdem gehen wir davon aus, dass sich der Boden in unseren Daten unten befindet, er also auf der Höhenachse die niedrigsten Koordinaten aufweist (ENU). Zunächst werden alle Punkte anhand der beiden anderen Koordinatensystemachsen in ein zweidimensionales Raster gruppiert. Für jede Rasterzelle werden dann die Koordinaten der z -Achse sortiert und der Wert des 0,05 Quantils als Höhenwert der Zelle genutzt. Wir verwenden das 0,05 Quantil und nicht den niedrigsten Wert, um mit möglichen Ausreißern, die sich z.B. aus Fehlmessungen ergeben, umgehen zu können.

Da nicht sichergestellt ist, dass jede Zelle überhaupt über Messwerte auf dem Boden verfügt, erfolgt nun ein Schritt, bei dem in einem rekursiven Vorgehen ausgehend von einer Startzelle versucht wird, alle Zellen des (Boden-)Rasters zu traversieren. Dabei wird ein Kriterium für die maximale Steilheit der Bodenebene angenommen. Wenn eine benachbarte Zelle der gerade aktuellen Zelle einen Höhenwert aufweist, dessen Abweichung vom Höhenwert der aktuellen Zelle dieses Steilheitskriterium nicht verletzt, gilt sie als traversierbar und wird anschließend als aktuelle Zelle verwendet. Zellen, die bei diesem Traversieren nicht erreicht werden können, werden aus dem Bodenraster entfernt. Bei der Ermittlung der Startzelle kommen drei Kriterien zur Anwendung.

1. Höhenwert der Zelle liegt zwischen dem 0,05 und 0,15 Quantil der Höhenwerte aller Zellen. Dies dient ebenfalls zur Abmilderung von Ausreißereffekten.
2. Anzahl der erreichbaren Nachbarzellen. Eine Zelle wird als Startzelle bevorzugt wenn diese Zahl höher ist.
3. Nähe zum Mittelpunkt des Rasters. Eine Zelle wird als Startzelle bevorzugt wenn sie mehr in der Mitte des Rasters liegt.

5.2 Segmentierung und Filtern

Für die Segmentierung wird ein Verfahren zum Regionenwachstum verwendet, bei dem Punkte, die in einem bestimmten Abstand zueinander liegen, als zum selben Segment gehörig angesehen werden. Das einzige Kriterium für die Segmentierung sind also die Koordinaten der Punkte bzw. deren euklidische Abstände.

Nach dem Segmentieren erfolgt ein Filtern, bei dem die Segmente anhand von Merkmalen wie Größe, Seitenverhältnis und Anzahl der Punkte gefiltert werden. Dies dient dazu, Segmente nicht weiter verarbeiten zu müssen, die entweder mit Sicherheit keiner Person zuzurechnen sind bzw. die sich nicht als Person erkennen lassen, da sie beispielsweise zu wenig Punkte aufweisen um für

eine Klassifizierung aussagekräftig genug zu sein. Es wird so also eine weitere Datenreduktion erzielt.

6 Personendetektion

Für die Detektion von Personen verwenden wir eine Variante des „Implicit Shape Model“-Ansatzes. Wir erhoffen uns davon sowohl mit der in Abschnitt 2 erwähnten geringen Datendichte umgehen zu können, als auch den Umfang der benötigten Trainingsdaten zu limitieren.

Die Personendetektion erhält als Eingabe die in der Vorverarbeitung ermittelten segmentierten Punktwolken, wobei wir jedes Segment als mögliche Person betrachten (im Folgenden Personenkandidat genannt). Die Aufgabe besteht also darin, diese Personenkandidaten den Klassen „Person“ und „keine Person“ zuzuordnen.

Die einzelnen Schritte unseres Verfahrens werden hier zunächst aufgelistet. Anschließend werden in mehreren Unterabschnitten Teilaspekte des Verfahrens näher erläutert. Die Schritte sind:

1. Extrahieren von Merkmalen für die 3D-Punkte des Personenkandidaten.
2. Für jedes extrahierte Merkmal: Suche nach einem dazu passenden geometrischen Wort im Wörterbuch.
3. Abstimmprozess: Jedes ermittelte geometrische Wort gibt eine oder mehrere Stimmen für die Position eines Objektes ab.
4. Gewichten der möglichen Objektpositionen basierend auf den Stimmen für die jeweilige Objektposition bzw. benachbarter Objektpositionen derselben Klasse (aktuell gibt es nur die Klasse „Person“. Das Verfahren ist aber in der Lage mit beliebig vielen Klassen umzugehen).
5. Entfernen aller Objektpositionen deren ermitteltes Gewicht unter einem bestimmten Schwellwert liegt.
6. Wenn unter den verbleibenden Objektpositionen derselben Klasse mehrere im geringen Abstand zueinander liegen werden sie zusammengefasst.
7. Ausgabe der verbleibenden Objektpositionen als detektierte Personen.

6.1 Merkmalsextraktion

Die bei der Merkmalsextraktion extrahierten Merkmale dienen zum Beschreiben der geometrischen Wörter des Wörterbuchs und zum lokalen Beschreiben der Punktwolken der verarbeiteten Personenkandidaten. Sie findet daher sowohl im Trainingsprozess als auch im eigentlichen Klassifikationsprozess statt, was im Hinblick auf die schritthaltende Verarbeitung Anforderungen an die Komplexität der Merkmalsextraktion stellt. Es ist natürlich möglich auch komplexe Merkmale zu verwenden, wenn diese dafür im Klassifikationsprozess nur für eine kleinere Auswahl der Daten extrahiert werden. Hierbei stellt sich dann jedoch die Frage, wie eine solche Auswahl der relevanten Daten erfolgt. Wie VELIZHEV et al. (2012) ausführen ist eine solche Auswahl und Ermittlung einiger weniger wohldefinierter Merkmale anfällig für Effekte, die sich durch Rauschen und Verdeckungen ergeben. Deswegen ist es vorzuziehen, Merkmale für sehr viele zufällig ausgewählte bzw. für alle Punkte einer Punktwolke zu extrahieren. In unserem Verfahren verwenden wir die von JOHNSON & HEBERT (1999) vorgestellten *Spin Images* als Merkmalsart, wobei wir planen in Zukunft auch die Verwendung anderer Merkmalsarten zu untersuchen.

6.2 Struktur des Wörterbuchs und Training

Das Wörterbuch umfasst alle geometrischen Wörter die im Klassifikationsprozess verwendet werden. Es stellt somit das Ergebnis des Trainingsprozesses dar. Die geometrischen Wörter selbst verfügen zum einen über ein Merkmalshistogramm, durch welches sie beschrieben werden, zum anderen über ein bzw. mehrere Stimmen für Objektpositionen. Diesen Stimmen wiederum ist neben einem Vektor vom Wort zur Objektposition auch ein Gewichtungsfaktor zugeordnet, es haben also nicht alle Stimmen dasselbe Gewicht.

Im Trainingsprozess erfolgen zunächst für die einzelnen Trainingsdatensätze eine Merkmalsextraktion und eine Bestimmung des Mittelpunkts. Anschließend wird für jedes extrahierte Merkmal ein neues Wort mit einer Stimme für die Objektklasse des Trainingsdatensatzes erzeugt. Dabei wird der ermittelte Mittelpunkt als Objektposition der Stimme genutzt und ihr Gewicht mit 1 initialisiert. Nachdem alle Trainingsdatensätze verarbeitet wurden, erfolgt ein k -means Clustering der erstellten Wörter, bei dem ihrer Merkmalsausprägung nach ähnliche Wörter zusammengefasst werden. Hierdurch wird der Umfang des Wörterbuchs reduziert.

Nach dem Zusammenfassen verfügen die hieraus resultierenden fusionierten Wörter über alle Stimmen der ursprünglichen Wörter aus denen sie entstanden sind. Die Stimmen eines Wortes werden anschließend gruppiert nach Objektklassen anhand der Ähnlichkeit der Positionen, für die sie stimmen, ebenfalls zusammengefasst. Hierbei gibt es einen Parameter der definiert bis zu welchem Abstand ähnliche Stimmen zusammengefasst werden. Die danach verbleibenden Stimmen eines Wortes werden so neu gewichtet, dass das Gesamtgewicht aller Stimmen des Wortes wieder 1 ergibt. Ein Wort, welches viele verschiedene Stimmen abgibt, wird hierdurch weniger stark für eine bestimmte Klasse und Position stimmen als ein Wort, welches wenige bzw. nur eine Stimme abgibt. Hierdurch wird das Stimmgewicht von Wörtern, die weniger aussagekräftig sind, gegenüber den aussagekräftigeren Wörtern reduziert.

6.3 Gewichten der möglichen Objektpositionen

Nach dem eigentlichen Abstimmprozess gibt es in unserem Verfahren eine Vielzahl möglicher Objektpositionen, da jedes gefundene geometrische Wort mindestens eine Stimme für eine Position abgegeben hat, meist jedoch sogar deutlich mehr als eine. Aus diesen sollen nun diejenigen bestimmt werden, bei denen es sich tatsächlich um die Position eines Objektes handelt. Dabei ist davon auszugehen, dass Stimmen, die für eine tatsächliche Objektposition abgegeben wurden, für Positionen gestimmt haben, die in einem kleinen Umkreis um die tatsächliche Objektposition liegen. Falsche Stimmen sollten hingegen für zufällig im Raum verteilte Positionen gestimmt haben. Daher wird ein hohes Stimmgewicht auf den Umkreis der tatsächlichen Objektpositionen fallen. Bei der Gewichtung der möglichen Objektpositionen betrachten wir daher das Gewicht anderer möglicher Objektpositionen in der Umgebung, wofür das folgende Vorgehen zur Anwendung kommt.

Zunächst findet eine Normalisierung der Stimmgewichte aller abgegebenen Stimmen dahingehend statt, dass ihre Summe 1 ergibt. Dies erlaubt es uns, später die Klassifizierungsentscheidung mit einfachen Schwellwerten zu treffen und dabei robust gegen die schwankende Anzahl abgegebener Stimmen zu sein. Es stellt jedoch auch einen Nachteil dar, wenn ein Personenkandidat tatsächlich

aus zwei oder mehr Personen besteht bzw. sehr viel Rauschen in der Punktwolke des Personenkandidaten vorhanden ist. In diesem Fall würde sich das zur Verfügung stehende Stimmgewicht weiter verteilen was die Gefahr birgt, dass der Schwellwert nicht mehr erreicht werden kann.

Nach der Normalisierung werden für jede mögliche Position die benachbarten möglichen Positionen gesucht und deren Stimmgewichte zu dem Stimmgewicht der Position selbst hinzuaddiert. Hierbei wird die Entfernung zwischen den beiden Positionen mithilfe der Gauß-Funktion zur Normalverteilung berücksichtigt. Je weiter die benachbarte Position entfernt ist, umso geringer ist der Anteil ihres Stimmgewichtes, der addiert wird.

Nach diesem Gewichtungsschritt der möglichen Objektpositionen findet die Klassifizierungsentscheidung anhand eines Schwellwertes statt. Alle Positionen, deren resultierendes Gewicht unter dem Schwellwert liegt, werden entfernt. Da anschließend häufig mehrere Objektpositionen in engem Umkreis zueinander verbleiben, werden diese dann noch zusammengefasst, wobei die Position mit dem höchsten Stimmgewicht dann die verbleibende ist.

7 Experimentelle Untersuchung

Unsere Verfahren bestehend aus der Vorverarbeitung und der Personendetektion wurden sowohl im Hinblick auf ihr Laufzeitverhalten als auch in Bezug auf die Güte ihrer Ergebnisse untersucht. Für die Experimente haben wir auf zwei Datensätze zurückgegriffen, die wir mit dem Messfahrzeug „MODISSA“ aufgezeichnet haben, welches in Abb. 4 dargestellt ist. Dieses Fahrzeug verfügt u.a. über vier LiDAR-Sensoren wobei es sich um zwei Velodyne HDL-64E und zwei Velodyne VLP-16 handelt. Bei unseren Experimenten verwendeten wir nur Messdaten der HDL-64E.



Abb. 4: Sensorfahrzeug „MODISSA“ von Fraunhofer IOSB. Dieses verfügt über verschiedene Sensoren, zu denen neben mehreren Kameras auch vier LiDAR-Sensoren gehören.

Die Datensätze bestehen jeweils aus mehreren Scans beider LiDAR-Sensoren. Als „Scan“ bezeichnen wir dabei eine Umdrehung des Sensors, wobei jeder Scan als einzelne Punktwolke vorliegt. Da die HDL-64E ca. 1,3 Millionen Messungen pro Sekunde durchführen und wir sie mit 10

Umdrehungen pro Sekunde betreiben umfasst ein einzelner Scan ca. 130.000 Messungen. Die daraus resultierenden Punktwolken sind jedoch kleiner, da nicht jede Messung zu einem verwertbaren Ergebnis führt (z.B. fehlende Pulsechos im Bereich des Himmels).

Bei dem ersten verwendeten Datensatz handelt es sich um Daten einer Messkampagne, welche im April 2016 auf dem Stammgelände der Technischen Universität München durchgeführt wurde. Dieser Datensatz enthält 3D-Messdaten einer Vielzahl an Personen im urbanen Umfeld (Campus der TU München). Für den zweiten Datensatz wurden Szenen mit jeweils einer Person gestellt. Dieser wurde primär im Rahmen des Trainings der Personendetektion verwendet. Weitere Trainingsdaten stammten jedoch auch aus dem ersten Datensatz.

7.1 Training und Parametrisierung der Personendetektion

Zur Erstellung der für den Personendetektor verwendeten Trainingsdaten wurden ganze Scans der Datensätze zunächst durch die Vorverarbeitung unseres Verfahrens segmentiert und gefiltert. Anschließend wurden die daraus entstehenden Segmente interaktiv durch visuelle Unterscheidung den Klassen „Person“ und „Keine Person“ zugeordnet, um dadurch die Trainingsdaten zu erzeugen. So entstanden 243 Trainingsdatensätze für die Klasse „Person“ und 1.095 für die Klasse „keine Person“, welche dann im Training verwendet wurden. Es erfolgte also ein Training sowohl mit Positiv- als auch mit Negativbeispielen, was sich gegenüber einem Training nur mit Positivbeispielen als vorteilhaft in Bezug auf die Qualität der Ergebnisse herausgestellt hat.

Wie im vorherigen Abschnitt erwähnt wird im Verlauf des Trainings zunächst für jedes ermittelte Merkmal in den Trainingsdatensätzen ein Wort und eine Stimme generiert, die dann entsprechend ihrer Ähnlichkeit zusammengefasst werden. Für den verwendeten Umfang an Trainingsdaten wurden so zunächst 353.782 Wörter und Stimmen generiert. Bei der anschließenden Zusammenfassung von Wörtern und Stimmen wurde die Verwendung unterschiedlicher Parameter empirisch untersucht. Dabei zeigt sich, dass ein zu starkes Zusammenfassen ähnlicher Wörter und damit ein zu kleines Wörterbuch negative Auswirkungen auf die Laufzeit während der Detektion haben. Dies liegt aber darin begründet, dass bei einem zu starken Zusammenfassen einzelne Wörter durchschnittlich über mehr Stimmen verfügen. Ein effizientes Durchsuchen auch eines größeren Wörterbuchs ist gut mit Suchstrukturen wie k -d-Bäumen möglich. Die Berücksichtigung einer größeren Anzahl abgegebener Stimmen im Detektionsprozess kann sich daher stärker auf die Laufzeit auswirken, als die Suche nach einem passenden Wort in einem größeren Wörterbuch. Gute Ergebnisse konnten wir mit einer Wörterbuchgröße von 10 % des ursprünglichen Umfangs erzielen, was sich in der Literatur auch für ähnliche Verfahren bestätigt (KNOPP et al. 2010). Für die einzelnen Wörter wurden Stimmen für Positionen, die bis zu 40 cm voneinander entfernt sind zusammengefasst. Im Ergebnis erhielten wir so ein Wörterbuch bestehend aus 35.378 Wörtern und 163.788 Stimmen, welches wir für die weiteren Untersuchungen verwendet haben.

Die im Hinblick auf die Detektionsrate und zum Teil auch Laufzeit hin wichtigsten Parameter unseres Verfahrens sind der Schwellwert für eine erfolgreiche Erkennung und der Umkreis, in dem benachbarte Stimmen bei der Gewichtung der möglichen Objektpositionen berücksichtigt werden. Hierfür haben wir in dem Experiment zuvor empirisch ermittelte Werte verwendet.

7.2 Laufzeit

Zur Untersuchung der Laufzeit wurden insgesamt 100 Scans verarbeitet und die dafür benötigte Zeit gemessen. Für diese Untersuchung wurde ein Computer mit einem Intel Core i7-3930K Prozessor verwendet. Die implementierten Verfahren liefen auf einer virtuellen Maschine, der sechs virtuelle Kerne des Prozessors zur Verfügung standen. Hierbei zeigte sich, dass das Ziel einer schritthaltenden Verarbeitung zumindest mit der verwendeten Hardware von uns aktuell noch nicht komplett erreicht wird. Im Schnitt benötigte die Verarbeitung eines Scans 1.287 ms, sie schwankt dabei jedoch stark zwischen 268 ms und 6.194 ms. Dieses schwankende Laufzeitverhalten muss von uns noch näher untersucht werden, liegt aber zum Teil in der jeweiligen Komplexität der aufgenommenen Szene begründet. In einigen Szenen lässt bereits die Vorverarbeitung nur noch wenige Daten übrig. In anderen, z.B. wenn viel Vegetation vorhanden ist, müssen mehr Segmente durch das Verfahren zur Personendetektion verarbeitet werden.

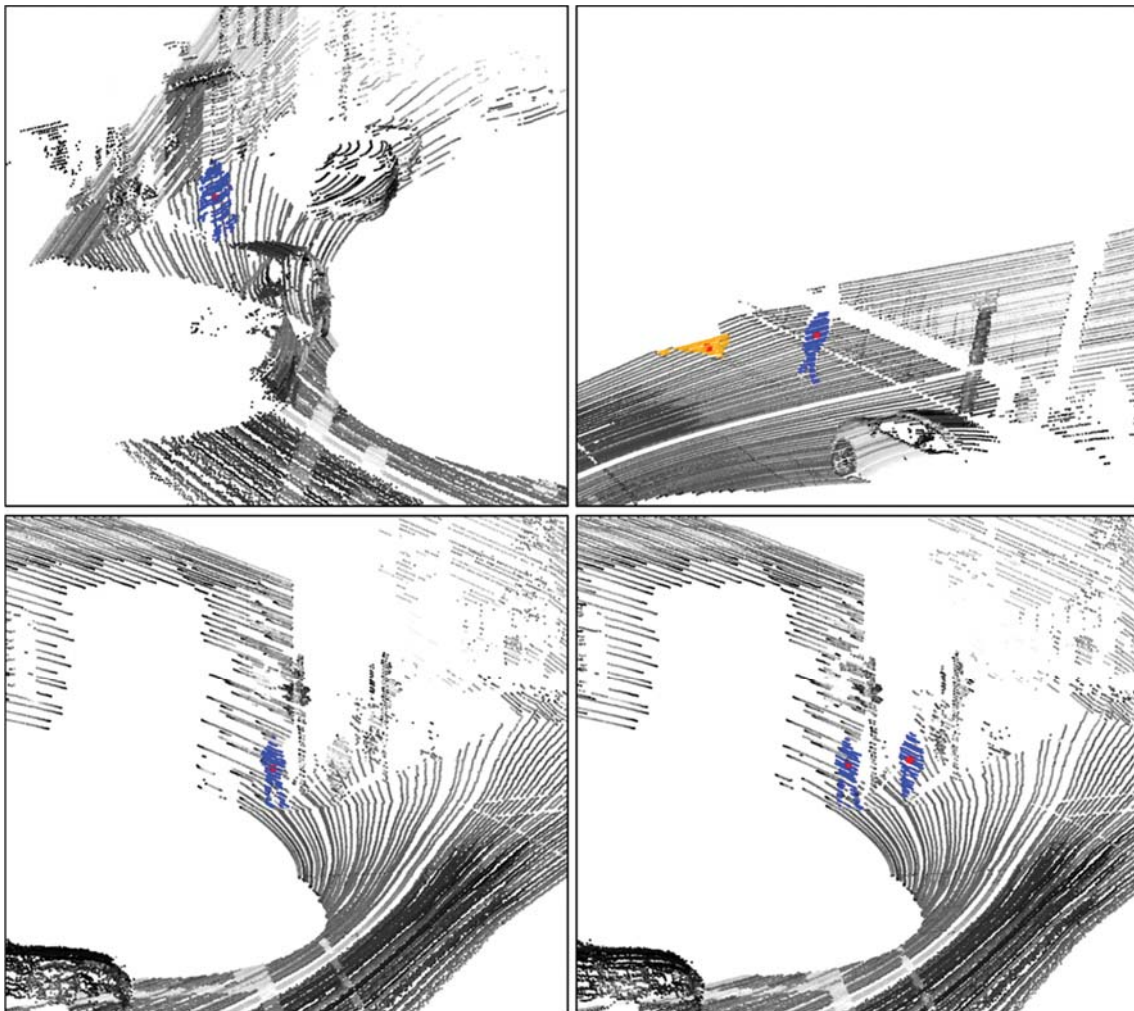


Abb. 5: Detektionsergebnisse: Oben links: Korrekt detektierte Person. Oben rechts: Korrekte (blau) sowie fehlerhafte Detektion (orange), wobei es sich um ein teilweise aufgenommenes Fahrzeug handelt. Unten: Eine Person wird in zwei aufeinanderfolgenden Scans nur im zweiten erkannt.

7.3 Qualität der Personendetektion

Zur Bewertung der Detektionsleistung wurden Ausschnitte aus den auch bei der Laufzeituntersuchung verarbeiteten Daten untersucht, um Auffälligkeiten bei den Resultaten festzustellen. Auch sollte näher untersucht und bestimmt werden, wie viele korrekt bzw. falsch erkannte Personen es in den Datensätzen gibt.

In den untersuchten Daten konnten wir ca. 76 % der vorhandenen Personen detektieren (Recall). Dabei lag der Anteil der korrekten Detektionen bei ca. 66 % (Precision). In Abb. 5 sind verschiedene Beispiele für Ergebnisse unseres Personendetektors dargestellt. Oben rechts sieht man ein Beispiel für eine Fehldetektion, bei der es sich um ein nur teilweise in der Punktwolke sichtbares Fahrzeug handelt. Diese Fehldetektion ist wohl u.a. auch auf Schwächen in unseren Trainingsdaten zurückzuführen, da in den verwendeten Negativbeispielen kaum Fahrzeuge vorhanden waren.

Im unteren Teil der Abbildung ist ein Phänomen beispielhaft dargestellt, welches wir bei der Untersuchung der Detektionsergebnisse mehrerer aufeinanderfolgender Scans häufiger beobachten konnten. Fehldetektionen aber auch Nichtdetektionen traten oft nur in einzelnen Scans auf. Hier besteht die Möglichkeit, durch die Berücksichtigung mehrerer im Zeitverlauf hintereinander aufgenommener Scans die Detektionsergebnisse in diesen gegeneinander zu plausibilisieren. So könnten Personen z.B. erst dann als detektiert gelten, wenn sie in mehreren Scans erkannt werden. Vorübergehende Nichtdetektionen ließen sich auf ähnliche Weise überbrücken, was insbesondere im Hinblick auf vorübergehende Verdeckungen wünschenswert ist. An dieser Stelle spielt natürlich auch das geplante Tracking der erkannten Personen eine Rolle.

8 Fazit und Ausblick

Dieser Beitrag stellt zunächst ein Design einer Methode vor, deren Ziel es ist personengefährdende Situationen mithilfe von LiDAR-Sensoren zu detektieren. Anschließend werden bereits realisierte Teile dieser Methode näher untersucht. Hierbei zeigen wir auf, dass die Detektion von Personen mithilfe von „Implicit Shape Models“ auch in 3D-Punktwolken mit verhältnismäßig geringer Dichtedichte grundsätzlich möglich ist. Wir zeigen jedoch auch, dass noch weitere Arbeiten und Verbesserungen an unserem Verfahren notwendig sind, um unser Ziel der schritthaltenden Verarbeitung zu erreichen und um die Rate von falsch bzw. nicht erkannten Personen zu reduzieren. Hierfür planen wir zunächst eine umfangreichere Untersuchung der verschiedenen Parameter, die in unser Verfahren eingehen, um für diese bessere Werte zu bestimmen. Außerdem planen wir die Verwendung anderer Merkmalstypen zu untersuchen sowie den Umfang der verwendeten Trainingsdaten zu vergrößern.

Wir erwarten, dass sich die Laufzeit unseres Verfahrens durch verschiedene Maßnahmen in der Zukunft noch verbessern lässt. Potenziale liegen dabei z.B. im Bereich der Parallelisierung. Diese wird zwar bereits genutzt, jedoch zeigt sich im Rahmen einer Untersuchung der Prozessorauslastung, dass es hier noch Verbesserungspotenzial gibt.

In zukünftigen Schritten planen wir zunächst das entwickelte Verfahren zur Personendetektion um die Detektion von Körperteilen und der Körperhaltung zu erweitern und anschließend auch die anderen Komponenten der entworfenen Methode zu realisieren.

9 Literaturverzeichnis

- GANDHI, T. & TRIVEDI, M. M., 2007: Pedestrian Protection Systems: Issues, Survey, and Challenges. *IEEE Transactions on Intelligent Transportation Systems* **8**(3), 413-430.
- JOHNSON, A. E. & HEBERT, M., 1999: Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(5), 433-449.
- JÜNGLING, K. & ARENS, M., 2011: View-invariant person re-identification with an Implicit Shape Model. 8th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 197-202.
- KELLER, C. G., DANG, T., FRITZ, H., JOOS, A., RABE, C. & GAVRILA, D. M., 2011: Active Pedestrian Safety by Automatic Braking and Evasive Steering. *IEEE Transactions on Intelligent Transportation Systems* **12**(4), 1292-1304.
- 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 2010, Springer-Verlag, 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.
- 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.
- 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. Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, 1297-1304.
- 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.
- SPINELLO, L., ARRAS, K. O., TRIEBEL, R. & SIEGWART, R., 2010: A Layered Approach to People Detection in 3D Range Data. Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, 1625.
- 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.

Bildqualität von optischen Fernerkundungsdaten

ANDREAS BRUNN¹, HORST WEICHEL¹ & RALF REULKE²

Zusammenfassung: Photogrammetrie und Fernerkundung bieten eine ganze Reihe unterschiedlicher Verfahren zur Ableitung geometrischer, radiometrischer und thematischer Informationen aus Bilddaten. Zur Erfassung der dafür benötigten Bilddaten stehen eine Vielzahl von Flugzeug- und Weltraumsensoren zur Verfügung.

Digitale Sensoren bieten auf Grund der Möglichkeiten der absoluten geometrischen und radiometrischen Kalibrierung vielversprechende Möglichkeiten zur Schaffung von Mehrwertprodukten wie digitale Höhenmodelle, Landnutzungskarten etc. Solche Kameras kombinieren die hohe geometrische Qualität mit den radiometrischen Standards von Erdbeobachtungssystemen. Für die Qualitätsbewertung optischer Fernerkundungsdaten sind verschiedene Standards und Spezifikationen verfügbar. Bei der Bestimmung der Bildqualität kann dabei zwischen (spektralen) radiometrischen und geometrischen Aspekten unterschieden werden. Normen enthalten verschiedene Messgrößen für Genauigkeitsprobleme (spektrale, radiometrische und geometrische Genauigkeit) sowie für Leistungsparameter wie SNR, MTF. Bildartefakte sind ein weiteres wichtiges Thema.

Der Beitrag führt in die Thematik ein und stellt einen neuen Ansatz zur effizienten radiometrischen Validierung und Kalibrierung operationeller satellitengetragener Fernerkundungssensoren vor.

1 Bestimmung der Bildgüte von Fernerkundungsdaten

Die Bestimmung der Güte von Fernerkundungsdaten lässt sich prinzipiell in spektrale, radiometrische und geometrische Aspekte unterscheiden. Für die Beschreibung lassen sich unterschiedliche Kenngrößen finden, die zum einen Genauigkeitsaspekte (z.B. Wandlung von DN in Strahlungsgrößen, geometrische Punktgenauigkeit wie z.B. CE90) aber auch Performancegrößen (SNR, MTF) betreffen. Hinzu kommen Artefakte und statistische Eigenschaften der Messfehler (Bildfehler).

Solche Bildfehler sind abhängig von

- Rohdaten (z.B. Detektor-Inhomogenitäten, Komprimierung oder Überstrahlung)
- der Vorverarbeitung (geometrische und radiometrische Korrekturen, Pansharpening, etc.)
- Fehler durch Algorithmen und mangelhafte Zusatzdaten (DGM, Atmosphärenparameter)
- Übertragungsfehler, etc.

Die Bestimmung der Güte kann auf unterschiedlichen Verarbeitungsniveaus erfolgen. Dies erfordert eine Vorverarbeitung der Bilddaten zur Herstellung einer vergleichbaren Datenqualität, die Gewinnung, Bereitstellung und Bearbeitung von Referenzdaten und Zusatzinformationen. Die Erfassung der Güte von Fernerkundungsdaten bezogen auf die vorgesehene Anwendung entspricht der Validierung.

¹ Planet Labs Germany GmbH, Kurfürstendamm 22, D-10719 Berlin,
E-Mail: [andreas.brunn, horst.weichel]@planet.com

² Humboldt-Universität zu Berlin, Institut für Informatik, D-10099 Berlin,
E-Mail: reulke@informatik.hu-berlin.de

Neben einer Charakterisierung und dem Vergleich von Fernerkundungssystemen bzw. Bilddaten können mit diesen Verfahren auch der Einfluss von Bildverarbeitungsalgorithmen (Kompression, Bildschärfung) untersucht werden. Für die Bewertung von Bildverarbeitungsalgorithmen sind oft nur vom Algorithmus abhängige spezielle Faktoren von Interesse. Die Untersuchung des Einflusses von Kompressionsalgorithmen kann sich beispielsweise auf die Untersuchung des Auftretens von Bildartefakten beschränken.

Die geometrische und radiometrische Genauigkeit sind ein Teil der Bildqualität und werden im Rahmen der Validierung bestimmt.

Der Beitrag stellt eine Methode zur radiometrischen Validierung und Kalibrierung satellitengetragener Kameras in der operationellen Phase vor, die sich auf Referenzmessungen am Boden stützt.

2 Radiometrische Kalibrierung

Für die meisten thematischen Anwendungen von Fernerkundungsdaten ist eine radiometrische Kalibrierung unerlässlich. Für die radiometrische Kalibrierung gibt es vielfältige wichtige Gründe. Ein Fernerkundungssensor liefert in der Regel Grauwerte, die keinen Bezug zu physikalischen Größen haben (SCHOWENGERDT 2007). Für eine effektive Anwendung der Daten, insbesondere wenn Daten unterschiedlicher Sensoren oder Daten aus unterschiedlichen Aufnahmezeitpunkten verglichen werden sollen, müssen die Daten in vergleichbare physikalische Größen umgerechnet werden. Diese physikalische Größe ist in der Regel die spektrale Strahldichte (radiance) und wird in der Einheit $W/m^2\mu m\ str$ angegeben. Die Bestimmung der Umwandlungskoeffizienten erfolgt durch Methoden der absoluten radiometrischen Kalibrierung.

Außerdem werden auch Methoden wie die Cross Kalibrierung verschiedener Sensoren oder das destripping von Zeilensensoren als radiometrische Kalibrierung bezeichnet. Diese Methoden werden hier aber nicht betrachtet.

Absolut radiometrische Kalibrierung beginnt für satellitengetragene Sensoren bereits vor dem Start in die Umlaufbahn am Boden im Labor. Dabei wird die spektrale Empfindlichkeit jedes Sensorelements in Abhängigkeit von der Bestrahlungsstärke gemessen und das Verhalten z.B. in Abhängigkeit von der Temperatur charakterisiert. Da die spektrale Empfindlichkeit gängiger Sensoren nicht stabil ist, sich insbesondere durch die Einflüsse wie Erschütterung und Temperaturunterschieden des Starts und in der frühen Orbitphase durch Ausgasung teilweise signifikant verändert, ist auch in der operationellen Phase des Instrumentes in der Umlaufbahn eine ständige Überwachung und ggf. erneute Kalibrierung erforderlich, wenn die radiometrische Genauigkeit erhalten bleiben soll. Während im Labor die Strahldichten am Detektor direkt bestimmt werden können, bedient man sich für die On-orbit Überwachung und Kalibrierung entweder auf der Erde befindliche Referenzflächen in Kombination mit Methoden zur Simulation und Entfernung atmosphärischer Einflüsse, oder Objekten, die sich außerhalb der Erdatmosphäre befinden.

2.1 Methoden der absoluten Kalibrierung

Einige Sensoren haben eingebaute Referenzquellen wie Lampen oder Sonnendiffusoren, die zur absoluten Kalibrierung genutzt werden können. Da diese sich normalerweise nicht im selben optischen Pfad wie die eigentlichen Ziele befinden und außerdem selbst gewissen Degradationen unterworfen sind, sind auch für solche Instrumente alternative Methoden von Vorteil.

Als solche alternativen Methoden haben sich die Kalibrierung anhand von Sternen und vor allem des Mondes sowie von homogenen Flächen am Boden (radiance based vicarious calibration) etabliert.

Für den Mond sind mittlerweile Modelle erstellt worden, mit denen sich die Irradianz des Mondes mit einer relativen Genauigkeit von ungefähr einem Prozent (BARNES et al. 2004; STONE & KIEFFER 2004) vorhersagen lässt. Für die Kalibrierung werden dann mit dem Sensor Bilder des Mondes aufgenommen und die so erzeugten Messwerte mit aus den Modellen der Mond Irradianz abgeleiteten Referenzwerten verglichen.

Für die Nutzung homogener Flächen am Boden werden gleichzeitig zur Aufnahme der Fläche mit der Fernerkundungskamera vom Orbit aus Referenzmessungen von Reflektanz und Irradianz am Boden aufgezeichnet. Aus diesen Referenzmessungen werden durch Modelle zum Strahlungstransfer durch die Atmosphäre, wie MODTRAN oder 6S, die am Sensor erwarteten Radianzen simuliert und mit den tatsächlichen Messwerten verglichen.

2.2 RadCaTS und RadCalNet

Ideal für die radiometrische Kalibrierung sind Flächen, die räumlich homogen, zeitlich invariant und atmosphärisch stabil, ohne große Wasserdampf- und Aerosolanteile sind. Eine solche Fläche ist z.B. Railroad Valley, das seit 1999 für die absolute Kalibrierung von Weltraumsensoren genutzt wird (THOME et al. 2003). Diese Testfläche wurde seitdem kontinuierlich als **Radiometric Calibration Test Site** (RadCaTS) weiterentwickelt, auf dem heute auch automatisiert Referenzmessungen der Boden Reflektanz und Irradianz im Wellenlängenbereich zwischen 400 und 2500nm durchgeführt werden (CZAPLA-MYERS et al. 2010).

RadCaTS ist Teil von RadCalNet welches als Initiative vom **Committee on Earth Observation Satellites** (CEOS) vier global verteilte automatisierte Sites zusammenfasst und die Daten derer gesammelt auf einem Internetportal zur Verfügung stellt. Neben Railroad Valley beinhaltet RadCalNet außerdem die Sites La Crau (Südfrankreich), Baotou (Innere Mongolei, China) und Gobabeb (Namibia). Das System ist derzeit noch im Testbetrieb und für Betatester auf Antrag zugänglich. Die operationelle Phase soll im zweiten Quartal 2017 beginnen. Instrumentierung und Messungen des RadCalNet sind dabei vereinheitlicht für alle Flächen identisch. Eine Zusammenfassung der verfügbaren Testgebiete findet sich in Tab. 1.

Tab. 1: Testflächen des RadCalNet (RADCALNET 2016)

Testfläche	Lat	Lon	Grösse
Railroad Valley	38.497	-115.690	1km x 1 km
Baotou	40.851	109.628	100m * 100m
La Crau	43.558	4.864	30m radius
Gobabeb	Noch nicht in Betrieb		

Auf jeder RadCalNet Testfläche werden täglich zwischen 9:00 und 15:00 Uhr Ortszeit Messungen im 30 Minuten Intervall gemacht. Gemessen wird dabei die Bodenreflektanz und die Sonneneinstrahlung aus der diverse atmosphärische Faktoren abgeleitet werden die wiederum dazu benutzt werden, um aus den gemessenen Bodenreflektanzwerten Reflektanzwerte über der Atmosphäre abzuleiten.

Das Datenpaket beinhaltet folgende Parameter: Reflektanz über der Atmosphäre (10 nm spektrale Auflösung), Reflektanz am Boden (10 nm spektrale Auflösung), Temperatur, Wasserdampf, Optische Dicke bei 550 nm und Angström Koeffizient. Für bis zu 13 wolkenfreie Messungen pro Tag. Diese Messungen können direkt mit simultan angefertigten kalibrierten Satellitenaufnahmen der jeweiligen Testfläche verglichen und dadurch die Genauigkeit der bisherigen Sensorkalibrierung abgeleitet und verbessert werden.

Bei Planet (früher BalckBridge AG) wird das Railroad Valley bereits seit 2009 als Kalibrierfläche für die RapidEye Sensoren genutzt. Zunächst ausschließlich mit manuell durchgeführten Messungen, später dann auch mit den atomatisierten RadCats Daten. Die Validierungen anhand von RapidEye Daten haben gezeigt, dass mit Hilfe der vicarious calibration Kalibrierungen mit einem Fehler von weniger als 4% möglich sind (BRUNN et al. 2016). Derzeit wird die Methode auf das gesamte RadCalNet Netzwerk und die gesamte Flotte von Planet Satelliten ausgebaut.

3 Fazit & Ausblick

In dem Beitrag wurde ein neuer Ansatz zur Kalibrierung von Fernerkundungssensoren gegeben. Durch Nutzung der automatisiert gewonnenen Referenzdaten von den RadCalNet Flächen steht eine zeitlich wesentlich höherfrequente Referenz zur Verfügung, die es erlaubt radiometrische Validierungen in wesentlich kürzeren Zeiträumen durchzuführen. Wir erwarten daraus eine Verbesserung der Genauigkeit sowie auch eine wesentliche Steigerung der Effizienz der Kalibrierprozesse. Außerdem wird durch die gemeinsame Nutzung derselben Flächen durch unterschiedliche Sensorbetreiber die Interoperabilität der Fernerkundungsdaten deutlich erhöht.

4 Literaturverzeichnis

- BARNES, R., EPLEE, R., PATT, F., KIEFFER, H., STONE, T., MEISTER, G., BUTLER, J. & MCCLAIN, C., 2004: Comparison of SeaWiFS measurements of the Moon with the U.S. Geological Survey lunar model. *Applied Optics* **43**(31), 5838-5854.
- BRUNN, A., WEICHEL, H., BAHLOUL, S., 2016: Absolute Calibration of a 5 Satellite Constellation Using Vicarious Calibration – 7 Years of Operational Experience. Conference on Characterization and Radiometric Calibration for Remote Sensing, Logan/Utah, Aug. 2016.
- CZAPLA-MYERS, J.S., THOME, K.J. & LEISSO, N.P., 2010: Radiometric calibration of earth-observing sensors using an automated test site at Railroad Valley, Nevada. *Canadian Journal of Remote Sensing* **36**(5), 474-487.
- SCHOWENGERDT, R., 2007: Remote Sensing, Models and Methods for Image Processing, Academic Press.
- RadCalNet, 2016: RadCalNet Quick Start Guide, https://www.radcalnet.org/resources/RadCalNetQuickstartGuide_20160915.pdf, last access on 19.01.2017
- STONE, T. & KIEFFER, H., 2004: Assessment of uncertainty in ROLO lunar irradiance for on-orbit calibration. SPIE 49th Annual Meeting on Optical Science and Technology, 300-310.
- THOME, K.J., BIGGAR, S.F. & WISNIEWSKI, W., 2003: Cross Comparison of EO-1 Sensors and Other Earth Resources Sensors to Landsat-7 ETM+ Using Railroad Valley Playa. *IEEE TGARS* **41**(6), 1180-1188.

DIN-Normungsarbeit in der Photogrammetrie und Fernerkundung – Stand und Perspektiven

SVEN BALTRUSCH¹ & RALF REULKE²

Zusammenfassung: Die Normungsarbeit ist eine Gemeinschaftsaufgabe, die von den Experten der interessierten Kreise und den Mitarbeitern von DIN erfüllt wird. Das DIN fungiert als „runder Tisch“, an dem Vertreter der "interessierten Kreise" (z. B. Wirtschaft, Anwender, Wissenschaft und Forschung, öffentliche Hand) konsensbasiert Normen und Standards markt- und zeitgerecht erarbeiten. sDIN-Normen sind Empfehlungen, deren Anwendung freiwillig ist. Eine Anwendungspflicht kann sich aufgrund von Rechts- oder Verwaltungsvorschriften sowie aufgrund von Verträgen oder sonstigen Rechtsgründen ergeben. DIN-Normen bilden einen Maßstab für einwandfreies technisches Verhalten.

Für die Bereiche Photogrammetrie und Fernerkundung existiert der Normungsausschuss NA 005-03-02 AA „Photogrammetrie und Fernerkundung“. Dieser Arbeitsausschuss bearbeitet die Terminologie-Norm DIN 18716, Photogrammetrie und Fernerkundung – Begriffe, die durch die Festlegung von Begriffsdefinitionen eine Grundlage für eindeutige fachspezifische Kommunikation schafft und bei der photogrammetrischen Auswertung angewandt wird.

Anforderungen an Photogrammetrische Produkte werden in der Normenreihe DIN 18740 festgelegt, die bereits vielfach bei der Vergabe von Dienstleistungen als anzuhaltender Standard eingesetzt wird. Beispielhaft stellt die Normenreihe einen wesentlichen Baustein bei den Vergabeleistungen von photogrammetrischen Dienstleistungen der deutschen Landesvermessungsbehörden (Bildflüge, Digitale Orthophotos und Digitale Höhenmodelle) dar.

Aktuelle Projekte beschäftigen sich mit der Bildqualität, mit Hyperspektralscannern und der Bodenbewegungsanalyse auf Basis der Radarfernerkundung. Der Beitrag gibt einen aktuellen Überblick und will die Perspektiven diskutieren.

1 Der DIN-Normungsausschuss NA005-03-02 „Photogrammetrie und Fernerkundung“

Der DIN-Normungsausschuss NA005-03-02 AA „Photogrammetrie und Fernerkundung“ setzt sich aus Vertretern von Behörden bzw. Verwaltungen, von Universitäten und Fachhochschulen sowie von Firmen zusammen. Außerdem werden projektbezogen Arbeitskreise eingerichtet, in denen Fachleute, zeitlich befristet, für das jeweilige Normungsvorhaben mitwirken.

Die Mitarbeit im DIN-Normungsausschuss erfolgt ehrenamtlich und wird von seinem Obmann Prof. Dr. rer. nat. Ralf Reulke koordiniert. Neben den Treffen der Ad hoc - Arbeitskreise finden zwei Sitzungen des Arbeitsausschusses im Jahr statt. Der Arbeitsausschuss ist weiterhin offen für neue Mitarbeiter. Internationale Aktivitäten sind in die Normungsarbeit integriert.

¹ Landesamt für innere Verwaltung Mecklenburg-Vorpommern, Amt für Geoinformation, Vermessungs- und Katasterwesen, 19059 Schwerin, Deutschland, E-Mail: sven.baltrusch@laiv-mv.de

² Humboldt-Universität zu Berlin, Institut für Informatik, 10099 Berlin, Deutschland, E-Mail: reulke@informatik.hu-berlin.de

Behörden und Industrie geben Themen vor, begleiten die Normentwicklung und wenden die Normen an. Die Normen vereinfachen das Verhältnis Auftraggeber – Auftragnehmer durch klare Festlegung der Ausgangssituation und der Angabe von Qualitätsmaßen. Ein entsprechendes Formblatt in der Normenreihe DIN 18740 formalisiert diese Beziehung.

Im Arbeitsausschuss werden aktuelle, dringende Probleme von Behörden und Industrie aufgegriffen und beraten (z.B. im Zusammenhang mit dem Normungsprojekt DIN 18740-6). Ein weiteres Beispiel ist die Beteiligung an einer Kameraevaluierung aktueller digitaler Kameras (DIN 18740-4), in Zusammenarbeit mit der EuroSDR und der DGPF.

Normen unterstützen die Lehrtätigkeit an den Universitäten, da sie fachliche Grundlagen komprimiert und übersichtlich darstellen.

2 DIN 18716

Die DIN 18716 legt Begriffe der Photogrammetrie und Fernerkundung zur Erfassung und Charakterisierung von Oberflächen mit abbildenden digitalen Sensorsystemen fest. Sie dient der Vereinheitlichung der Grundbegriffe und Benennungen.

Die Norm 18716 beinhaltet folgende Abschnitte:

- Allgemeine Begriffe
- Physikalische Grundlagen
- Aufnahmeverfahren
- Auswerteverfahren

Sie enthält Begriffe, die für mehr als eine Norm der Normreihe DIN 18740 relevant sind. Spezielle Begriffe, die nur für eine Norm bzw. einen Normteil relevant sind, werden nur in dieser Norm definiert. Bei der Überarbeitung einer Norm der Reihe DIN 18740 sind somit ggfs. Anpassungen in DIN 18716 erforderlich. Es wird in so einem Fall ein Änderungsblatt der DIN 18716 parallel veröffentlicht. In regelmäßigen Abständen werden die Änderungen in einer neuen Vollversion der DIN 18716 zusammengeführt.

3 Die Normenreihe DIN 18740

Die Normenreihe DIN 18740 legt Qualitätsanforderungen an photogrammetrische Produkte fest. Veröffentlicht wurden:

- DIN 18740-1:2001-11, Photogrammetrische Produkte - Teil 1: Anforderungen an Bildflug und analoges Luftbild (ohne Ersatz zurückgezogen)
- DIN 18740-2:2005-02, Photogrammetrische Produkte - Teil 2: Anforderungen an das gescannte Luftbild (ohne Ersatz zurückgezogen)
- DIN 18740-3:2015-08, Photogrammetrische Produkte - Teil 3: Anforderungen an das Orthobild
- DIN 18740-4:2016-04, Photogrammetrische Produkte - Teil 4: Anforderungen an digitale Kameras für Luftbild und Weltraumphotogrammetrie

- DIN 18740-5:2012-08, Photogrammetrische Produkte - Teil 5: Anforderungen an die Klassifizierung optischer Fernerkundungsdaten
- DIN 18740-6:2014-12, Photogrammetrische Produkte - Teil 6: Anforderungen an digitale Höhenmodelle
- DIN 18740-7:2014-12, Photogrammetrische Produkte - Teil 7: Anforderungen an das PAN-Sharpening

Grundsätzlich werden die Fachbegriffe, die in Ergänzung zu den in DIN 18716 definierten Begriffen für die jeweilige Norm erforderlich sind, definiert. Begriffe beinhalten stets eine englische Begriffsübersetzung.

In Anhängen sind Formblätter enthalten, welche die Definition des entsprechenden photogrammetrischen Produktes unterstützend ermöglichen sollen.

Im Folgenden werden die aktuellen Normen der Normenreihe DIN 18740 kurz vorgestellt.

3.1 DIN 18740-3: Anforderungen an das Orthobild

Diese Norm legt Anforderungen an die Herstellung von Orthobildern aus Luft- und Satellitenbildern und daraus abgeleitete Orthobildprodukte fest. Sie dient dazu, die Qualitätsanforderungen an Orthobilder und an Orthobildprodukte festzulegen und die Abhängigkeit der Qualität dieser Produkte von den Teilprozessen und den einzelnen Einflussgrößen zu beschreiben.

Gegenüber der Vorgängerversion (DIN 18740-3:2003-10) wurden folgende Änderungen vorgenommen:

- a) die überarbeitete Norm reduziert sich auf die Herstellung digitaler Endprodukte und deren Anforderungen;
- b) Anforderungen an analoge Endprodukte wurden gestrichen;
- c) Anhänge A und C wurden normativ;
- d) in Abschnitt 3 wurden die englischen Begriffe hinzugefügt;
- e) die Norm wurde redaktionell überarbeitet

Orthobilder sind die geometrische und weitgehend auch semantische Grundlage für die Herstellung und Laufendhaltung von Karten, die Gewinnung und Aktualisierung topographischer Grunddaten und die Datenverknüpfung in Geoinformationssystemen. Darüber hinaus dienen sie als Interpretations- und Planungshilfe in unterschiedlichen Fachgebieten.

DIN 18740-3 behandelt drei Schwerpunktthemen:

- allgemeine Anforderungen;
- Prüfung;
- Spezifikation für Orthobildprodukte.

3.2 DIN 18740-4: Anforderungen an digitale Kameras für Luftbild und Weltraumphotogrammetrie

Die DIN 18740-4 gilt für digitale Kameras auf Flugzeug- und Satellitenplattformen, die Bildaufnahme und die dabei entstehenden Bilddaten. Diese Norm gilt für Bilddaten von digitalen Flächen- oder digitalen Zeilenkameras, die für messtechnische Zwecke eingesetzt werden.

Die DIN 18740-4 wurde auf Basis der Vorgängerversion (DIN 18740-4:2007-09) geändert:

- a) In der überarbeiteten Version wird bei den Anforderungen kein Unterschied zwischen Luftbild- und Satellitenkameras gemacht;
- b) der Titel wurde geändert;
- c) die Norm wurde generell überarbeitet und an den Stand der Technik angepasst.

Eine digitale Messkamera stellt einen Sensor dar. Sie ist integriert in ein Sensorsystem, welches u. a. auch die Plattform, das Positionsbestimmungssystem und die Navigation sowie die Datenspeicherung umfasst. Detektoren erfassen das durch das optische System projizierte Bild. Die Erfassung erfolgt über einzelne Detektorelemente oder Detektorarrays, die in der Fokalebene angeordnet sind (räumliche Diskretisierung). Die aus den Detektorelementen ausgelesenen elektrischen Signale werden anschließend digitalisiert (radiometrische Diskretisierung) und als zweidimensionales geordnetes Datenfeld für die Weiterverarbeitung zur Verfügung gestellt. Die einzelnen Elemente dieses Datensatzes werden als Pixel bezeichnet.

Das auf diese Weise entstandene digitale Bild ist Grundlage für photogrammetrische und fernerkundliche Anwendungen sowie für Visualisierungsaufgaben.

DIN 18740-4 behandelt drei Schwerpunktthemen:

- Kamerasystem;
- Bildaufnahme;
- digitales Bild

3.3 DIN 18740-5: Anforderungen an die Klassifizierung optischer Fernerkundungsdaten

Diese Norm gilt für Landbedeckungsklassifizierungen auf der Grundlage multispektraler Luft- und Satellitenbilddaten, die mittels digitaler optischer Sensoren (Flächen- oder Zeilenkameras) hergestellt werden. Diese Sensoren müssen messtechnische Anforderungen (radiometrisch und geometrisch) erfüllen.

Satellitengestützte Sensoren stellen inzwischen eine geometrische Auflösung zur Verfügung, die bisher den Luftbildkameras vorbehalten war. Andererseits haben sich seit der Einführung digitaler Luftbildkameras auch die verwendeten geometrischen Auswertemethoden weitgehend angenähert, so dass frühere Unterscheidungsmerkmale zwischen flugzeug- und satellitengestützten Sensoren ihre Bedeutung verloren haben.

Die hohe geometrische Auflösung dieser digitalen Systeme bedingt dabei neue Algorithmen und Verfahren für die Auswertung und Interpretation der Bilddaten, insbesondere auch im Bereich der Klassifizierung und computergestützten Objekterkennung.

Digitale bildgebende Satellitensensoren sind im Gegensatz zu analogen (filmbasierten) Luftbildkameras in der Regel absolut kalibriert. Die radiometrischen Kalibrierinformationen ermöglichen eine Umrechnung der Bildgrauwerte (DN = „Digital Number“) in absolute Strahldichten. Für Satellitensysteme wird auf Grund des großen Abstandes zwischen Sensor und Objekt üblicherweise der Atmosphäreinfluss korrigiert.

Aus der absoluten radiometrischen Kalibrierbarkeit digitaler Sensoren ergeben sich neue Möglichkeiten für die Erstellung abgeleiteter Produkte, (z. B. bei der Klassifizierung), da diese Kameras die Eigenschaften hoher geometrischer Qualität mit denen radiometrischer Standards von Erdbeobachtungssystemen verbinden.

Die Klassifizierung setzt eine aus der fachlichen Sicht des Anwenders erfolgte genaue Festlegung und Beschreibung der zu klassifizierenden Objekte voraus. Sie erfordert eine

Objektidentifikation und Objektbeschreibung durch multispektrale und strukturelle Merkmale im Bild, die gegebenenfalls durch multitemporale Objekteigenschaften (z. B. Zu- oder Abnahme eines Merkmalswertes in einem festgelegten Zeitintervall) in ihrer Aussagefähigkeit ergänzt werden. Der Klassifizierungsprozess im engeren Sinne realisiert die Abbildung der aus den Bilddaten abgeleiteten Objekte auf die fachlichen, anwenderspezifischen Objektfestlegungen. Die Klassifizierung erfordert eine Vorverarbeitung der Bilddaten zur Herstellung einer vergleichbaren Datenqualität, die Gewinnung, Bereitstellung und Bearbeitung von Referenzdaten und Zusatzinformationen sowie die Nachbearbeitung.

DIN 18740-5 behandelt drei Schwerpunktthemen:

- Anforderungen an Sensoren und Aufnahmebedingungen;
- Anforderungen an Algorithmen und Verfahren;
- Genauigkeitsanforderungen.

3.4 DIN 18740-6: Anforderungen an digitale Höhenmodelle

Diese Norm gilt für die Ableitung von digitalen Höhenmodellen, die auf der Grundlage terrestrischer oder flugzeug- bzw. satellitengestützter Erfassungsmethoden hergestellt werden. Diese Sensoren müssen geometrische Anforderungen erfüllen. Es werden sowohl Digitale Geländemodelle als auch Digitale Oberflächenmodelle behandelt. Die Höhenmodelle finden u.a. Anwendung in geobezogenen Aufgabenstellungen der Landesvermessung, der Umweltverwaltung, der Lagerstättenforschung und der Archäologie.

Die dreidimensionale Punktmenge eines Digitalen Höhenmodells (DHM) kann durch verschiedene Methoden erfasst werden. Hierbei reicht die Bandweite von der Digitalisierung analoger Quellen über photogrammetrische Methoden bis zum Airborne Laserscanning. Die Verfahren kommen je nach Projektdefinition hinsichtlich der erforderlichen Genauigkeit und Punktdichte unter Beachtung der Wirtschaftlichkeit zum Einsatz.

Ebenfalls aufgabenspezifisch werden die erforderlichen Elemente, wie z.B. Massenpunkte und geländecharakteristische Strukturelemente zur Beschreibung von DHM festgelegt. Die Oberflächenmodellierung ist letztendlich vom mathematischen Ansatz, z.B. der Wahl des Interpolationsverfahrens, Methodik der Vermaschung etc., abhängig.

Zunehmend gewinnt die dritte Dimension bei der Erfüllung von geo-bezogenen Aufgabenstellungen an Bedeutung. Das Anforderungsprofil der tatsächlichen Anwendung setzt hierbei die Genauigkeitsanforderungen und folglich die Klassenzugehörigkeit fest.

DIN 18740-6 behandelt drei Schwerpunktthemen:

- Erfassungsmethoden
- Klassifizierung von DHM
- Produktableitung aus DHM-Datensätzen und ihr praktischer Einsatz

3.5 DIN 18740-7: Anforderungen an das PAN-Sharpning

Diese Norm gilt für Bilddaten, die durch Datenfusionsverfahren erzeugt wurden. Je nach eingesetztem Verfahren bilden fusionierte Bilddaten die Grundlage für Anwendungen in der Photogrammetrie, Kartographie und Fernerkundung.

Zur Verbesserung der Messgenauigkeit und damit auch zur Erweiterung der Verwendbarkeit von Sensoren ist es zweckmäßig, spektrale Daten unterschiedlicher Sensoren miteinander zu

fusionieren, um eine verbesserte geometrische und radiometrische spektrale Datengrundlage für eine digitale Weiterverarbeitung zu erhalten.

Daneben gibt es in der Photogrammetrie und Fernerkundung eine Vielzahl von zusätzlichen Anwendungen, die Fusionstechnologien verwenden:

- Auflösungsverbesserung
- Fusion von Bild- und Abstandsdaten (Orthophotoerstellung)
- Monitoring (Change Detection, Objektverfolgung/Tracking) mit Multisensorsystemen
- Orientierung (Kombination aus GPS, IMU, etc)
- Integration von Daten verteilter und mobiler Sensoren auf der Basis von Geo-Sensornetzwerken, Sensorschwärme, Muster- und Ereigniserkennung in verteilten Geodaten (-banken)

Diese Norm betrifft abbildende Sensoren und legt Anforderungen an Verfahren zum Pansharpening (Fusion auf Daten-/Signalniveau), der Fusion von niedriger aufgelösten multispektralen mit höher aufgelösten panchromatischen Daten, fest. Bei der Kombination von Daten unterschiedlicher Sensoren ist die exakte Koregistrierung der unterschiedlichen Datensätze im Subpixelbereich von entscheidender Bedeutung.

4 Aktuelle Normungsvorhaben

Die Normenreihe DIN 18740 wird entsprechend der Einschätzung der technischen Entwicklungen in regelmäßigen Abständen aktualisiert. Neben einer Fortschreibung der vorhandenen Normen werden weitere Normen erarbeitet. Im Folgenden werden aktuelle Normungsvorhaben kurz vorgestellt.

4.1 DIN 18740-8: Anforderungen an die Bildqualität (Güte optischer Fernerkundungsdaten)

Dieses Normungsvorhaben gilt für die Bestimmung der Güte optischer Fernerkundungsdaten auf der Grundlage multispektraler Luft- und Satellitenbilder, die mittels digitaler optischer Sensoren (Flächen- oder Zeilenkameras) hergestellt werden. Diese Sensoren müssen messtechnische Anforderungen (radiometrisch und geometrisch) erfüllen.

Satellitengestützte Sensoren stellen inzwischen eine geometrische Auflösung zur Verfügung, die bisher den Luftbildkameras vorbehalten war. Deshalb gilt das Normungsvorhaben sowohl für Daten von digitalen Luftbildkameras als auch von Satellitensystemen (z.B. SPOT, IKONOS, QUICKBIRD oder RapidEye). Die hohe geometrische Auflösung dieser digitalen Systeme bedingt dabei neue Algorithmen und Verfahren für die Auswertung und Interpretation der Bilddaten.

Die Bestimmung der Güte von Fernerkundungsdaten lässt sich prinzipiell in (spektral, true color) radiometrische und geometrische Aspekte unterscheiden. Für die Beschreibung lassen sich unterschiedliche Kenngrößen finden, die zum einen (spektral radiometrische und geometrische) Genauigkeitsaspekte (Wandlung von DN in Strahlungsgrößen, geometrische Punktgenauigkeit, z.B. CE90) aber auch Performancegrößen (SNR, MTF) betreffen. Hinzu kommen Artefakte und statistische Eigenschaften der Messfehler (Bildfehler).

Solche Bildfehler sind abhängig von

- der Vorverarbeitung (Rohdaten, PRNU, etc. korrigierte, Pansharpening),
- Fehler durch Algorithmen und mangelhafte Zusatzdaten (DGM, Atmosphärenparameter)
- Übertragungsfehler, etc.

Die Bestimmung der Güte kann auf unterschiedlichen Verarbeitungsniveaus erfolgen. Dies erfordert eine Vorverarbeitung der Bilddaten zur Herstellung einer vergleichbaren Datenqualität, die Gewinnung, Bereitstellung und Bearbeitung von Referenzdaten und Zusatzinformationen.

Die Erfassung der Güte von Fernerkundungsdaten bezogen auf die vorgesehene Anwendung entspricht der Validierung.

Neben Charakterisierung und Vergleich von Fernerkundungssystemen bzw. Bilddaten kann mit diesen Verfahren auch der Einfluss von Bildverarbeitungsalgorithmen (Kompression, Bildschärfung) untersucht werden. Für die Bewertung von Bildverarbeitungsalgorithmen sind jeweils nur, vom Algorithmus abhängige, spezielle Faktoren von Interesse. Bsp: Die Untersuchung des Einflusses von Kompressionsalgorithmen beschränkt sich auf die Untersuchung des Auftretens von Bildartefakten.

Die geometrische und radiometrische Genauigkeit sind ein Teil der Bildqualität und werden im Rahmen der Validierung bestimmt.

4.2 Normungsvorhaben: Anforderungen an Hyperspektralscanner

Hyperspektrale abbildende Sensoren erfassen einen zusammenhängenden Wellenlängenbereich mit einer hohen Anzahl von schmalbandigen Kanälen (abbildende Spektrometer). Die Wellenlängenbereiche entsprechen denen in der Fernerkundung (UV, VIS, NIR, SWIR, MIR, LWIR). Das Normungsvorhaben bezieht sich auf terrestrische Flugzeug- und Satellitensysteme. Dieses Normungsvorhaben befasst sich mit hyperspektralen Sensorsystemen und der Vorverarbeitung der damit gewonnenen Datensätze.

Eines der großen Probleme von hyperspektralen Sensoren ist die radiometrische und geometrische Kalibrierung. Einen Überblick zu Designs und wesentlichen Parameter zur Charakterisierung eines hyperspektralen Systems findet man in den Beitrag von SPIEB & REULKE (2017).

Die Kalibrierung eines hyperspektralen Sensors bezieht sich auf die jeweiligen spektralen und radiometrischen Messbereiche, des Blickwinkels, der spektralen, geometrischen und radiometrischen Auflösung, der Messgenauigkeit und der Stabilität und Zuverlässigkeit der Daten. Das bedeutet auch, dass die Kalibrierung eines Sensors in regelmäßigen Abständen kontrolliert und ggf. auch aktualisiert werden muss, um die Zuverlässigkeit der Daten über längere Operationszeiträume garantieren zu können. Das sollte entweder vom Hersteller des Sensors selbst, vom Anwender nach entsprechenden Vorschriften des Herstellers oder von autorisierten Labors durchgeführt werden.

4.3 Normungsvorhaben: Anforderungen an die Bodenbewegungsanalyse

Inhalt dieser Norm ist die satellitengestützte Bodenbewegungsüberwachung. Typische Anwender von radargestützten Beobachtungen des Untergrundes sind die Bergbau- und Energieindustrie, Anwendungen für Infrastruktur und Ingenieur- oder Bauwesen. Nutzer von solchen Systemen sind weiterhin Versicherungen sowie die Verteidigung und der Heimatschutz.

Fragestellungen beziehen sich meist auf Überwachungsaufgaben im Bergbau oder Pipelines, Überwachung für Baustellen, Deiche und Dämme. Natürlich ist die Identifizierung und Veränderungsdetektion ein wichtiges Problem.

Die Parameter des Sensors, aber auch der Prozessierung haben Einfluss auf die Qualität des Endprodukts. Zu den Sensorparametern zählen zum Beispiel Radiometrie und SNR des Systems. Geometrisch ist die Genauigkeit der Lokalisierung im Vergleich zu Referenzinformationen von Bedeutung.

Analog zu optischen Systemen lassen sich eine Reihe von Anforderungen an die Datenakquisition formulieren. Dazu zählen Aufnahmebedingungen (Probleme z.B. durch Kohärenzverlust infolge von Schnee, Wasser), hochgenaue Positionsbestimmung, Aufnahmeplanung unter Berücksichtigung von Orbit und topographische Gegebenheiten am Beobachtungsort.

Weitere Probleme, die im Rahmen einer Norm behandelt werden, sind die interferometrische Auswertung und die Anforderungen an die Koregistrierung, Abschließend beschäftigt sich die Norm mit einer Qualitätsbewertung der Daten, der Zwischen und Endprodukte. Die Produkte werden mit einer Dokumentation ausgeliefert.

5 Zusammenfassung und Ausblick

Die Standardisierungsprojekte werden unter Kollegen gut angenommen. Die Normen unterstützten die Vorbereitung, Durchführung und Bewertung von kommerziellen und wissenschaftlichen Projekten.

Die DIN Normen dienen dem Wissenstransfer und unterstützen den Transfer von wissenschaftlichen Ergebnissen hin zur praktischen Anwendungen von Firmen und der zuständigen Verwaltung. Die entwickelten Standards dienen auch als offizielle Spezifikationen für Ausschreibungen der öffentlichen Verwaltungen in Bezug auf die Dienstleistungsindustrie in Photogrammetrie, Fernerkundung und Geomatik.

6 Literaturverzeichnis

DIN 18740-3:2015-08, 2015: Photogrammetrische Produkte - Teil 3: Anforderungen an das Orthobild, Beuth Verlag GmbH.

DIN 18740-4:2016-04, 2016: Photogrammetrische Produkte - Teil 4: Anforderungen an digitale Kameras für Luftbild und Weltraumphotogrammetrie, Beuth Verlag GmbH.

DIN 18740-5:2012-08, 2012: Photogrammetrische Produkte - Teil 5: Anforderungen an die Klassifizierung optischer Fernerkundungsdaten, Beuth Verlag GmbH.

DIN 18740-6:2014-12, 2014: Photogrammetrische Produkte - Teil 6: Anforderungen an digitale Höhenmodelle, Beuth Verlag GmbH.

DIN 18740-7:2014-12, 2014: Photogrammetrische Produkte - Teil 7: Anforderungen an das PAN-Sharpener, Beuth Verlag GmbH.

DIN 18740-8 (Entwurfsstatus): Anforderungen an Digitale Höhenmodelle

DIN 18740-9 (Entwurfsstatus): Anforderungen an das PAN-Sharpener

SPIESS, CH. & REULKE, R., 2017: Vergleich der optischen Eigenschaften von hyperspektralen Kameras. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 26, T. Kersten (Hrsg.), 37. Wissenschaftlich-Technische Jahrestagung der DGPF, 8.-10. März 2017 an der Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Tagungsband auf CD-ROM, 390-401.

Use Cases and their Requirements on the Semantic Modeling of 3D Supply and Disposal Networks

IHAB HIJAZI^{1,2}, TATJANA KUTZNER¹ & THOMAS H. KOLBE¹

Abstract: Different utility network data models were developed by different industries to provide means to represent, exchange and store utility networks. Often, these network data models are developed to meet the needs of specific domains without considering the integrated representation of different network systems. Moreover, mutual relations between networks as well as their embedding into 3D urban space are not always supported. The CityGML extension UtilityNetworkADE aims at meeting the requirements of the complex modern urban utilities.

This paper aims at evaluating the capability of the different network models including the UtilityNetworkADE to meet the needs of network infrastructure. The requirements were extracted from a list of use cases that was defined by a number of experts from different domains. The investigation points out that, in most cases, the UtilityNetworkADE is capable of linking different network systems, providing an integrated view to understand the relation to city entities. Only regarding connectivity rules there is no exact concept to represent them in the UtilityNetworkADE up to now. Many examples are discussed for the limitations and capabilities of different network data models.

1 Introduction

Modern society depends on a stable and a complex array of networks to deliver fuel, water and wastewater, electricity, gas and communication. Utility network infrastructures require an improved model to manage their relation to other network systems and to provide an integrated view to understand the interaction between city entities and utility networks.

In GIS and CAD, distribution systems are typically modeled as networks. Such networks form a directed graph (also called digraph), where each connector, fixture or outlet is viewed as a vertex. Each uninterrupted stretch of wire or pipe is viewed as an edge. Flow networks are represented by a special kind of graph, called a tree, where the nodes are reachable from one starting node and where no cycles exist. When creating a graph for a flow network, questions that deal with the reachability in a digraph G could be answered like this (GOODRICH et al., 2014): Given vertices u and v , determine whether u reaches v (i.e. does a chain between u and v exist?). Find all the vertices of G that are reachable from a given vertex s . Determine the ancestor vertices for a given list of vertices. However, networks in traditional GIS data models do not manage the third and fourth (i.e. temporal) dimensions and also do not consider the representation of network objects together with city entities.

In recent years, the development of semantic 3D city models has allowed for new approaches to town planning and urban management (BILJECKI et al. 2015) such as emergency and catastrophe

¹ Technische Universität München, Lehrstuhl für Geoinformatik,
Arcisstraße 21, D-80333 München, E-Mail: [kutzner, ihab.hijazi, thomas.kolbe]@tum.de

² An-Najah National University, Department of Urban Planning Engineering, Nablus, Palestine

preparation planning, checking building developments, and utility networks. These 3D city models provide a representation of cities that goes beyond visualization in terms of the application that they can support. One standard for representing semantic 3D city models is the international OGC standard CityGML (GRÖGER et al. 2012) which is currently extended to support applications in the context of urban planning and geo-design. These extensions include: 1) a General Indicator Model which allows for linking indicators to city objects and for evaluating scenarios (ELFOULY et al. 2015); 2) a Dynamizer concept which extends static 3D city models by the ability to support variations of individual feature properties and associations over time (CHATURVEDI & KOLBE 2016); 3) an EnergyADE (KADEN et al. 2015) which aims at storing information for energy simulations and energy system modelling as well as solar energy analysis; the EnergyADE deals with different data qualities, levels of details and urban energy model complexities; and 4) a UtilityNetworkADE (BECKER et al. 2011; BECKER et al. 2012; KUTZNER & KOLBE 2016) that supports the modeling and simulation of supply and disposal networks in 3D city models as well as the fluxes of the commodities such as water, electricity and gas within these networks.

In this paper, utility networks in urban areas are considered. The paper documents use cases defined by a number of expertise from different views and domains such as: planning and simulation of district heating, electricity and freshwater networks, planning and operation of smart energy, supply and disposal networks, vulnerability assessment and disaster management, city system simulation and Smart Cities as well as facility management. The aim is to determine the requirements of these use cases. Different requirements of these situations were defined and a list of requirements was documented complying with their semantics, dimension, visualization, relationships, connection to sensors, and scale level. These are analyzed to evaluate data models that match the requirements in the best way.

The paper is subdivided into five sections. Section 2 presents use cases for utility networks and corresponding requirements. Section 3 introduces different utility network data models. Section 4 reviews these network models with the aim to identify what kind of information modeling is applied and if the methods comply with the identified utility network requirements. Section 5 concludes with a discussion about future research on utility network modeling.

2 Use cases and their requirements

2.1 Utility network use cases

To be able to determine the information need for utility network applications, an inventory of possible applications has been made. The use cases were developed in discussion with experts from different domains, such as storm drainage, water, electricity, energy planning, and facility management. The use cases presented below do not pretend to be complete, but reveal some tendencies. The following cases have been identified.

2.1.1 Storm drainage network

By nature, as cities thrive, urban areas expand. The permeable earth of urban areas that provide the natural drainage or soak the storm water has been replaced by roads, buildings, and other hard surfaces that no longer allow rainwater to soak into the ground. Therefore, now in urban areas,

man – not nature – must control the surface water as well as protect the water supply. So the storm drainage system (cf. figure 1) collects surface water into underground pipes and conveys it to a surface watercourse, streams, rivers and lakes. The storm drainage network system is comprised of pipes, manholes, and catch basins. Catch basins are the buried basins that collect runoff from the streets and other ground surfaces, and are typically located below curb grates seen in the streets. As long as everything functions properly, there is little concern about this

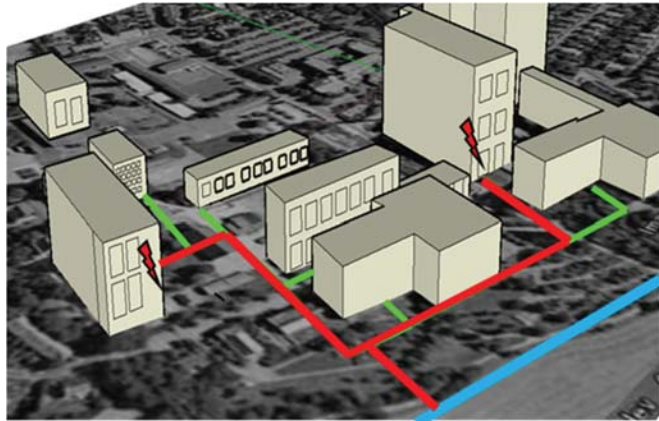


Fig. 1: Building discharge into the storm drainage network

system until a heavy storm occurs and water cannot drain rapidly enough from the streets. The capacity of the drains is determined by the flow that can be discharged by the drainage system and the storage available. When the hydraulic capacity is insufficient in a certain discharge situation, flooding occurs and water discharges to the surface. The occurrence of these events must be minimized, particularly because a lack of drainage systems is related to public health hazards (HIJAZI et al. 2012).

Water authorities have the responsibility to plan and manage the storm drainage network, they need to reduce the overflow in the storm drainage system, and the aim is to decrease the water overflow and to secure the city from flooding. The water authorities have the responsibility to know the building sites (area/volume, private/public), building roofs and sewer systems that are connected to the storm drainage network, the areas of roofs and the land surface type. This will allow the city to estimate the amount of water that will be discharged into the storm drainage system. Therefore, there is a need to calculate the building roof areas, the different built up areas and their covering materials. Moreover, this information is of great importance to the city in order to calculate the fees they need to charge from citizens.

The water authorities do not have yet a tool that allows them to search the network and quickly access the information that enables them to connect to the buildings and get information about its roof properties; also they need to get information about built up areas that are used as parking or as asphalt yards. The water authority needs to be able to get the area of buildings and non-permeable surfaces that are connected to specific parts of the network. Retrieval of the relevant information about the roofs and other surfaces is important in order to take actions to change these surfaces to a permeable one where water can soak in. Also, a textual description that is providing a reference to the locations of the building and their owners would be useful. Information about the area of the roofs, surfaces and building uses is of great importance.

2.1.2 Clean water act

The second use case refers to the inspection of waste water. City authorities perform regular inspections of some buildings in the city (e.g. chemical labs, factories) to ensure that the water discharged from these buildings does meet the safety standards of public water resources. The

inspection team needs to find the location of these elements inside these buildings to test whether they are working properly. The team needs a reliable tool that allows them to select a point in the network. The system should identify the buildings that contain the devices that are connected to the defined point and need inspection, and it should also show the exact location of these devices in the buildings and how they connect to the external network. Therefore, the proposed system should provide a suitable description for the location of these devices inside the building, e.g. the building and the storey they are within. Also, there is a need to establish a linkage to natural water resources such as streams. A method that provides the water authority with the ability to force rules on the network that can be linked to the natural water resource is useful in this use case. Using connectivity rule techniques, the network authority can control how to connect to other systems (HIJAZI et al. 2012)

2.1.3 Vulnerability assessment and disaster management emergency response

This use case refers to disaster management in different stages, i.e. vulnerability assessment and emergency situation. Vulnerability assessment is of great importance, as cities need to know the effects of natural disasters or man-made disasters on utility networks. Important information to know is the area that can be affected and how this can affect specific network systems. The city needs to know the buildings and city facilities that will be out of service based on this natural or man-made disaster, and also the interdependency between a specific network system and other network systems is of great importance. Having the different network systems and the city objects linked in one data model (cf. figure 2) will facilitate vulnerability assessment of the city infrastructure in order to develop mitigation plans.

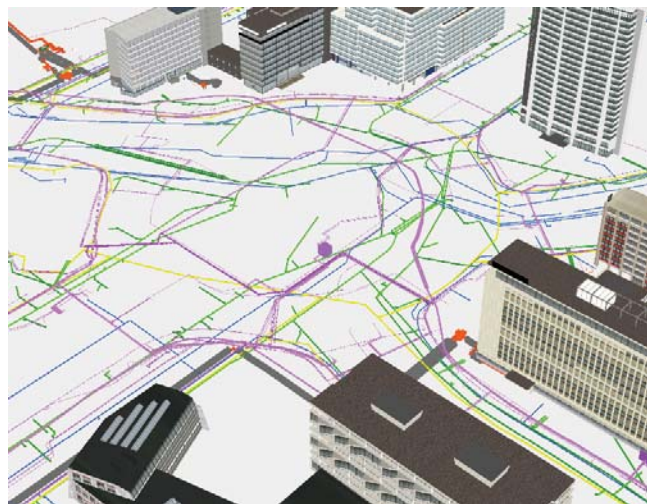


Fig. 2: Integrated representation of city objects and different network systems at the Ernst-Reuter-Platz in Berlin as realized in the SIMKAS 3D project (BECKER et al. 2012)

In emergency situations, when the precise location of a shut-off valve and the response in a timely manner are key issues, e.g. during fire incidents, the crew team should be able to disconnect any part(s) of the service system. The operational workflow starts with a notification of the facility management to help. The facility management needs an information system that can be used to define the location where the accident has occurred. It must identify which device (switch or transformer) should be turned off, so as to disconnect the relevant part(s) of the service system in the building or any city facility; this will also include a shut-off of the flow in other network systems that can be affected. Moreover, the system should also generate a notification report, list all the buildings and facilities that will be affected by the shut-off, and communicate this list to all the occupants of these buildings and facilities users. The list should include detailed information describing the location of the shut-off with reference to the building structure or any other city

objects, and also information on the building itself. Decisions based on unreliable information about the location of a shut-off can result in a delayed response to a critical situation and additional costs will result from the extra damage. In addition, damage in network systems can affect accessibility to the city and can therefore affect city life.

2.1.4 Maintenance operations

This use case focuses on warning residents about a scheduled or unscheduled maintenance. Facility managers need to perform maintenance operations, which can be either caused by a failure in the network or by planned (preventive) operation (due to date of expiration or cleaning). Both cases will cause an outage of service, because replacing of elements is required. Therefore, occupants of private buildings or public facilities need to be warned. The operational workflow starts by announcing the maintenance operation prior to its date after which the location of the shut-off valve must be defined and submitted to the field crew. The process in the field starts by closing the shut-off valve. In some cases, there is more than one option, and the best one to select will be that which affects the least number of occupants and city facilities (although this is not easy to define in the current system). The facility management teams need a tool that allows them to search the network and quickly access the information that enables them to contact the persons in the buildings or part of the buildings that will be affected; or, at best, since retrieval of the relevant information takes a sizable amount of time, and, thus, not all the information can be retrieved in time, the team has to make an assumption and generalize the announcement. The team needs to be able to input the location of the maintenance and have the GIS return an information product that includes a 3D view that describes the location of the shut-off and provides a textual description of the location of the shut-off in a human-oriented form. The view should provide a perspective of the space where the shut-off is, including the structure elements (e.g. manholes, walls or slabs) and the network segments under suspicion, the segments being connected to the structure elements both upstream and downstream, as well as any other structure elements that immediately surround the shut-off location. Finally, the team needs to define the network elements and their contained space, or city objects that would be out of service when there will be a shut-off. Also a textual description is needed providing a reference to the location of the space within the city feature (e.g. building or a long street).

2.1.5 Smart energy planning, simulation and operation

This use case considers issues related to electrical grid, heating and cooling simulations. Electricity planners and energy authorities need to know how a change in land use can affect the energy consumptions and production, therefore they need to relate a supply area to a specific object that provides the commodities to the area – this object can be considered the source for the supply area. Supply areas are in particular useful, when the detailed modeling of the supply lines is not available (KUTZNER & KOLBE 2016). Another important information is related to the ability to perform network analyses of the different utilities considering different scale levels. Moreover, another important issue to consider for simulation purposes is the ability to provide network objects themselves with information about their potential and current supply of commodities to them. In addition, considering the issue of security is of great importance, therefore the availability of supply areas can be important in this matter. This use case also requires that energy planning

authorities can define rules of how to connect network features to city entities and other network types. Furthermore, there is a need of coupling smart grids with buildings to simulate various issues concerning energy production, consumption and distribution.

2.1.6 Facility management

Facility management departments that are responsible for maintaining utility networks need an integrated modeling of the utilities with other city features (cf. figure 2). In addition, an important part of their work is their ability to access utility network objects for maintenance or repair, which can be done by humans or robots. Facility managers need to have a tool that enables them to navigate within networks for utility network analysis and to find easy access points, which allows performing the maintenance operations and minimizing disturbances.

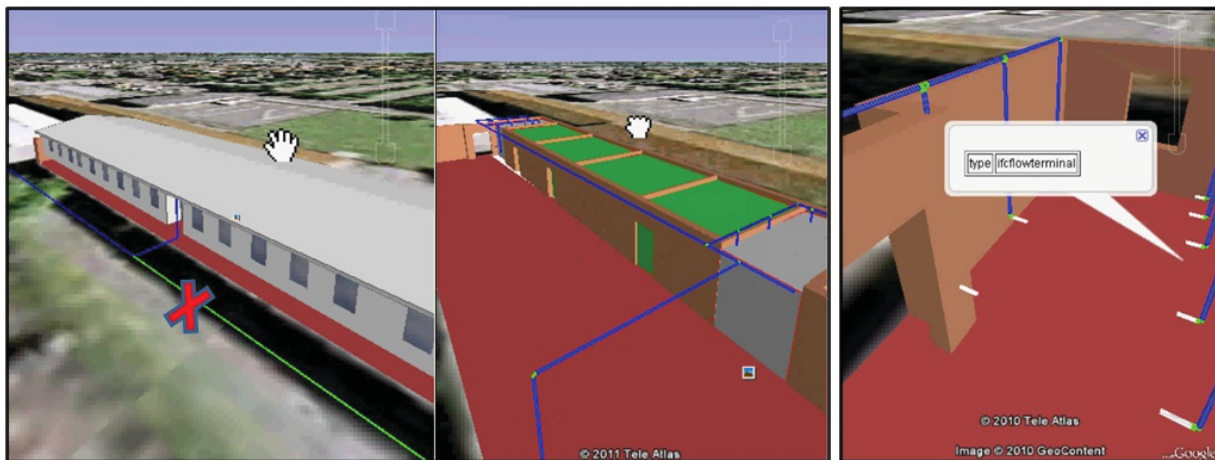


Fig. 3: Outdoor and indoor network integrated with 3D building features (HIJAZI et al. 2012)

2.2 Utility network requirements

On the basis of the use cases presented in the previous section, we define aspects that are relevant to the use cases. These aspects can be considered as modeling criteria or data modeling requirements regardless of the application or the final goal. These aspects will be used for the evaluation of selected network models that will be discussed in the next section. In all, a set of 18 generic and sub generic requirements were identified and subdivided into five groups:

- *Spatial scope*: The use cases provide us with the spatial extent that is required to facilitate the workflow in the use case. Some of the use cases operations need to be modeled on city block level, others on city level or district level. Providing a modeling mechanism to aggregate or disaggregate the models to different scopes and extents is of importance.
- *Visualization*: Topographic representation is one of the most important aspects for 2D and 3D representation of network objects. Realism and interaction are necessary for information to be understood quickly. In this part of the requirements investigation, we consider the following aspects:
 - Resolution and true 2D and 3D representation of network objects and other city objects
 - Representation methods such as iconic, mapped and realistic visualization.
 - Levels of detail as another method used to decrease the complexity of 3D representations.

- *Spatial relationships*: The investigation of the operation to be performed on the different use cases provides us with the relationships that should be preserved. These spatial relationships are relationships between network objects itself within the same network, between different networks or between network objects and other city features. These relationships can be summarized as follows:
 - There is a need to link in the data level between the network fixtures (e.g. lamps, sinks), and the city features e.g. space, street in order to select the spaces - or part of buildings or city features that would be out of service.
 - There is a need for a link between network elements itself in order to be able to trace the commodities that are moving through the network objects.
 - There is some network-to-network linkage that must be made between the network systems; i.e. the logical relationship between the hot water and the electricity network that needs to be maintained in the network.
 - The relation between the exterior and interior networks needs to be maintained.
 - Finally, there is a need to force constraints on how to establish the different relationships mentioned above.
- *Semantic representation*: Most of the use cases require a semantic categorization of the network objects based on their role in the network. Moreover, other semantic information related to the dimensionalities, materials, volume and geometric properties of network objects are required.
- Connection to *sensors* and time-variant attributes: Some of the decisions related to network management or planning requires up-to-date information about the status of the network elements or the amount of stocks available at a specific time.

3 Related network models

A range of data models for representing, exchanging, analyzing and storing utility network infrastructures exist already. This section provides a short overview of those models which are currently most relevant in the geospatial domain.

- The EU Directive INSPIRE provides the *INSPIRE Utility Networks* model (JRC 2013a) which is based on the INSPIRE Generic Network Model (JRC 2013b). The INSPIRE Utility Networks data model defines a 2D topological relationship between network objects and allows for representing five different types of networks (water, electricity, waste water, district heat and oil/gas/chemicals). However, the semantic categorization of network objects is basic, i.e. the data model defines, for instance, pipes and cables, but does not provide a further domain-specific classification of these network elements.
- The *Industry Foundation Classes* (IFC) is an ISO standard (ISO 16739:2013) which is predominantly used in Building Information Modeling. IFC provides a 2D and 3D representation of network objects. Relationships between network objects are described using a connectivity concept, which comprises both the physical and logical connectivity. Therefore, with the IFC data model it is possible to establish a linkage between different network types. IFC provides a rich semantic categorization of network objects based on their role in the network. However, the IFC data model was developed with the intention to provide a way to

model utilities at the building level. The integration of city scale network on small scale (large areas) is not supported.

- ArcGIS provides two sets of network data models to manage the logical and physical relations in a network. The *ESRI Geometric Network* model (ESRI 2017a) represents the basic structure for any utility network type. The network is composed from nodes and junctions that can be generated automatically by ArcGIS based on the physical connectivity of network objects that are represented as points and lines in the database. *ArcGIS Schematics* (ESRI 2017b) provides a mechanism to represent the logical relation in a network. Using a relationship class that represents relationships between network objects, it is possible to generate a graph that represents the linkage between different network objects. ArcGIS has a set of industrial-specific domain data models for gas, water and electricity that are customized based on the Geometric Network model. However, the model is lacking the topographic representation of network objects in 3D and also managing the logical relation between network objects is a challenge.
- The *CityGML UtilityNetworkADE* proposed by BECKER et al. (2011), BECKER et al. (2012) and KUTZNER & KOLBE (2016) aims, on the one hand, at providing “a common basis for the integration of the diverse models in order to facilitate joint analyses and visualization tasks” (BECKER et al. 2012), but, on the other hand, also intends to overcome shortcomings of existing network models with respect to the following characteristics: The data model should allow for the representation of heterogeneous networks, i.e. not only for specific types of networks, for a dual representation of network topography as well as topology and for a representation of topographic/graphic aspects (including 3D) as well as of functional aspects. Furthermore, the data model is to allow for a hierarchical modelling of networks and subnetworks as well as of components and subcomponents and for modelling interdependencies between network features and city objects.
- Another ISO standard, which allows for representing utility networks, is *SEDRIS* (Synthetic Environment Data Representation and Interchange Specification) (SEDRIS 2006). SEDRIS focuses on the representation and exchange of synthetic environments and allows for modelling networks for electricity, water and wastewater as well as for oil, gas and chemicals. SEDRIS was developed for training simulation and is to date only applied in the military domain. SEDRIS exhibits a similar good support regarding the characteristics mentioned above for the *CityGML UtilityNetworkADE*; disadvantages, however, are the high representational ambiguity of the format at runtime and the limited software support.
- PipelineML is a GML-based data interchange standard for the exchange of pipeline data focusing on the oil and gas industry which is currently under development by the OGC (OGC 2016). In its current stage of development, the standard focuses on distribution components and 2D geometries only, terminal elements such as pump stations are not considered, neither is a topological representation of networks.
- Data models for networks also are proposed in scientific literature. HALFAWY (2010) for instance, presents data models for water and wastewater networks, taking hereby also into account lifecycle aspects of network components such as maintenance operations or performance assessment.

4 Discussion

In the following, the use cases presented in section 2.1 and the data models introduced in section 3 are analyzed with respect to the requirements that were defined in section 2.2.

4.1 Evaluation of use cases with respect to requirements

The evaluation in this section considers the requirements with respect to the criteria that were defined in section 2. The columns in table 1 indicate which information will be required to analyze the use cases. Some of the use cases require the linkage to other city objects and the linkage to the specific elements within the city object, for example storm drainage network use cases required the connection to the building roofs. In addition, some information important for other use cases that include planning, is the ability to have a hierarchical view of the network in order to be able to do the analysis at different scale levels.

Tab. 1: Requirements relevant to different use cases

	Storm Drainage Network	Clean Water Act	Vulnerability and Emergency	Maintenance Operation	Facility Manage- ment	Smart Energy	
Sensors	•	++	+	•	++	++	
Semantics	++	++	++	++	++	++	
Spatial Relationships	• Network for indoor navigation	–	•	++	+	–	
	• Indoor to outdoor network	•	+	+	++	++	
	• Connectivity rules	+	++	+	•	++	
	• Network to City features	+	++	++	+	++	
	• Network itself	+	+	+	+	+	
	• Network to Network	–	+	++	+	+	++
Visualization	• Mapped	+	+	+	++	•	
	• Iconic	–	+	+	+	•	
	• LOD	++	++	+	+	+	
	• 2D utilities	++	++	++	++	+	•
	• 3D utilities	++	•	++	++	++	•
	• 2D city features	++	•	+	+	+	+
	• 3D city features	+	+	++	++	++	•
Spatial Scope	• City	++	++	+	•	–	+
	• Block	+	•	++	+	++	++
	• Building	•	•	++	++	++	+

– = not needed, • = basics, + = needed, ++ = very much needed

Table 1 clearly shows that the information requirements strongly depend on the application. Emergency response poses the highest requirements in total. To be able to support emergency and disaster management and perform rescue, response and navigation in timely manner, there is a need to have a linkage between different network systems and also to have a connection to city

objects. Therefore, geometrical, topological (connectivity and adjacency) and semantic information shall be available. Information about movable objects such as cars, city furniture, or indoor furniture is critical for transporting equipment for maintenance or emergencies.

Many applications need a time component indicating the status of the network objects, e.g. the availability of commodities, either through the supply line or through storage, at a specific point in time.

All applications need semantic information to be able to address the network operations in the best possible way. Although not of general interest, connectivity rules (how network objects connect to each other or other network types) are critical for planning and simulation of different scenarios. The table shows that several applications need real-time information and connection to sensors.

4.2 Evaluation of data models with respect to requirements

The overview of data models presented in section 3 does not pretend to be complete, but it reveals some tendencies. As mentioned above, most of the data models are based on the graph representation of utility network objects in order to provide the ability to perform reachability analysis. Table 2 provides an estimation of the fitness of the different data models for the requirements under consideration.

Tab. 2: Support of requirements by different data models

	INSPIRE Utility Networks	IFC	ArcGIS Utility Networks	SEDRIS	Pipeline ML	CityGML UtilityNetwork ADE
Sensors	–	•	–	+	–	•
Semantics	+	++	•	+	•	++
Spatial Relationships	• Network for indoor navigation	–	•	–	–	•
	• Indoor to outdoor network	–	–	–	•	–
	• Connectivity rules	–	•	++	–	–
	• Network to City features	–	•	–	•	–
	• Network itself	+	+	++	++	+
	• Network to Network	–	++	–	–	–
Visualization	• Mapped	•	++	–	+	•
	• Iconic	•	–	++	–	•
	• LOD	–	–	–	•	–
	• 2D utilities	+	+	++	+	+
	• 3D utilities	–	++	–	+	–
	• 2D city features	++	–	++	+	–
	• 3D city features	–	–	+	+	–
Spatial Scope	• City	+	–	++	+	++
	• Block	+	–	++	+	++
	• Building	–	++	•	•	–
– = no support, • = basic support, + = sophisticated support, ++ = comprehensive support						

The IFC data model is a sophisticated generic data model that provides a semantic categorization of network objects based on their role in the network and a mechanism to derive specialized classes from specific network objects. Also, the IFC data model provides a detailed 3D representation of network objects, and in addition, a graph representing logical and physical connectivity relationships can be extracted to generate a directed graph for flow analysis.

The CityGML UtilityNetworkADE integrates characteristics of different models. The UtilityNetworkADE is the only model that manages the interdependencies between network objects and city entities. In addition, the data model allows the analysis of the network systems at different scales i.e. city, block or building and at different levels of detail. In addition, the data model provides the basic classes to differentiate network objects based on their role in the network. The UtilityNetworkADE provide a basic mechanism to control how network objects can connect to each other. However, this mechanism still needs further development to force different rules to control how network objects can connect to each other or other network types.

5 Conclusion

This paper discusses utility network issues and applications that can be found among city settings. Current data models available for planning and management of utility networks in GIS consider 2D, such as 2D maps and 2D analysis functionalities, and model specific network types. Therefore, this paper introduces use cases for utilities in the built environment, which demonstrate the need of 3D, i.e. 3D virtual reality, 3D analysis functionalities and 3D spatial relationships, and which also require the integration of different utility network types.

Several use case scenarios, “maintenance operation”, “emergency response”, “inspection operation”, “storm drainage network” and “energy planning and simulation” were described. The use cases are used to define the requirements in order to evaluate selected network data models. These requirements can be summarized as follows: there is a need for a holistic modeling framework that supports the management of utility networks at different scales e.g. city or building. Also, there is a need to link networks to other network types, to interior networks, and to city entities. Such a model will help us determining the effects of maintenance operations undertaken on specific utilities on other networks types, as well as to work out the cause of waste materials discharged from networks into the natural resources. It will also allow us to investigate both, the result of damage to the city structure and to another utility network, and the effect of different maintenance operations in different locations within a city on utility service systems. Moreover, it will allow planning the network in conjunction with other city entities, as a change in land use can affect the energy demand on specific utility network systems. The model will help us to see the whole picture of the network, use it for planning the future demand and maintenance and for enhancing management.

The CityGML Utility Network ADE represents a suitable data model for modeling heterogeneous networks in the context of 3D city models. The model is matured and considers most of the requirements defined by the uses cases. However, the model still needs further development which includes:

- Connectivity rules: to constrain the type of network features that may be connected to one another and the number of features of any particular type that can be connected to features of another type. Also this will allow managing of how different network types can connect to each other. This will provide utility managers with the ability to selectively validate features in the network and generate reports as to which features in the network are invalid, i.e. are violating one of the connectivity or other rules. Currently, the CityGML UtilityNetworkADE provides basic support of this concept which still needs further development.
- Application-specific extensions: The current UtilityNetworkADE data model defines base feature types and properties which are applicable to all areas of application. Specific use cases such as district heating simulations or electricity grid simulations, however, might require more specific properties which are not part of the general UtilityNetworkADE data model. To fully support these use cases as well, suitable extensions need to be developed for the UtilityNetworkADE.
- Reference network objects: there is a need to couple indoor navigation models with utility network data models. Successful routing in case of utilities is of great importance for facility management – to respond in timely manner or to reduce the cost of the associated maintenance operation. External networks, which follow the street network path, can be easily referenced to these one-dimensional objects. However, the built environment, is complicated and needs a method to reference objects to this complex structure. This is important for repair and maintenance operations and emergency response in a timely manner.
- 3D models are a valuable contribution, but it is also apparent that they are not needed for all types of applications. Utility network planning on small scale is easier to undertake and visualize with 2D maps. However, 3D is of importance for complex irregular structures and in cases when vertical information needs to be considered (height), for routing above and below objects, without touching a surface. There is a need to provide a mechanism to aggregate networks to one node in 2D. In addition, an investigation is required to support modeling of networks in different levels of detail.
- Connection to the CityGML EnergyADE and to sensors: The CityGML UtilityNetworkADE needs to be further extended to support sensors and time-variant data. Moreover, a further investigation is required on how to link the UtilityNetworkADE with the EnergyADE.

Acknowledgements

The authors would like to thank the participants of the 1st Joint SIG 3D and OCG Workshop on the CityGML UtilityNetworkADE (UTILITYNETWORKADEWIKI 2017) for their valuable contributions to the use cases presented in this paper.

6 References

- BECKER, T., BARTELS, M., HAHNE, M., HEMPEL, L. & LIEB, R. 2012: Cascading Effects and Interorganisational Crisis Management of Critical Infrastructure Operators. Proceedings of the 8th International Conference on Geo-Information for Disaster Management, Enschede, 105-116.
- BECKER, T., NAGEL, C. & KOLBE, T.H., 2011: Integrated 3D Modeling of Multi-utility Networks and Their Interdependencies for Critical Infrastructure Analysis. Advances in 3D Geo-Information Sciences, Lecture Notes in Geoinformation and Cartography, KOLBE, T.H., KÖNIG, G. & NAGEL, C. (eds.), Springer, Berlin Heidelberg, 1-20.
- BECKER, T., NAGEL, C. & KOLBE, T.H., 2012: Semantic 3D Modeling of Multi-Utility Networks in Cities for Analysis and 3D Visualization. Progress and New Trends in 3D Geoinformation Sciences, Lecture Notes in Geoinformation and Cartography, Pouliot, J., Daniel, S., Hubert, F. & Zamyadi, A. (eds.), Springer, Berlin Heidelberg, 41-62.
- 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.
- BOCCALETTI, S., LATORA, V., MORENO, Y., CHAVEZ, M. & HWANG D.-U., 2006: Complex networks: Structure and dynamics. Physical Reports **424**(4-5), 175-308.
- CHATURVEDI, K. & KOLBE, T. H., 2015: Dynamizers – Modeling and implementing dynamic properties for semantic 3D city models. 3rd Eurographics Workshop on Urban Data Modelling and Visualisation.
- CHATURVEDI, K. & KOLBE, T. H., 2016: Integrating Dynamic Data and Sensors with Semantic 3D City Models in the context of Smart Cities. ISPRS-Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, **IV-2/W1**, 31-38.
- CITYGMLUMLMODEL, 2016: ISO-compliant definition of the CityGML 2.0 UML model using Enterprise Architect. https://github.com/opengeospatial/CityGML-3.0/blob/master/WP%2001%20Resources/CityGML_2.x_new_2016_03_09.eap, last access on 19/01/2017.
- CITYGMLWIKI, 2017a: NetworkComponents UML model. <http://www.citygmlwiki.org/>, last access on 19/01/2017.
- CITYGMLWIKI, 2017b: NetworkProperties UML model. <http://www.citygmlwiki.org/>, last access on 19/01/2017.
- ELFOULY, M., KUTZNER, T. & KOLBE, T.H., 2015: General Indicator Modeling for Decision Support based on 3D City and Landscape Models using Model Driven Engineering. Peer Reviewed Proceedings of Digital Landscape Architecture at Anhalt University of Applied Sciences, Wichmann, Berlin Offenbach, 256-267.
- ESRI, 2017a: What are geometric networks? <http://desktop.arcgis.com/en/arcmap/latest/manage-data/geometric-networks/what-are-geometric-networks-.htm>, last access on 19/01/2017.
- ESRI, 2017b: What is Schematics? <http://desktop.arcgis.com/de/arcmap/latest/extensions/schematics/what-is-schematics-.htm>, last access on 19/01/2017.
- GOODRICH, M.T., TAMASSIA, R. & GOLDWASSER, M.H., 2014: Data Structures and Algorithms in Java. Wiley Publishing.

- 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.
- HALFAWY, M. R., 2010: Municipal information models and federated software architecture for implementing integrated infrastructure management environments. *Automation in Construction* **19**(4), 433-446.
- HIJAZI, I. H., EHLERS, M. & ZLATANOVA, S., 2012: NIBU: A new approach to representing and analysing interior utility networks within 3D geo-information systems. *International Journal of Digital Earth*, **5**(1), 22-42.
- ISO 16739:2013 Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries.
- JRC, 2013a: D2.8.III.6 INSPIRE Data Specification on Utility and Governmental Services – Technical Guidelines. European Commission Joint Research Centre.
- JRC, 2013b: D2.10.1: INSPIRE Data Specifications – Base Models – Generic Network Model, Version 1.0rc3. European Commission Joint Research Centre.
- KADEN, R., ELFOULY, M. & KOLBE, T. H., 2015: The CityGML Energy ADE – An International Standardization Effort for the Extension of 3D City Models to Support Energetic Building Analysis. *Innovations for Energy Systems, Mobility, Buildings and Materials*, 5th Colloquium of the Munich School of Engineering.
- KUTZNER, T. & KOLBE, T. H., 2016: Extending Semantic 3D City Models by Supply and Disposal Networks for Analysing the Urban Supply Situation. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, Band **25**, KERSTEN, T. P. (Hrsg.), 382-394.
- LÖWNER, M.-O., BENNER, J. & GRÖGER, G., 2014: Aktuelle Trends in der Entwicklung von CityGML3.0. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, Band **23**, SEYFERT, E., GÜLCH, E., HEIPKE, C., SCHIEWE, J. & SESTER, M. (Hrsg.), Beitrag 168.
- NEWMAN, M.E.J., 2003: The structure and function of complex networks. *SIAM Review* **45**(2), 167-256.
- OGC, 2016: PipelineML. <http://www.pipelineml.org/>, last access on 19/01/2017.
- SEDRIS, 2006: SEDRIS standards. <http://standards.sedris.org/>, last access on 19/01/2017.
- UTILITYNETWORKADEWIKI 2017: Workshop Munich 2016. http://en.wiki.utilitynetworks.sig3d.org/index.php/Workshop_Munich_2016, last access on 19/01/2017.

Filtern bewegter Objekte aus mobilen LiDAR-Daten auf Grundlage volumetrischer Modelle

JOACHIM GEHRUNG^{1,2}, MARCUS HEBEL¹, MICHAEL ARENS¹ & UWE STILLA²

Zusammenfassung: Die automatische Änderungsdetektion in mobil durchgeführten Lasermessungen hat das Potential, ein wichtiges Werkzeug für die Aktualisierung von Stadtmodellen zu werden. Eine Voraussetzung ist das Entfernen von mobilen Objekten, welche zusammen mit der Bausubstanz in den Messdaten enthalten ist. Im Rahmen dieser Arbeit wird ein Ansatz vorgestellt, der sich eine Volumenrepräsentation zunutze macht, um auf Basis von Einzelmessungen bewegte mobile Objekte zu erkennen und vom statischen Hintergrund zu separieren. Die Methode wurde anhand zweier Datensätze einer qualitativen Untersuchung unterzogen. Die Resultate zeigen, dass eine hohe Anzahl der Oberflächenpunkte aller bewegten Objekte sowie des Hintergrunds korrekt erfasst wird.

1 Einleitung

Heutige Mobile-Mapping-Systeme verfügen neben Kameras in der Regel auch über aktive Lasersensoren bzw. Laserscanner zur direkten 3D-Erfassung der Umgebung. Mit derartigen Systemen durchgeführte Messungen verfügen über eine hohe Genauigkeit und 3D-Punktdichte und eignen sich daher hervorragend für die Änderungsdetektion urbaner Strukturen, wie sie beispielsweise für die automatische Aktualisierung von Stadtmodellen benötigt wird. Die mit mobilen Laserscannern (MLS) in urbanen Gebieten durchgeführten Aufnahmen enthalten neben der zu erfassenden Bausubstanz jedoch auch eine Reihe von mobilen Objekten. Hierbei handelt es sich beispielsweise um Fußgänger, Fahrräder, aber auch abgestellte PKWs, die alle nicht für die Stadtmmodellierung von Interesse sind und daher separiert werden müssen. Ein erster Schritt in diese Richtung ist die Entfernung bewegter Objekte, die eine vergleichsweise einfach zu detektierende Sonderklasse der mobilen Objekte darstellen. Gerade diese haben beim Einsatz von Laserscannern einen großen negativen Einfluss auf die Messergebnisse, da sie zu Messartefakten und Mehrfachbeobachtungen führen, die sich über die Zeit hinweg akkumulieren und so einen Schweif entlang der Bewegungstrajektorie des Objekts bilden (vergleiche „rot“ in Abb. 1). In belebten Innenstädten ist dieser Sachverhalt aufgrund von Fußgängern und sich bewegendem Fahrzeuge besonders ausgeprägt und stellt einen großen Störfaktor für die automatische Änderungsdetektion bzgl. der Bausubstanz dar.

Aktuelle Ansätze zur Identifikation mobiler Objekte unterscheiden entweder sehr grob zwischen diesen und dem Hintergrund oder aber arbeiten mit einem kurzen Zeithorizont, welcher zu einer fehlerhaften Segmentierung führen kann, sobald ein Objekt für eine längere Zeitdauer seinen Standort hält. Im Rahmen dieses Papers wird ein Ansatz vorgestellt, welcher eine volumenbasierte

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 Photogrammetrie und Fernerkundung, Technische Universität München, Arcisstr. 21, 80333 München, E-Mail: stilla@tum.de

Hintergrundrepräsentation der Umgebung dazu nutzt, die in deren Erstellung eingegangenen Einzelmessungen möglichst präzise in statischen Hintergrund und bewegte Objekte zu segmentieren. Der hierbei berücksichtigte Zeithorizont umfasst die komplette Messung, sodass auch über längere Zeitabschnitte hinweg stehende Objekte nach wie vor als bewegte Objekte erkannt werden.

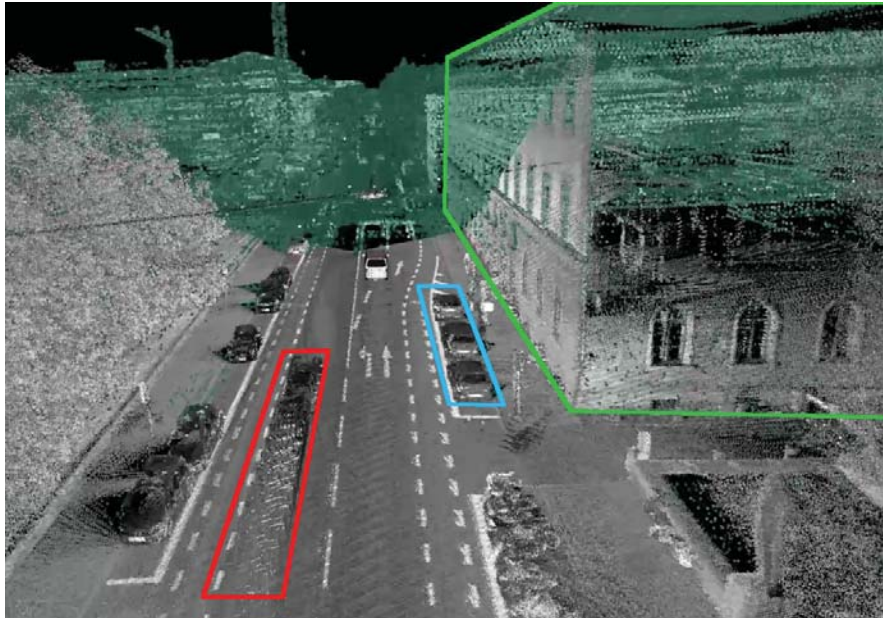


Abb. 1: Eine Beispielszene mit statischen Bausubstanzen (grün), statischen mobilen Objekten (blau) und bewegten mobilen Objekten (rot). Die Mehrfachbeobachtung (rot) ist durch das Akkumulieren mehrerer Einzelmessungen entstanden.

2 Stand der Forschung

2.1 Aufbereitung von Punktwolken

Die Unterscheidung zwischen dynamischen und statischen Strukturen ist in vergleichbarer Form in verschiedenen Bereichen der Forschung relevant. Viele vom Sinn her thematisch ähnliche Arbeiten sind im maschinellen Sehen angesiedelt, konkret im Bereich der Objektverfolgung auf zweidimensionalen Bildern. Im Hinblick auf 3D-Daten sind Parallelen zur Detektion und Verfolgung von mobilen Objekten (engl.: Detection and Tracking of Mobile Objects, DATMO) zu erkennen. Hierbei handelt es sich um einen Themenbereich, der vor allem in der Robotik und dem autonomen Fahren untersucht wird.

Im Kontext von letzterem identifizieren (VALLET et al. 2015) mit einem auf der Dempster-Shafer-Evidenztheorie basierten Ansatz mobile Objekte durch die schritthaltende Suche von Konflikten über mehrere Einzelmessungen hinweg. Die Anzahl der für den Vergleich gewählten Epochen ist dabei relativ gering, sodass nur ein relativ kurzer Zeithorizont abgedeckt wird. Da im Rahmen dieser Arbeit in erster Linie langfristig statische Strukturen gesucht werden, ist das Verfahren daher für die angestrebte Anwendung nur bedingt geeignet.

AZIM & AYCARD (2012) erzeugen eine Repräsentation der Umgebung auf Basis eines octree-basierten 3D-Belegungsgitters. Anhand dieser werden Einzelmessungen im Hinblick auf Konflikte

überprüft. Anhand dieses Vorgehens wird sukzessiv eine Liste von Voxeln aktualisiert, welche in weiterführenden Verarbeitungsschritten zu Objekthypothesen zusammengefasst werden. Das Verfahren ist laut den Autoren anfällig für Rauschen, welches durch ein Schwellwertverfahren auf Ebene der Objekthypothesen oder in späteren Schritten zu beheben ist. Da sich auch dieses Verfahren schnell an Veränderungen in der Umgebung anpasst, ist es ebenfalls nur unzureichend zur Extraktion langfristig stabiler Strukturen geeignet.

Bei der simultanen Lokalisierung und Umgebungskartierung (engl.: Simultaneous Localization and Mapping, SLAM) kann das Entfernen bewegter Objekte dazu beitragen, die Konsistenz der Umgebungserfassung zu erhöhen. LITOMISKY & BHANU (2013) entfernen bewegte Objekte aus Punktwolken, indem Cluster in Punktwolken zweier Epochen gebildet und Korrespondenzen dazwischen gesucht werden. Anhand dieser Korrespondenzen werden bewegte Objekte ermittelt und entfernt. Hierbei bleibt ein hoher Prozentsatz der statischen Objekte vorhanden, bei weitem jedoch nicht alle. Da das Verfahren primär für SLAM konzipiert ist, wird mehr Wert auf Konsistenz als auf exakte Repräsentation der Umgebung gelegt.

Alle hier erläuterten Ansätze haben zum Ziel, entweder eine konsistente Repräsentation unter Billigung von Informationsverlusten zu generieren oder aber sich schnell auf Änderungen in der statischen Umgebung einstellen zu können. Beide Kriterien sind im Hinblick auf die möglichst exakte Änderungsdetektion in urbanen Umgebungen nicht zielführend, da hier der Fokus auf langfristig stabilen Strukturen liegt, die möglichst exakt und vollständig zu erfassen sind.

2.2 Volumenmodelle

Zur Repräsentation des Hintergrunds ist eine punktuelle Oberflächenbeschreibung, wie sie durch Laserscannerdaten gegeben ist, unzureichend. Stattdessen wird ein auf Belegungsgittern basierender Ansatz herangezogen, da dieser zur Beschreibung ganzer Raumbereiche in der Lage ist. MORAVEC & ELFES (1985) erzeugen anhand von ultraschall-basierten Entfernungsmessungen ein probabilistisches Belegungsgitter, welches dem Zustand nach *freie* und *belegte* Raumbereiche in der Umgebung auf Höhe des Sensors beschreibt. Obwohl das kegelförmige Schallsignal eine eher unpräzise Form der Abstandsmessung darstellt, sind die auf diese Art und Weise erzeugten Umgebungsrepräsentationen überraschend exakt.

Eine geometrische Modellierungstechnik von beliebigen 3D-Objekten wird von MAEGHER (1982) vorgeschlagen. Diese umfasst eine auf einem Octree basierende Datenstruktur, welche binäre Belegungsinformationen enthält. PAYEUR et al. (1997) ergänzen diesen Ansatz um probabilistische Informationen. Um eine schnellere Anpassung an eine sich ändernde Umgebung zu gewährleisten, führen HORNING et al. (2013) eine als „Clamping“ bezeichnete Begrenzung der Wahrscheinlichkeiten ein, welche eine Kompression der Datenstruktur durch das Zusammenfassen von Kinds-knoten erlaubt.

Auf dieser Basis wurde in einer vorangegangenen Arbeit ein Modellierungskonzept vorgestellt, welches die globale Verwaltung von probabilistischen Volumenrepräsentationen ermöglicht und ein Konzept zu deren effizienter Abspeicherung vorstellt (GEHRUNG et al. 2016).

3 Filtern bewegter Objekte

Bei der Auswertung von LiDAR-Daten eines mobilen Sensors im Kontext der automatischen Änderungsdetektion stellen bewegte Objekte unerwünschte Störungen dar, die zu einer fehlerhaft diagnostizierten Änderung führen können. Die im Rahmen dieser Veröffentlichung vorgestellte Methode macht sich die in (GEHRUNG et al. 2016) vorgestellte volumetrische Umgebungsrepräsentation zunutze, um eine Hypothese des nicht-beweglichen Hintergrunds aus Sensoreinzelmessungen zu erstellen. Letztere liegen in Form von Entfernungsmessungen vor. Diese Hintergrundrepräsentation wird in einem weiteren Verarbeitungsschritt dazu eingesetzt, die darin eingeflossenen Einzelmessungen in bewegte Objekte und statische Strukturen zu segmentieren. Erstere umfassen alles, was sich zum Zeitpunkt der Messung zumindest eine Zeit lang in Bewegung befunden hat. Eine Annahme hinsichtlich Struktur oder Klasse bewegter Objekte wird nicht getroffen. Die statischen Strukturen umfassen in erster Linie urbane Strukturen wie beispielsweise Gebäude, Schilder und Straßenlaternen sowie in eingeschränktem Maß den Pflanzenbestand. Ebenfalls betroffen sind jedoch auch sonst mobile Objekte wie stillstehende Personen und abgestellte Fahrzeuge, die sich jedoch zum Zeitpunkt der Messung nicht in Bewegung befinden.

3.1 Volumenbasierte Repräsentation

Zum besseren Verständnis des hier vorgestellten Ansatzes wird kurz die in (GEHRUNG et al. 2016) verwendete volumenbasierte Repräsentation von Umgebungen umrissen (Abb. 2). Der zu repräsentierende Raum wird in würfelförmige Kacheln zerlegt, welche je einen Octree enthalten. In diesem sind Voxel in unterschiedlichen Auflösungsstufen abgelegt, welche den Grad der Belegtheit des davon abgedeckten Raumbereichs beschreiben. Dies erfolgt durch eine im weiteren Verlauf als Belegungswahrscheinlichkeit bezeichnete Fließkommazahl, welche das Resultat eines Zwei-Klassen-Entscheidungsproblems darstellt. Es handelt sich hier um die Division der Wahrscheinlichkeit für eine Belegung durch diejenige für eine Nicht-Belegung, auf welche der Logarithmus angewandt wurde. Das logarithmierte Verhältnis zweier Wahrscheinlichkeiten wird auch als Log-Odd bezeichnet.

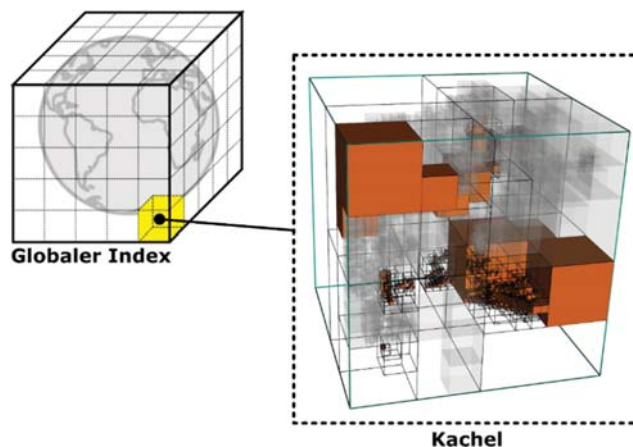


Abb. 2: In einem globalen Referenzgitter werden einzelne Kacheln abgelegt. Jede dieser Kacheln enthält eine probabilistische Umgebungsrepräsentation in Form eines Octree.

Für jede einzelne Entfernungsmessung werden die traversierten Voxel bestimmt und die Belegungswahrscheinlichkeit darin entsprechend verringert. Weiterhin wird das Voxel bestimmt, in welchem der Strahl endet und darin die Belegungswahrscheinlichkeit erhöht. Die Werte für das Absenken und Erhöhen werden im Rahmen des Sensormodells festgelegt. Die Belegungswahrscheinlichkeit ist auf ein festes Intervall begrenzt. Dieser Prozess wird als Klammerung (engl.: Clamping) bezeichnet. Wird die untere Grenze des Intervalls erreicht, so gilt das Voxel als frei; beim Erreichen der oberen Grenze gilt es als belegt.

3.2 Konvergenz der volumetrischen Umgebungsrepräsentation

Dem Ansatz liegt die Annahme zugrunde, dass die volumetrische Umgebungsrepräsentation für längere Beobachtungszeiträume zu einem Punkt hin konvergiert, an dem nur noch die statischen Elemente der Umgebung darin enthalten sind. Dieses erwartete Verhalten lässt sich dadurch begründen, dass ein in Bewegung befindliches Objekt jedes Voxel auf seiner Bahn nur für einen relativ kurzen Augenblick durchläuft. Für jedes der besagten Voxel liegen im Schnitt betrachtet also nur relativ wenige Messungen vor, die für deren Belegung sprechen. Der Großteil der Messungen impliziert hingegen, dass das Voxel Freiraum umfasst.

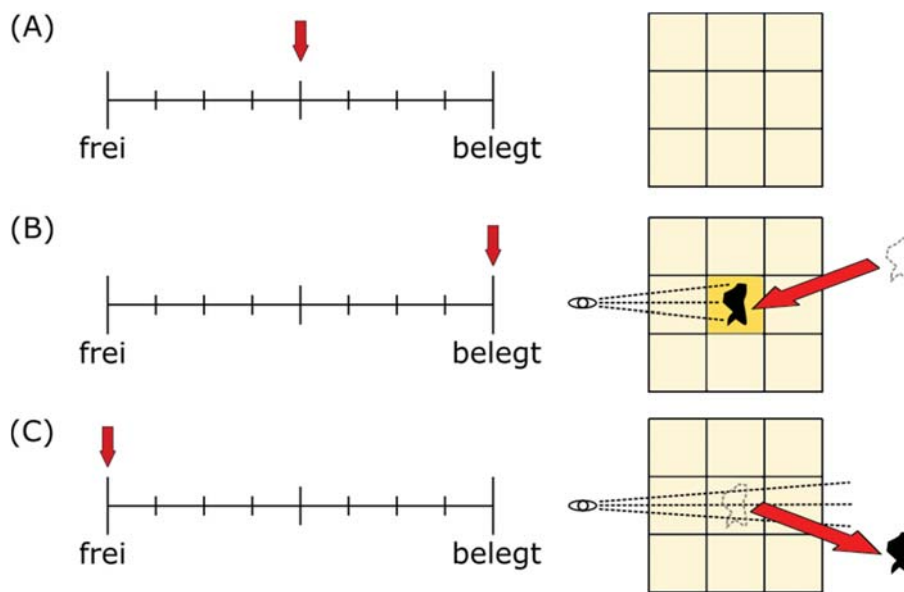


Abb. 3: (A) Ein Voxel wird mit Unwissen instanziiert. (B) Ein Objekt bewegt sich in das Voxel, die Belegungswahrscheinlichkeit wird erhöht, bis die Obergrenze des Intervalls erreicht ist. (C) Das Objekt hat das Voxel verlassen, weitere Messungen senken die die Belegungswahrscheinlichkeit sukzessive ab.

In Bezug auf das im letzten Abschnitt beschriebene Repräsentationskonzept findet folgender Prozess statt (vergleiche auch Abb. 3). Einige Messungen implizieren die Belegung eines Voxels, worauf der entsprechende Wert des Sensormodells auf die Belegungswahrscheinlichkeit aufaddiert wird. Hierbei kann es vorkommen, dass die obere Intervallgrenze erreicht wird und das Voxel als belegt angesehen werden kann. Hat das belegte Objekt besagtes Voxel wieder verlassen, so sorgen

alle durchlaufenden Messstrahlen dafür, dass dessen Belegungswahrscheinlichkeit um den im Sensormodell definierten Wert dekrementiert wird. Traversieren ausreichend viele Messungen das Voxel, so wird der untere Schwellwert erreicht, woraufhin das Voxel wieder als frei betrachtet wird. Die Annahme ist unter einer Reihe von Randbedingungen gültig. So muss eine ausreichend lange Messung für den von einem Objekt durchlaufenen Raumbereich vorhanden sein. Diese hängt von mehreren Faktoren ab; so etwa von der Anzahl der Einzelmessungen und der Parametrisierung des Sensormodells. Auch die Bewegungsgeschwindigkeit des Objekts spielt eine Rolle. Ist das Objekt sehr langsam unterwegs, so kann eine zuverlässige Erkennung bewegter Objekte aus zeitlicher Sicht letzten Teils der Entfernungsmessungen nur erfolgen, wenn es den Messbereich des Laserscanners verlassen hat. Auch kann es problematisch sein, wenn sich sehr viele Objekte sehr dicht zueinander bewegen, da bestimmte Voxel immer wieder von Objekten verlassen und kurz darauf durch andere Objekte betreten werden.

3.3 Filtern der Einzelmessungen

Zunächst wird aus den Einzelmessungen eine volumetrische Hintergrundrepräsentation erstellt. Erstere liegen in Form von Punktwolken mit Messwegen von Laserpulsen vor, die hierfür wie oben beschrieben als Sichtstrahlen interpretiert werden. Unter Voraussetzung der Korrektheit der obigen Annahme konvergiert die volumetrische Repräsentation zu einem stabilen Zustand, welcher die Eigenschaft der Umgebung hinsichtlich des Grades der Belegtheit zutreffend wiedergibt. Kurzzeitige Störungen durch bewegte Objekte werden durch Akkumulation der Messungen beseitigt. Das Resultat stellt eine Hypothese über den statischen Hintergrund der gemessenen Szene dar.

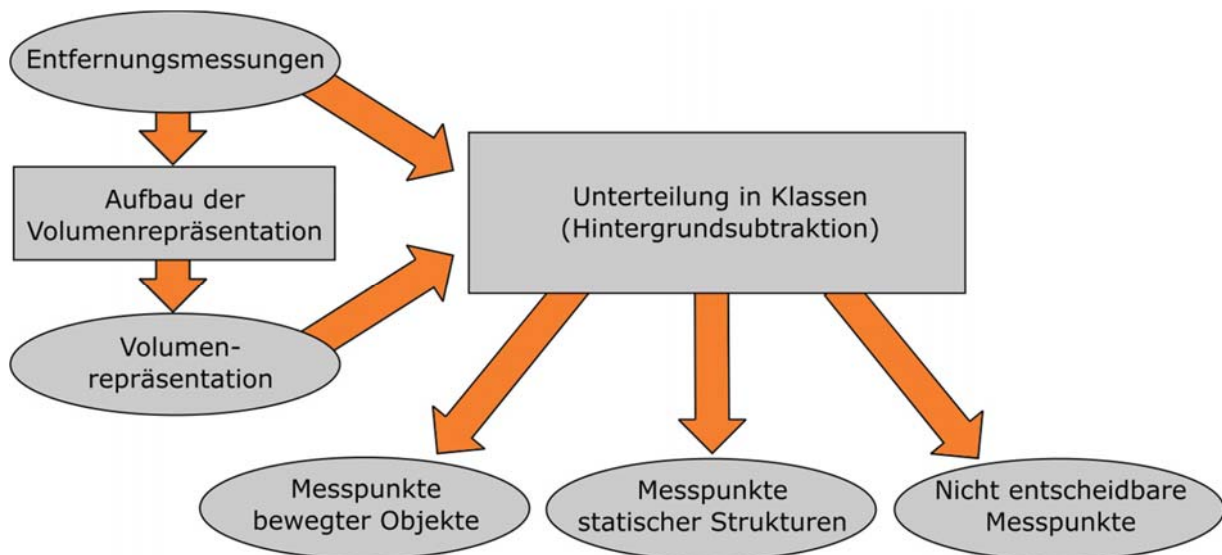


Abb. 4 Ablaufdiagramm der Hintergrundsubtraktion. Aus den Entfernungsmessungen wird zunächst die Volumenrepräsentation hergeleitet. Anhand dieser werden die Einzelmessungen in verschiedene Klassen aufgeteilt.

In einem weiteren Schritt werden die ursprünglich in die Erstellung der Volumenrepräsentation eingeflossenen Einzelmessungen in die Klassen bewegte Objekte, Hintergrund sowie in eine Restklasse für nicht entscheidbare Messungen unterteilt. Ein Ablaufdiagramm ist Abb. 4 zu entnehmen. Der Prozess ähnelt im weitesten Sinn der im Bereich des Maschinensehens angewandten Hintergrundsubtraktion. Für jede Einzelmessung, d.h. jeden Punkt der Punktwolke, wird das Voxel bestimmt, in dem es liegt. Anhand der darin abgelegten Belegungswahrscheinlichkeit wird die Zugehörigkeit des Messpunktes zu einer der oben genannten Klassen bestimmt. Liegt die Belegungswahrscheinlichkeit am oberen Rand des Intervalls, so ist der Messpunkt Teil einer statischen Struktur. Befindet sich die Belegungswahrscheinlichkeit hingegen am unteren Ende, d.h. das Voxel wird gegenwärtig als frei betrachtet, so handelt es sich um einen Messpunkt, der von einem bewegten Objekt stammt. Messpunkte, deren Belegungswahrscheinlichkeit an keiner der Intervallgrenzen liegt, werden der Restklasse zugeordnet.

4 Experimente

4.1 Mobile-Mapping-System

Der vorgestellte Ansatz wurde mit Messdaten mobiler Laserscanner evaluiert, die auf dem Stammgelände der TU München und entlang öffentlicher Gebäude im Umkreis aufgenommen wurden. Zur Erfassung wurde das Mobile-Mapping-System MODISSA des Fraunhofer IOSB eingesetzt (siehe Abb. 5). Das Fahrzeug verfügt insgesamt über vier Laserscanner. Zwei Velodyne HDL-64E sind in einem Winkel von 35° auf dem vorderen Dach angebracht. Diese Konstellation erlaubt es, sowohl einen Teil der Straße vor dem Fahrzeug als auch die Fassaden der umliegenden Gebäude zu erfassen. Zwei weitere Velodyne VLP-16 Puck befinden sich auf dem hinteren Teil des Dachs und sind ebenfalls geneigt angebracht. Acht an den Ecken des Dachs angebrachte Rundumkameras dienen unter anderem zur Texturierung der Lasermessungen. Eine Georeferenzierung der Messdaten erfolgt anhand der Navigationsdaten, die ein Applanix POS LV Navigationssystem mit zwei GNSS-Antennen, einer inertialen Messeinheit und einem Raddrehgeber liefert. Die Navigationsdaten wurden nachprozessiert, um die Genauigkeit der Georeferenzierung zu erhöhen.



Abb. 5: Das Messfahrzeug MODISSA, mit dem die im Rahmen dieser Arbeit verwendeten Datensätze aufgenommen wurden.

4.2 Qualitative Bewertung

In diesem Abschnitt wird eine qualitative Untersuchung der vorgestellten Methode beschrieben. Die Lasermessungen werden wie im vorherigen Abschnitt dargestellt durch Hintergrundsubtraktion in verschiedene Klassen unterteilt. Das Resultat wird anhand folgender Kriterien durch einen Betrachter beurteilt:

- Sind für jedes bewegte Objekt zumindest einige Messpunkte als zu dieser Klasse zugehörend erfasst?
- Zu welchem Prozentsatz sind die Oberflächenpunkte eines bewegten Objektes erfasst?
- Wie verhält sich der Ansatz bei feinen beweglichen Strukturen wie z.B. Vegetation?
- Wie hoch ist die Fehlerkennungsrate; sprich die Anzahl der Messpunkte, die als zu bewegten Objekten gehörend gezählt werden?
- Wie hoch ist der Anteil derjenigen Messpunkte, für die keine Aussage über ihre Zugehörigkeit zu einer der beiden Klassen *bewegt* oder *statisch* getroffen werden kann?
- In welchem Ausmaß wurde der statische Hintergrund erfasst?
- Treten sonstige Artefakte auf?

Die Beantwortung der oben genannten Fragen erfolgt anhand eines Datensatzes, der mit dem Mobile-Mapping-System MODISSA im Bereich des Stammgeländes der TU München erfasst wurde. Es handelt sich um einen belebten Innenhof auf dem Campus. Die Messung hat eine Dauer von knapp 4 Minuten und umfasst mehrere Personengruppen sowie sitzende und umhergehende Einzelpersonen. Ebenfalls in den Aufnahmen enthalten sind mehrere Bäume, die zum Zeitpunkt der Aufnahme über ein ausgeprägtes Blattwerk verfügten. Abb. 6 zeigt einen exemplarischen Ausschnitt aus dem Datensatz.

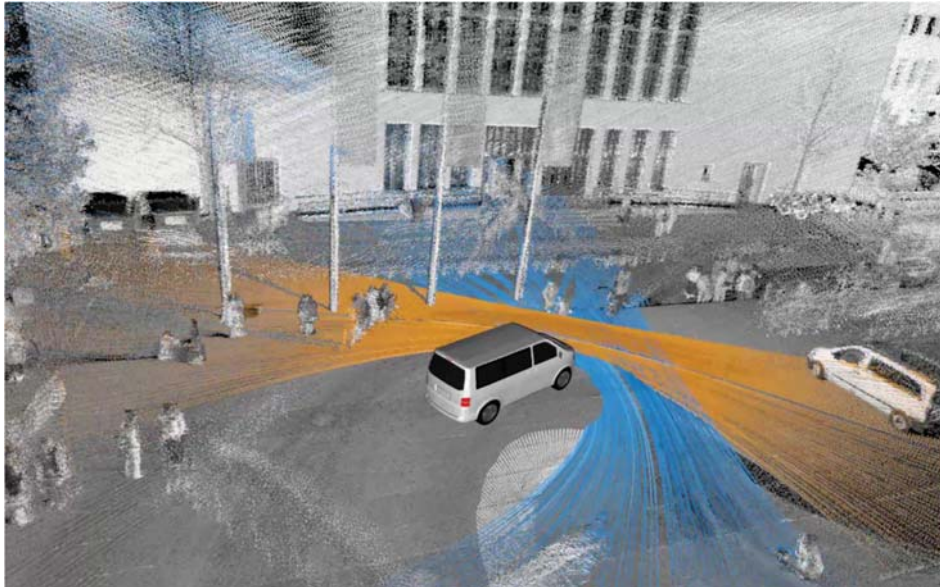


Abb. 6: Akkumulierte Lasermessungen aus dem Innenhof-Datensatz. Die zum Zeitpunkt der Aufnahme aktuellen Messungen der vorderen Laserscanner sind farblich hervorgehoben.

Neben der qualitativen Betrachtung wird auch die Laufzeit des Ansatzes für unterschiedliche Auflösungsstufen untersucht. Hierbei wird die Hintergrundsubtraktion an beiden Datensätzen durchgeführt und die mittlere Dauer pro Umdrehung des Laserscanners bestimmt, welche ca. 100.000 Einzelmessungen umfasst.

5 Diskussion der Ergebnisse

Die hier vorgestellten Ergebnisse basieren auf einer Volumenrepräsentation mit 40 cm Auflösung. Die maximale Länge der berücksichtigten Messstrahlen wurde auf 30 m festgelegt, um die Effekte von Ungenauigkeiten in der Geolokalisierung zu minimieren. Um die Anzahl der Messpunkte ohne Zuordnung zu reduzieren und das Gesamtergebnis zu verbessern wurde das Intervall der Belegungswahrscheinlichkeit für die Einordnung in bewegte Objekte und Freiraum auf $[-0.5, 0.5]$ gesetzt.

5.1 Grad der Erfassung

Alle bewegten Objekte wurden im Rahmen einer manuellen Betrachtung identifiziert und einer genauen Betrachtung unterzogen. Hierzu wurden die Ergebnisse der Segmentierung übereinandergelegt und farblich so kodiert, dass sich anhand der Farbe die Zugehörigkeit zur entsprechenden Klasse erkennen lässt. Alle so aufbereiteten Einzelmessungen wurden durch einen menschlichen Betrachter ausgewertet. Zusammenfassend lässt sich festhalten, dass auf allen bewegten Objekten Oberflächenpunkte detektiert wurden.

Der Prozentsatz korrekt zugeordneter Messpunkte pro bewegtem Objekt liegt schätzungsweise bei 90-95%. Gelegentlich kommt es zu Fehlzuordnungen, wenn eine Person ein belegtes Voxel oder ein Voxel, das nicht zugeordnet werden kann, durchläuft. In einem solchen Fall ist für eine kurze

Zeitspanne von ca. 1-2 Sekunden ein größerer Anteil der Messpunkte falsch zugeordnet. Im vorliegenden Datensatz wurde dies allerdings nur an sehr wenigen Stellen beobachtet. Häufiger fehlen Messpunkte in Grenzbereichen zu anderen Flächen, wie etwa der Bodenebene. Dies ist auf Diskretisierungsfehler zurückzuführen. In der vorliegenden Konfiguration der Volumenrepräsentation liegt die Größe eines Voxel bei $40 \times 40 \times 40 \text{ cm}^3$. Befindet sich die Bodenebene nun mittig in einem Voxel, so fallen die Füße einer Person in ein als solide angesehenes Volumen und werden daher falsch klassifiziert. Beispiele für eine farblich kodierte Einzelmessung sowie die hier beschriebenen Beobachtungen sind Abb. 7 zu entnehmen.

5.2 Kurzzeitig stationäre Objekte

An mehreren Stellen im Datensatz wurde beobachtet, wie Personen kurzfristig an einem Ort verweilen, um dann ihren Weg fortzusetzen. Die Person ist also stationär, ohne dass in den Lasermessungen eine größere Bewegung erkennbar ist. Die zugehörigen Messdaten wurden in besagten Fällen ohne zeitliche Unterbrechung als zu einem bewegten Objekt gehörend eingestuft. Dieser Sachverhalt ist darauf zurückzuführen, dass die Messdaten im Kontext der volumenbasierten Hintergrundrepräsentation betrachtet werden. Der Raumbereich, in den die Messungen fallen, wurde größtenteils als frei beobachtet und war nur kurz durch besagte Personen belegt.

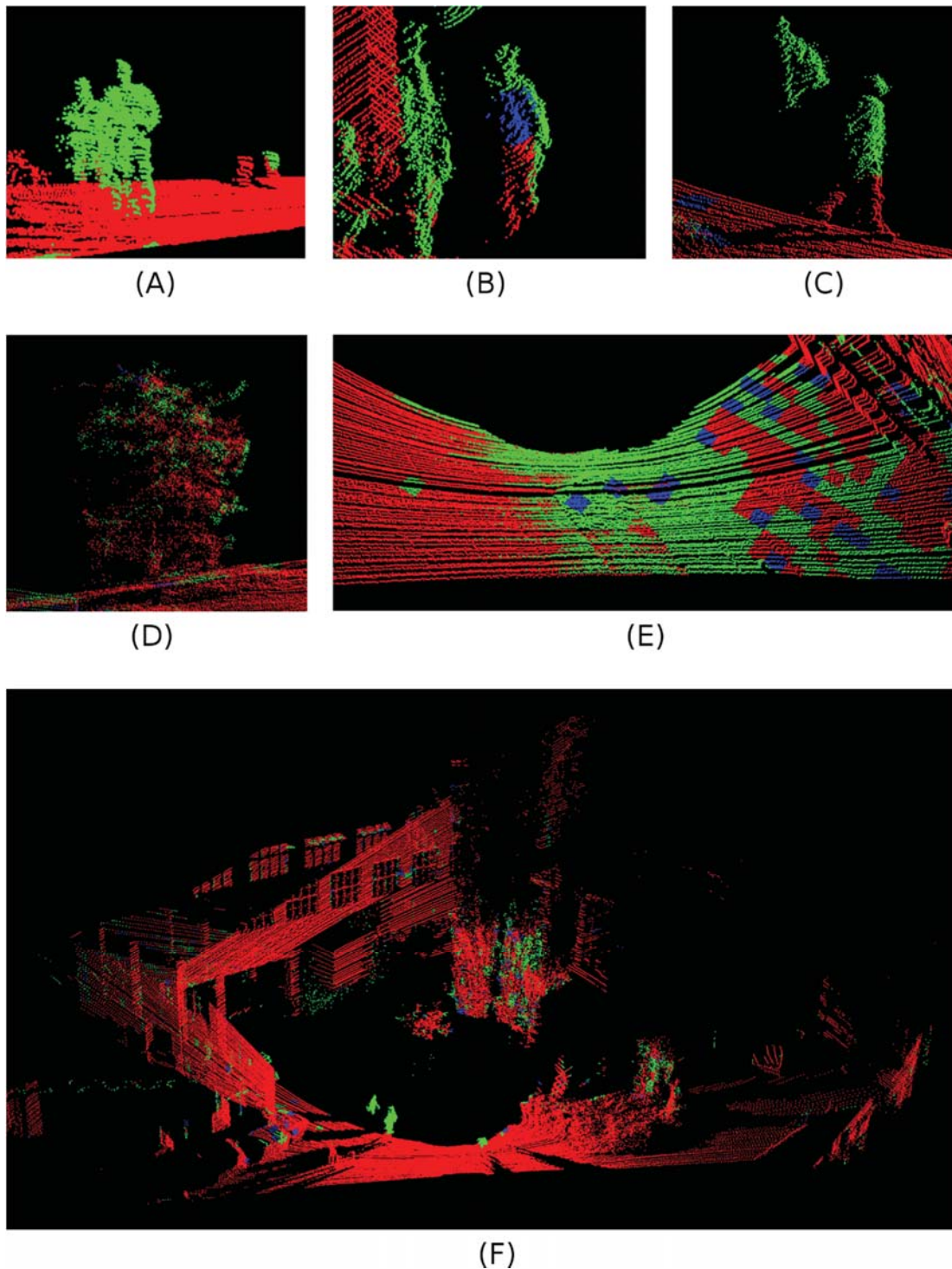


Abb. 7: Die Messpunkte sind in die Kategorien *bewegtes Objekt* (grün), *statische Struktur* (rot) und *keine Zuordnung möglich* (blau) unterteilt. (A) Segmentierung einer bewegten Person. (B) Segmentierungsfehler aufgrund falscher Zuordnung eines Voxels. (C) Diskretisierungsfehler. (D) Vegetation. (E) Systematischer Fehler, voraussichtlich ausgelöst durch flach auftreffende Messstrahlen. (F) Segmentierter Einzelscan eines Velodyne-Laserscanners.

5.3 Fehlererkennung bewegter Objekte

Dieser Abschnitt behandelt mögliche Gründe für die Fehlererkennung bewegter Objekte und der dazugehörigen Objektpunkte beim beschriebenen Verfahren. Wird die Auflösung des Volumenmodells zu hoch gewählt, so kommt es aufgrund von Mess- und Registrierungsungenauigkeiten zu einer hohen Fehldetektionsrate. Eine Auflösung von 40 cm hat sich beim betrachteten Messsystem als guter Kompromiss zwischen Rauschunterdrückung und Anfälligkeit bzgl. Segmentierungsfehlern erwiesen.

Neben dem eher zufällig verteilten Hintergrundrauschen wurden eine Reihe von Artefakten beobachtet (vergleiche auch Abb. 7, D bis E). So werden teilweise durchlässige Strukturen wie Vegetation oder gitterartige Objekte zum Teil als bewegt erkannt, da hier die Messstrahlen gelegentlich das Objekt durchdringen. Ebenfalls beobachtet wurden mehrere als bewegt eingestufte Flecken an statischen Strukturen, vorwiegend an waagrechten oder senkrechten Ebenen. Deren Kantenlänge variiert von ca. einem halben Meter bis hin zu mehreren Metern. Als Ursache hierfür werden Messstrahlen angenommen, die in einem sehr flachen Winkel auf die Ebene treffen. Im Zuge dessen werden andere zu der Ebene gehörende Voxel traversiert, die als belegt eingestuft sind. Messung für Messung wird so der Grad der Belegtheit umgekehrt, bis die traversierten Voxel als frei angesehen werden. Eine grafische Veranschaulichung des hier beschriebenen Effekts ist Abb. 8 zu entnehmen.

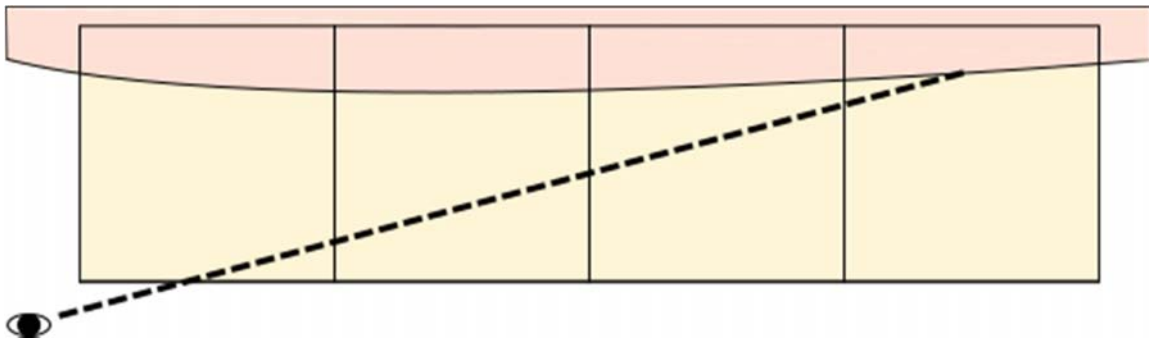


Abb. 8: Flach auf eine Oberfläche (rot) auftreffende Messstrahlen führen dazu, dass eigentlich als statisch angesehene Voxel sukzessive als frei angesehen werden.

5.4 Erfassung des Hintergrunds

Das Resultat der Hintergrundsubtraktion ist eine Menge von Lasermessungen, die sich in Bezug auf die Dichte mit einer unbearbeiteten Messung vergleichen lässt. Vegetation ist ebenfalls in Teilen noch enthalten, auch wenn diese wie oben beschrieben zu einem Großteil den bewegten Objekten zugeordnet wird. Personen sind keine mehr in den Messdaten vorhanden, lediglich deren Schatten. Gelegentlich kann es vorkommen, dass die Füße einer oder mehrerer Personen sichtbar bleiben.

5.5 Messpunkte ohne eindeutige Zuordnung

Messpunkte lassen sich nicht eindeutig einer der beiden oben diskutierten Klassen zuordnen, wenn die Belegungswahrscheinlichkeit des Voxels, in welches ein Messpunkt fällt, an keiner der beiden Intervallgrenzen liegt. Wie oben beschrieben wurden die Intervallgrenzen für die Segmentierung angepasst, sodass auch nicht eindeutige Messungen einer der beiden Klassen zugeordnet werden. Das Resultat ist, dass die Anzahl der Messpunkte ohne Zuordnung bei unter 5% liegt. Es handelt sich meist um einzelne Messpunkte oder aber um Flecken in Form und Größe eines Voxels.

5.6 Laufzeitbetrachtung

Die Laufzeitmessung wurde auf einem Rechner mit 32 GiB an RAM und einem Intel i7 mit 3.5 GHz und 12 Kernen durchgeführt. Die Implementierung wurde mittels OpenMP parallelisiert. Nach Initialisierung des Programms werden schrittweise die benötigten Elemente der volumetrischen Hintergrundrepräsentation geladen. Sobald diese komplett geladen ist beträgt die Laufzeit für jedes Bündel aus ca. 100.000 Einzelmessungen ca. 450 ms. Experimente für mehrere Auflösungsstufen von grob nach fein haben ergeben, dass die Tiefe im Octree keinen nennenswerten Einfluss auf die Laufzeit hat.

6 Fazit & Ausblick

Das im Rahmen dieser Arbeit vorgestellte Verfahren ermöglicht es, einzelne Lasermessungen in die Klassen *bewegtes Objekt* und *statischer Hintergrund* zu unterteilen. Hierfür wird ein einfacher Ansatz ohne Modellannahmen auf Basis einer Volumenrepräsentation verwendet, der zuerst aus Lasermessungen aufgebaut und dann zu deren Unterteilung in die beiden Klassen eingesetzt wird. Die Methode wurde anhand eines durch ein Messfahrzeug aufgenommenen Datensatzes qualitativ untersucht. Für jedes bewegte Objekt wird eine gewisse Mindestanzahl von Messpunkten richtig zugeordnet. Der prozentuale Anteil an korrekt zugeordneten Messpunkten pro Objekt liegt bei ca. 90-95%, wobei es an Übergängen zu statischen Strukturen wie etwa der Bodenebene zu Diskretisierungsfehlern kommen kann. Weiterhin treten Artefakte an Vegetation, teildurchsichtigen Strukturen sowie aufgrund von flach auftreffenden Messstrahlen auf. Zum statischen Hintergrund gehörende Messungen werden abgesehen von wenigen Diskretisierungsfehlern korrekt zugeordnet. Der Anteil der Messpunkte, die nicht zugeordnet werden können, ist mit weniger als 5% relativ gering. In weiterführenden Arbeiten sind Ansätze zu untersuchen, um das durch flach auftreffende Messstrahlen ausgelöste Artefakt sowie durch Vegetation bedingte Fehlzuordnungen zu beheben. Weiterhin ist es denkbar, die den bewegten Objekten zugeordneten Messpunkte zu gruppieren und damit einen Klassifikator zu trainieren, um auch unbewegte Instanzen der Objektklassen automatisiert aus den Lasermessungen entfernen zu können.

7 Literaturverzeichnis

- 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. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **XLII-2/W1**, 45-51.
- HORNUNG, A., WURM, K. M., BENNEWITH, M., STACHNISS, C. & BURGARD, W., 2013: OctoMap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees. Autonomous Robots **34**(3), 189-206.
- LITOMISKY, K. & BHANU, B., 2013: Removing Moving Objects from Point Cloud Scenes. Advances in Depth Image Analysis and Applications, Springer Berlin Heidelberg, 189-206.
- MEAGHER, D., 1982: Geometric modeling using octree encoding. Computer Graphics and Image Processing **19**, 129-147.
- MORAVEC, H. & ELFES, A., 1985: High Resolution Maps from Wide Angle Sonar. Proceedings of the 1985 IEEE International Conference on Robotics and Automation, 116-121.
- PAYEUR, P., HEBERT, P.; LAURENDEAU, D. & GOSSELIN, C. M., 1997: Probabilistic octree modeling of a 3D dynamic environment. Proceedings of the IEEE International Conference on Robotics and Automation, 1289-1296.
- VALLETT, B., XIAO, W. & BRÉDIF, M., 2015: Extracting Mobile Objects in Images Using a Velodyne LIDAR Point Cloud. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, **II-3/W4**, 247-253.

Dynamically Extending Spatial Databases to support CityGML Application Domain Extensions using Graph Transformations

ZHIHANG YAO¹ & THOMAS H. KOLBE¹

Abstract: As a domain extendable 3D city model the OGC standard CityGML has been increasingly employed as a dominant consensus over different application areas. An essential challenge encountered in many research and commercial activities in the field of 3D city modelling is to find a reliable approach for achieving high-efficient storage and management of data models according to CityGML with its Application Domain Extensions (ADE) in order to ensure interoperable data access across broad application domains. Based on graph transformation, this paper introduces a new approach along with an extensive database structure that allows for dynamically extending the spatially-enhanced relational databases for handling arbitrary CityGML ADEs by means of graph transformation systems. With this approach, relational database schemas with simplified and optimized table structures can be automatically generated from the XML application schemas of CityGML ADEs by performing a series of user-defined graph transformation rules which can describe complex mapping rules for transforming object-oriented data models to relational database models in a fully declarative way. The proposed approach has been successfully implemented and tested based on a number of different CityGML ADEs like Energy ADE, Dynamizer ADE, and UtilityNetwork ADE.

1 Introduction and motivation

Over the recent years, the OGC standard CityGML has been increasingly used as a semantic 3D city model for describing the relevant entities in the landscape and urban space. It is designed as an application schema based on the Geography Markup Language 3 (GML3) and includes a wide range of feature classes along with their interrelationships as well as their relevant thematic and spatial properties which can cover the most varying needs of different application domains. However, in many practical applications, an essential issue often encountered while using CityGML is that many additional feature classes or extra attribute types are required to represent some certain objects or data structures for performing domain specific analysis or simulations such as energy demand calculations, utility network analysis, facility management, noise propagation simulations etc. For this reason, CityGML provides an extension mechanism called “Application Domain Extension (ADE)” allowing third parties to dynamically and systematically extend the existing CityGML data models by attaching extra application schemas which can contain additional feature classes and attribute properties defined for specific application domains. Since 3D city model objects usually have well-defined identifiers which are usually kept stable throughout the lifetime of the referenced real-world objects, the complete 3D city models being attached with diverse domain-specific information are hence considered as a good basis for

¹ Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstraße 21, D-80333 München, E-Mail: [zhihang.yao, thomas.kolbe]@tum.de

building an integrative platform for realizing interoperable data exchange across different application domains.

In order to efficiently store and manage CityGML-based 3D city models regarding their complex data structures and spatial characteristics, spatially-enhanced relational database management systems (SRDBMS) were often chosen to serve as the central data repositories as they are usually featured with extensive spatial capabilities compared to other types of database systems. A representative relational database solution is the 3D City Database (3DCityDB), an Open Source software for the storage, management, and analysis of 3D city models according to the CityGML standards. It is shipped with a compact relational database schema resulted from a careful mapping of the object-oriented data model of CityGML 2.0 to a fixed database structure which has been employed as standard database implementation in many production environments and commercial projects all over the world. However, the 3DCityDB database schema has no support for CityGML ADEs and the underlying design decisions strongly rely on many manual steps, for example recognizing a certain complex model structure and mapping it to a particular target database structure, which make the automation of the mapping process for CityGML ADEs much harder than the generic solutions proposed in the many literatures. Furthermore, since CityGML ADE models can be arbitrarily defined with very complex model structures in practical situations, a challenging task is to find a way for dynamically and automatically extending the 3DCityDB database schema in order to efficiently store and maintain geospatial data of arbitrary CityGML ADEs (cf. Fig. 1).

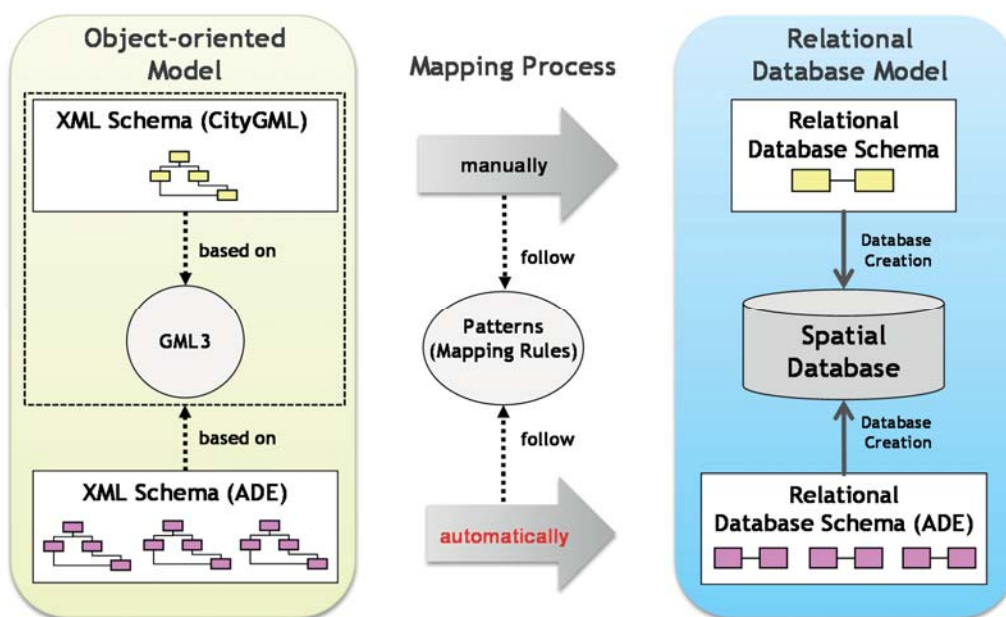


Fig. 1: Conceptual workflow for creating a dynamically extendable 3D geo-database for storage and management of CityGML ADEs using spatially enhanced relational database management systems

Within this paper, we will present a new graph-based framework which uses typed and attributed graphs for semantically representing the object-oriented data models of CityGML ADEs and utilizes graph transformation systems to automatically generate compact relational database

schemas extending the 3DCityDB. The transformation process is performed by applying a series of fine-grained graph transformation rules which declaratively describe user-definable mapping rules for transforming complex object-oriented data models to relational database schemas. With this approach, the 3DCityDB database schema has been structurally enhanced to be dynamically extendable for arbitrary CityGML ADEs and can, hence, be used as an integrative information backbone for interoperable data access for a wide range of domain-specific information according to the CityGML standard.

The rest of this paper is structured as follows: Section 2 addresses the relevant concepts for the modelling and development of CityGML ADEs and provides the theoretical foundation for our work. Section 3 introduces an extended database structure for 3DCityDB which becomes to be dynamically extendable for handling multiple CityGML ADEs simultaneously. In section 4 we propose a conceptual approach for schema transformation using graph transformation systems for realizing the automatic generation of relational database schemas from CityGML ADE application schemas. In addition, a technical implementation of the proposed conceptual approach is detailed in section 5. The last section draws the conclusions about the presented study and outlines the relevant aspects of our future research and development work.

2 Model Driven Approach for the development of CityGML ADEs

A fundamental understanding of the mechanisms behind the development of CityGML ADEs is the major prerequisite for the establishment of the graph-based framework presented in this paper. Basically, the data models in software and systems engineering are nowadays usually developed by following the well-known design approach called “Model Driven Architecture (MDA)” issued by the Object Management Group (OMG). MDA aims to provide a systematic way for facilitating the development of software and data models using Computer-Aided Software Engineering (CASE) tools (GA et al. 2006). The key idea of this design approach is that an abstract platform-independent model (PIM) representing the underlying conceptual models shall be first created at the earlier stage of the development lifecycle and then be implemented for different application platforms at a later stage by transforming it into the corresponding platform-specific models (PSM) such as XML schema and database schema etc. In order to ensure the unambiguity and consistency of the definition and description of model structures, the Unified Modelling Language (UML) was chosen as the standard modelling language for describing the platform-independent models since it comes with a rich set of graphic notations and syntaxes allowing to visually represent complex model structures as well as to fully represent the respective semantic meanings.

In the field of geographic information modelling, the combined use of MDA and UML has been predominantly adopted as the main instrument for the development of geospatial information models on the basis of the so-called “ISO 19100 standards family” issued by the Technical Committee (TC) 211 of the International Organization for Standardization (ISO). This bundle of standards comprises a series of abstract specifications which jointly provide a general guideline for geographic information modelling regarding the definition and description of geographic phenomena in order to ensure the interoperability of data exchanges across different application platforms. For example, the ISO standard 19109 “*Rules for Application Schema*” comprehensively specifies the rules on modelling of real-world features along with their attribute properties as well

as their interrelationships such as generalization/specialization and association relationships like aggregations and compositions (ISO 19109:2005). In particular, the spatial properties of features such as their geometric-topological characteristics can be comprehensively expressed using the geometry and topology models defined in the ISO standard 19107 “*Spatial Schema*” specification (ISO 19107:2003). Other relevant modelling aspects such as temporal schema, metadata, and coverage etc. are explicitly covered in the standards ISO 19108, ISO 19115, and ISO 19123 respectively. Based on these abstract standards, the ISO 19136 standard “*Geography Markup Language (GML)*” was released which provides an open and manufacturer independent framework for the definition of geospatial data models on the application level (ISO 19136:2007). The main GML model components like geometries, topologies, coverages, coordinate- and time reference systems etc. are wholly and partly drawn from the conceptual models defined in the ISO 19100 standard family and implemented as a XML-based schema (GML Schema) according to the XML encoding conventions specified in the ISO 19118 standard (ISO 19118:2011). Moreover, the ISO 19136 standard additionally specifies an extensive UML profile together with a normative UML-to-XML encoding rule set which builds an MDA-compliant framework allowing for the development of a variety of application schemas based on a standardized exchange interface using GML (cf. Fig. 2).

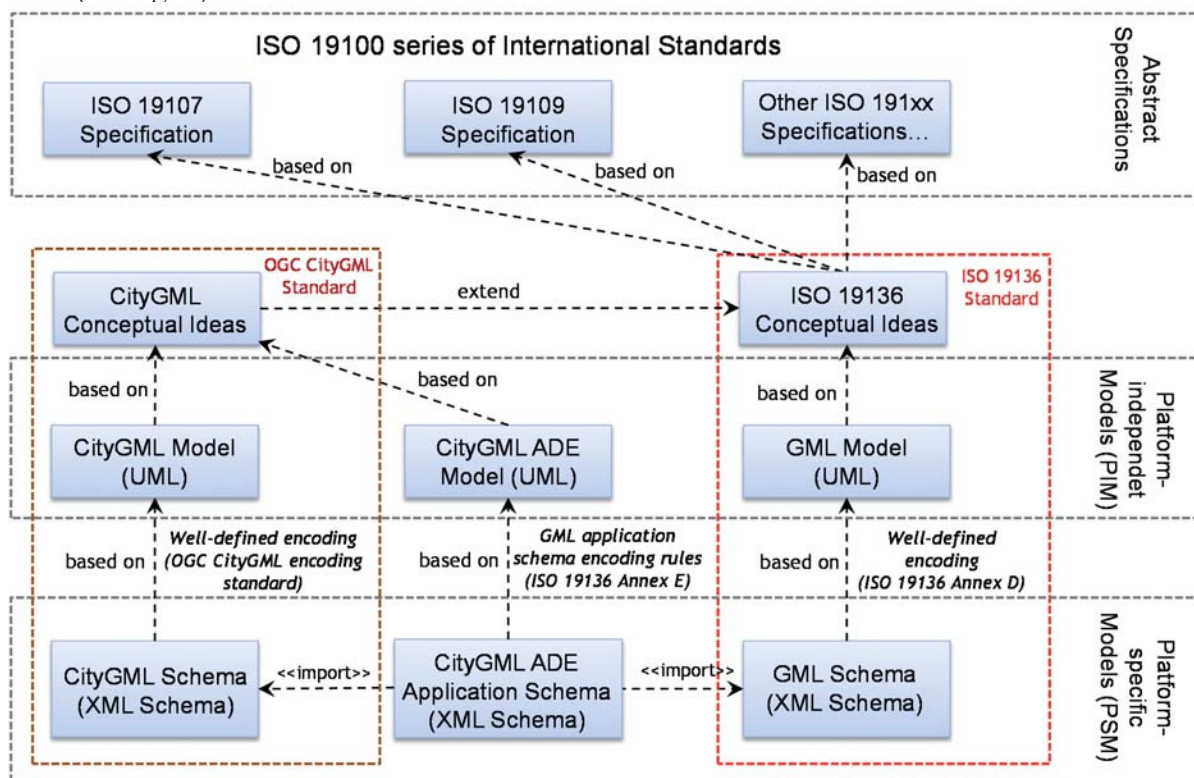


Fig. 2: Relationship between the ISO 19100 standard family and CityGML ADEs

As a GML application schema, CityGML specifies domain specific concepts for 3D city models formally by extending GML’s data models for the spatio-semantic modelling of 3D city and landscape objects. It is a platform-independent geospatial information model that has been realized as a XML-based application schema (CityGML Schema) in compliance with GML schema.

According to the CityGML specification, CityGML is featured with an extension mechanism called “Application Domain Extension (ADE)” allowing to dynamically enrich the CityGML’s predefined data models with extra classes and attribute properties to develop a domain-specific GML application schema. According to literature (VAN DEN BRINK et al. 2013), CityGML ADEs can also be originated and developed from the start with the design of a platform-independent model using UML models by following the MDA approach. An XML-based application schema pointing to the CityGML and GML schemas can later be automatically generated by means of the UML-to-XML encoding rules defined in the ISO 19136 standard (SHAPECHANGE 2016). Although the UML model can be directly serialized to a text-based document using the standard format “XML Metadata Interchange (XMI)” for information exchange, CityGML ADEs developed by third parties are mostly provided by means of XML schema definition (XSD) files instead. This is because, on the one hand, the XML schema provides a sophisticated syntactical structure which is able to fully represent the semantics of object-oriented models with complex data structures and can also be easily interpreted and parsed by many XML-schema-aware software tools and programming libraries. On the other hand, the XML instance documents can also be directly checked against their meta-models using the respective XML schemas to ensure the validity of the corresponding geospatial data.

3 Extendable relational database structure for CityGML ADEs

In 3D GIS applications, typically database management systems (DBMS) are employed for achieving high-efficient data storage and management as well as interoperable data access for large-volume geospatial data. Since the data models of CityGML ADEs may comprise heterogeneous information about the domain-specific features along with their spatial and non-spatial properties, a number of open-source and commercial database products with spatial extensions have been developed and also intensively investigated in research and development projects to ascertain their key functional capabilities and limitations in handling GML-compliant geospatial data. These database products can be roughly categorized into two types according to their native database structure, namely non-relational and relational databases. The former one can be further classified into object-oriented databases, document-oriented databases, and graph databases etc. According to the literature (MAO et al. 2014, AGOUB et al. 2016), such non-relational databases are currently still more or less limited in their capabilities and performance of performing certain kinds of spatial operations and coordinate transformations which are of great importance for the use of CityGML ADEs in practical applications. Thus, due to their extensive abilities for storing, analyzing, and processing spatial data elements, the spatially-enhanced relational database management systems (SRDBMS) such as the commercial Oracle with ‘Spatial’ license and the open-source PostgreSQL with ‘PostGIS’ extension etc. are nowadays predominantly employed in many enterprise applications and services to maintain GML-conformant data.

As a representative relational database solution, the open-source 3D geodatabase “3DCityDB” was developed for CityGML. It can be operated using the Oracle and or PostgreSQL database management system. The database schema of 3DCityDB results from a careful, manual mapping of the object-oriented data model of CityGML onto a compact relational database structure optimized with respect to database complexity, operating performance, and semantic

interoperability. Over the past years, the 3DCityDB has been widely deployed in many commercial production environments to manage virtual 3D models for many cities worldwide like Munich, Berlin, Zurich, Rotterdam, Helsinki, Singapore, and London etc. (KOLBE et al. 2016). However, the current version of the 3DCityDB does not provide a generic solution for handling datasets of CityGML ADEs. This limitation hinders the usage of 3DCityDB to be disseminated for many domain-specific application fields which in turn have a strong need for dynamically extending 3DCityDB for arbitrary CityGML ADEs. In order to achieve this objective, an extended database structure based on the original 3DCityDB database schema was proposed in the course of this research work and the conceptual database design is sketched in the following figure in terms of an informal package diagram.

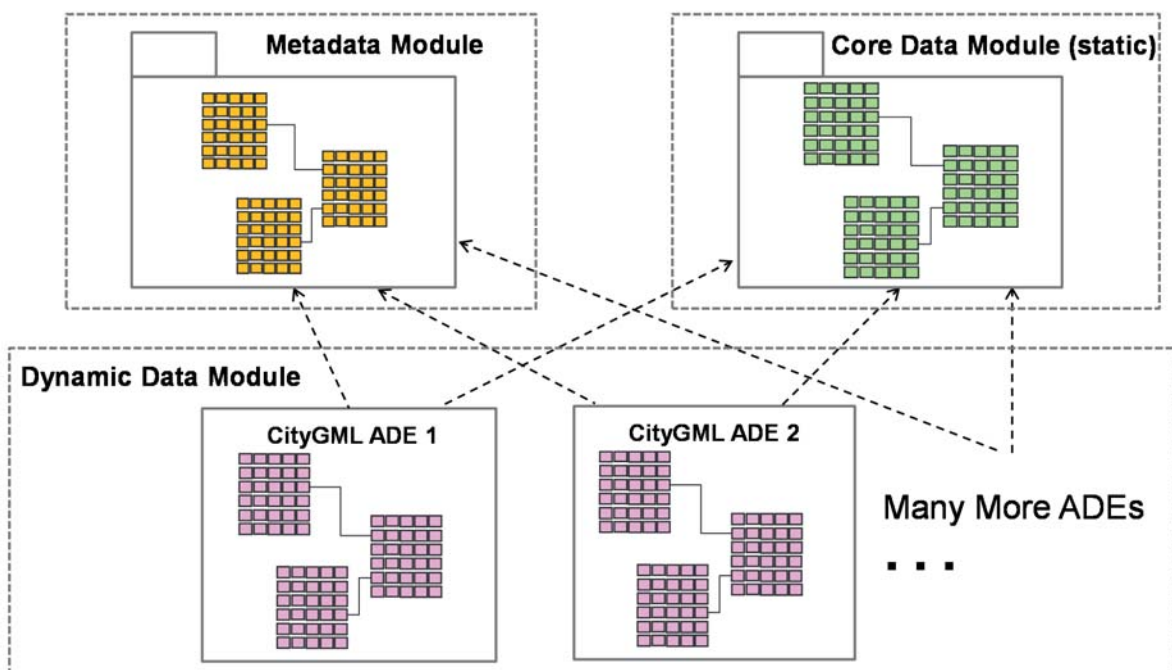


Fig. 3: Dynamically extendable database structure for storing CityGML ADE data in the 3DCityDB

As shown in the figure above, the new 3DCityDB database schema is designed in a modular fashion. It consists of three parts, namely *Metadata Module*, *Core Data Module*, and *Dynamic Data Module*. The green grids enclosed in the *Core Data Module* represents those database tables that are already included in the current version of the 3DCityDB database schema which is responsible for storing the standard CityGML models such as *Building*, *Tunnel*, *Transportation*, *CityFurniture*, *CityObjectGroup*, *Generic*, *Appearance* etc. For a given CityGML ADE, an additional group of database tables forming as a separate module belonging to the *Dynamic Data Module* (pink grids in the figure) shall be created and attached to the 3DCityDB database schema and the relationships (e.g. generalization/specialization and associations) among the model classes of CityGML and CityGML ADEs are adequately reflected using database foreign key constraints which can also ensure the data integrity and consistency within the database system. The *Metadata Module* associated with the *Dynamic Data Module* is used for storing the relevant meta-information (e.g. the XML namespaces, class affiliations etc.) about the application schema of the

registered CityGML ADEs as well as about the database objects (e.g. indexes, columns, foreign key constrains) created in the corresponding *Dynamic Data Module*. In this way, the SQL statements for e.g. dropping the database tables of an individual CityGML ADE can be directly derived from the database schema which hence become to be dynamically manageable for handling multiple CityGML ADEs within the one database instance.

4 Automatic derivation of relational database schemas for CityGML ADEs using a Graph Transformation System

4.1 General concepts

Concerning the fact that domain-specific data models of CityGML ADEs may define arbitrarily complex data structures within their XML schema definition files, one of the key challenges regarding the use of the database structure presented in the previous section is to find a way to automatically derive efficient database schemas. Whereas the XML schema of a CityGML ADE natively represents an object-oriented data structure, the target database schema has a relational table structure which should result from a corresponding model transformation process to perform the mapping of an object-oriented model (input model) to a relational database model (output model). For this, both the input and output models have to be mapped onto some computer-interpretable representations such that the model transformation process can be automatically carried out by applying a set of predefined mapping rules within a computer-aided transformation engine. A couple of commercial and open-source software systems like Go Loader (SNOWFLAKE 2016) and Degree (DEEGREE 2016) etc. have been developed for this problem which are capable of reading and parsing GML-compliant application schemas and automatically generating the desired relational database schemas for different types of modern database management systems such as Oracle and PostgreSQL/PostGIS.

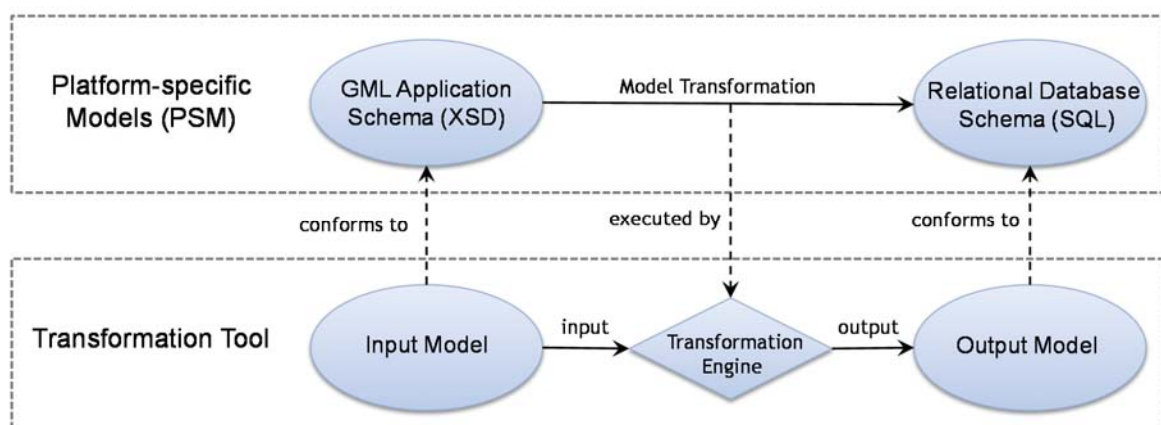


Fig. 4: General approach for deriving relational database schemas based on model transformation

However, these software systems are usually limited in performing such model transformation processes by following a number of simple and rigid mapping rules, i.e. each GML class or complex data type is mapped onto one individual database table, and an association between two GML classes is represented using an associative database table connecting with the database tables

of the respective GML classes. Such kinds of mapping rules can easily result in a large amount of database tables and will, therefore, lead to time-consuming data retrieval processes when, for example, performing a complex query on those data contents that are distributed over many database tables, because a large number of database joins are required which have a significant negative impact on the overall database performance. This issue has been taken into account by the software system “Go Loader” which provides a special optimization strategy that maps multiple associated feature classes onto one database table in order to ensure that the overall number of the generated database tables will not exceed a given maximum number. However, this solution lacks the ability to treat some special cases in a special way and can only rigidly reduce the database complexity without concerning the semantic clarity. For example, in case two non-abstract feature classes have an inheritance relationship and both are attached with different attribute properties, each feature class shall be mapped to an individual database table rather than a shared one, because such separated table representation is not only space-efficient for the database but also has much more clear class affiliation to facilitate the data access to the respective database tables for ETL software tools.

In order to achieve a good trade-off between database complexity and semantic clarity, the design decisions applied during the manual development of the 3DCityDB database schema must be captured and abstracted to a set of mapping rules which will be adopted in the generation process of relational database schemas for CityGML ADEs. The general idea of this design decision is that a group of classes with particular characteristics shall be first simplified into an optimal model structure which can be transformed to a relational database model with a simple mapping step (cf. STADLER et al. 2009). However, such design decision process strongly relies on many manual operations, because the class groups usually have varying characteristics and complex structures which are hardly formalized as programming code implemented in traditional software systems for performing automatic operations. Moreover, the mapping process can become even more complicated due to the possible iterative procedures in case that the simplified class groups need to be again further optimized if they satisfy some further simplification criteria. Concerning the above-mentioned issues, a novel approach is needed to overcome the above-mentioned issues with the help of an appropriate formalism to represent the source and target models as well as their mapping rules.

4.2 Mathematical background and conceptual solution

Since both the object-oriented model and the relational database model intrinsically have a graph structure, the semantic meanings and mapping relationships of both models can be represented using a special kind of graph whose nodes and edges shall be assigned with types and attributes. With such graph, the model transformation can be adequately abstracted as an algebraic graph transformation comprising a series of user-defined graph transformation rules $\{r_1, r_2, \dots, r_n\}$ that can be used to declaratively describe the model transformation rules and be successively applied using a graph transformation system (or graph rewriting system) for typed and attributed graphs. According to the fundamental theory of this system (EHRIG et al. 2004), each graph transformation rule is equivalent to a match morphism in the form of $r: L \rightarrow R$ where L is called left-hand side (LHS) graph, whereas R is called right-hand side (RHS) graph. The LHS graph can be seen as a match pattern which could be algebraically isomorphic to one of those graphs $\{G'_1, G'_2, \dots, G'_n\}$

that are a subset of the given host graph G_S , where $L \cong G'_x$ and $\{G'_1, G'_2, \dots, G'_n\} \subseteq G_S$. Per default, the given transformation rules are ordered randomly by the graph transformation system and each will be selected and checked against the host graph G_S . Alternatively, the processing sequence can also be scheduled by grouping them into a set of layers sorted in descending order according to the user-defined priorities and as such will be processed successively. When a graph transformation rule is being processed and if a match of the LHS has been found in the host graph, this graph transformation rule will be interpreted as being applicable and the matched subgraph will be substituted by the RHS immediately. In the subsequent operation, the modified host graph G_i will in turn be treated as the input for the next transformation step and the entire graph transformation process will hence be carried out within an iterative procedure which will be terminated until no further applicable transformation rule can be found. In the last step, and the latest state of the host graph G_T will be treated as the final result of the graph transformation process (cf. Fig. 5).

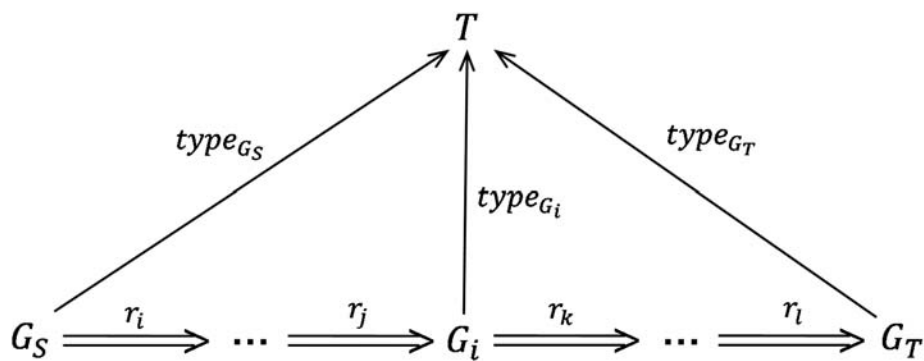


Fig. 5: General process chain in graph transformation systems (cf. TAENTZER et al. 2006)

In addition, a transformation rule can be further constrained by defining additional conditions which are generally categorized into two types: the attribute-based prerequisite condition called *Positive Application Condition* (PAC) and the graph-based prohibitive condition called *Negative Application Condition* (NAC). The PACs are combined with the logical operator “AND” and thus a graph transformation rule will only be triggered when all its PACs are fulfilled. In the contrary, the NACs use the logical operator “OR” and make a transformation rule not to be applicable if one of the NACs is satisfied. Therefore, the proper use of PACs and NACs offers a variety of possibilities for specifying a graph transformation rule being subject to different kinds of application conditions. For instance, in practical applications, there is a frequently applied NAC which is algebraically identical with the RHS of the respective transformation rule and can be used for avoiding the unexpected infinite loop of running the same graph transformation rule. Moreover, the graph transformation system supports the declaration of types for graph objects (e.g. nodes and edges) and also the definition of an abstracted meta-graph (or called *type graph*) T (cf. Fig. 5) which has a similar fashion compared to the UML meta-model in model-driven engineering and can serve as a global constraint on the declared node and edge types by prescribing the structural relationships among them, i.e. the associations and inheritance hierarchies between node types, the multiplicities and role names of the edges etc. Due to its rich semantic structure the meta-graph is well capable of formally describing the meta-models of the object-oriented models according to

the ISO 19100 standard family as well as the underlying conceptual models of the relational database models along with the model mapping structures. In the following figure, an excerpt of our developed meta-graph is depicted in which the yellow nodes represent elements of the object-oriented (source) schema and the green nodes represent elements of the relational (target) database schema. Based on this graph representation, a model-driven framework can be established for the automatic derivation of a relational database schema from a GML application schema using graph transformation systems.

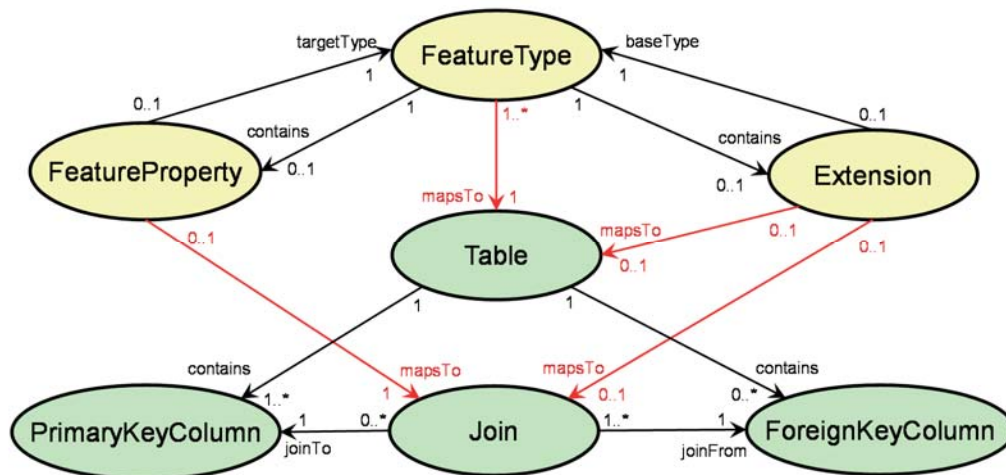


Fig. 6: An excerpt of the meta-graph for representing the conceptual model mapping structure

The yellow nodes together with the associated edges among them constitute a meta-subgraph for representing the structural information (e.g. associations and inheritance relationships) of those classes that are transitively derived from the GML class “FeatureType”. The definition of this subgraph follows the XML encodings defined in the GML specification (ISO 19136) in accordance with the conceptual models of the ISO 19109 General Feature Model (GFM). First, the inheritance relationship is represented by means of the node type *Extension* being connected with the node type *FeatureType* via the edge types *contains* and *baseType* both of which are connected with the node type *FeatureType*. This graph structure is adopted from the syntax of the extension mechanism of XML schema according to which a subclass defined as a XML `<complexType>` shall enclose an `<extension>` XML element having a reference to the respective super class. Second, the association such as aggregation and composition between two feature classes can be explicitly mapped onto a node with the type *FeatureProperty* which links the node type *FeatureType* via the edge types *contains* and *targetType*. This graph representation follows the so-called “feature-property-feature” encoding structure where an aggregate class represented as an XML element can contain a child element acting as a “property” element which shall have a reference to the member class in order to establish the class association.

The green nodes and their associated edges are meant to be used for semantically representing the meta-model of the relational database schema especially regarding the relations between database tables. In general, a database table shall have only one primary key acting as a unique record identifier which can be made up of one or more columns known as “primary key columns”. In order to establish a connection between two database tables, one of them should have a special

column called “foreign key column” pointing to a primary key column in another one by defining a so-called “foreign key constraint”. It is also allowed to define multiple foreign key constraints on different columns to link with the one primary key column within the same table. Based on this conceptual database structure, a corresponding graph representation is designed in the following way: A node type *Table* representing the database table is created and connected with the node types *PrimaryKeyColumn* and *ForeignKeyColumn* to reflect the ownership using the edge type *contains*. The foreign key constraint is explicitly represented using the node type *Join* which links the *ForeignKeyColumn* with the *PrimaryKeyColumn* using the edge type *joinFrom* and *joinTo* respectively.

Concerning the conceptual models of the mapping relationships between object-oriented models and relational database models, the edges with the type *mapsTo* are utilized to link the corresponding entities of both models according to our following mapping principles:

- A feature class shall be mapped onto a database table which also allows to be shared by multiple feature classes if they have inheritance relationships among them.
- In case that two feature classes with inheritance relationship are mapped onto one database table, the inheritance relationship should also be mapped onto this database table. For other cases, an additional foreign key constraint shall be used for the representing the inheritance relationship.
- The association between two feature classes shall be always mapped onto a foreign key constraint.

Above, a variety of graph transformation rules can be flexibly designed by developers in a declarative way. In total, we have defined 11 of these rules to cover all occurring mappings. For the illustrative purpose, we have chosen a representative graph transformation rule realizing a relatively complex model transformation whose underlying idea is shown in the following figure.

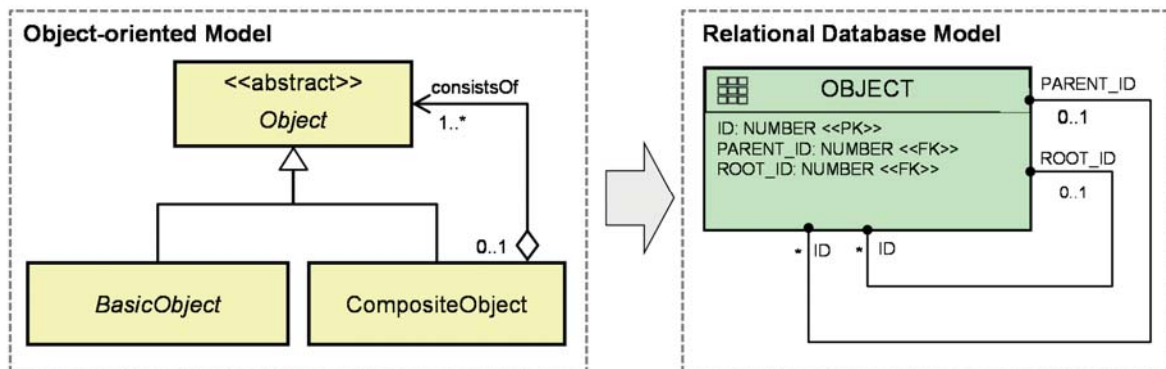


Fig. 7: General idea for the mapping of an object-oriented model with the composite pattern onto an efficient relational database model

The object-oriented data model sketched on the left of the figure builds a recursively composite data structure according to the design pattern “Composite Pattern” (cf. GAMMA et al. 1995). The target database structure on the right side was originally introduced in the 3DCityDB documentation (KOLBE et al. 2016) and has been successfully implemented for the 3DCityDB database schema to perform fast and efficient queries on composite geospatial data stored in relational databases. The key idea of this database design is, on the one hand, to utilize one database

table for the mapping of all the involved feature classes along with their inheritance relationships. On the other hand, a foreign key column PARENT_ID is used for representing the composition relationship. Additionally, this database table receives a foreign key column ROOT_ID which holds the ID of the root element of each composite hierarchy and hence allows for fast retrieving of all its child elements by querying on the attribute ROOT_ID in order to avoid time-costly recursive database joins. More details about the corresponding database implementation can be found in the technical documentation of 3DCityDB. The following figure gives an overview of how the presented mapping concept is formulated as a graph transformation rule with respect to the meta-graph illustrated in figure 6. The LHS graph shown on the left of figure 8 stands for the object-oriented model, whereas the RHS graph receives a copy of the LHS graph in addition with a set of new nodes and edges which represent the target database model and the explicit mapping information between the entities of both models. Moreover, this graph transformation rule can be additionally equipped with a NAC being structurally identical with the RHS in order to ensure that this transformation process will be executed only one time.

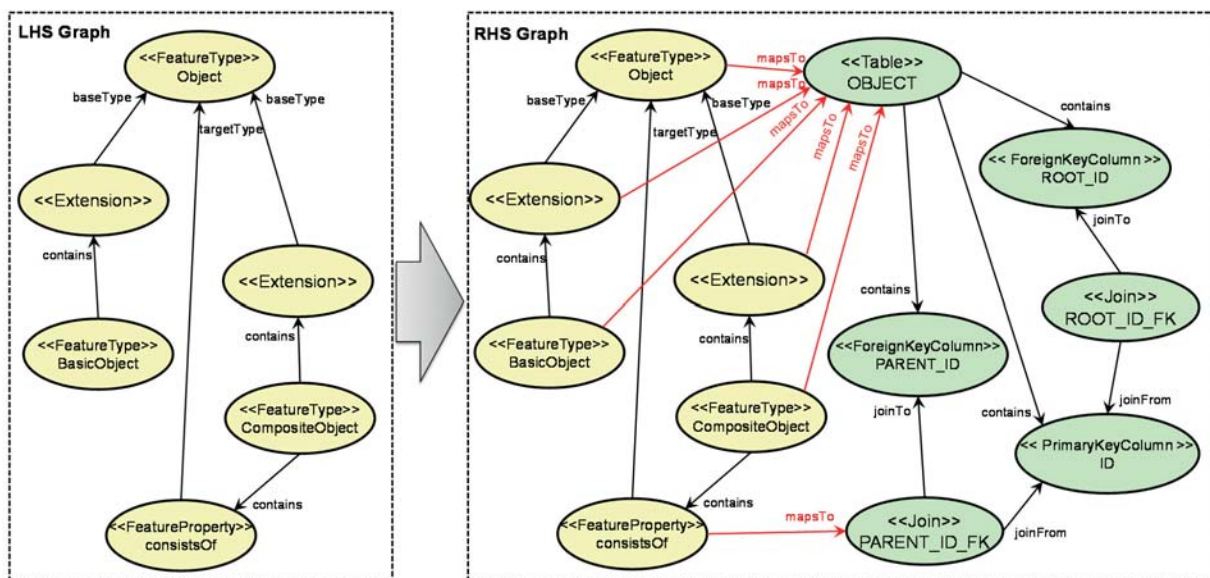


Fig. 8: Graph transformation rule in accordance with the mapping concept illustrated in Fig. 7

4.3 Prototypical implementation

This section introduces a converter tool which was implemented based on the presented conceptual solution presented in the previous sections regarding the automatic derivation of a 3DCityDB compliant relational database schema from the GML application schema of a CityGML ADE. This converter tool is as a stand-alone Java application which employs the AGG graph transformation system (TAENTZER et al. 2003) for performing the dedicated graph transformation process. For test purposes, a number of existing valid CityGML ADEs such as Energy ADE (NOUVEL et al. 2015), UtilityNetwork ADE (KUTZNER & KOLBE 2016), and Dynamizer ADE (CHATURVEDI & KOLBE 2016) have been used for automatically deriving their corresponding relational database schemas. These schemas are given as SQL scripts. When they are executed, new tables according to the CityGML ADEs are being created in the 3DCityDB. The tables extend to the latest version (3.3.0)

of the 3DCityDB database schema and the SQL scripts are online available from a web-based GitHub repository². The following figure gives an overview of the entire process chain of this converter tool which is logically decomposed into three parts and each is explained in detail in the following paragraph.

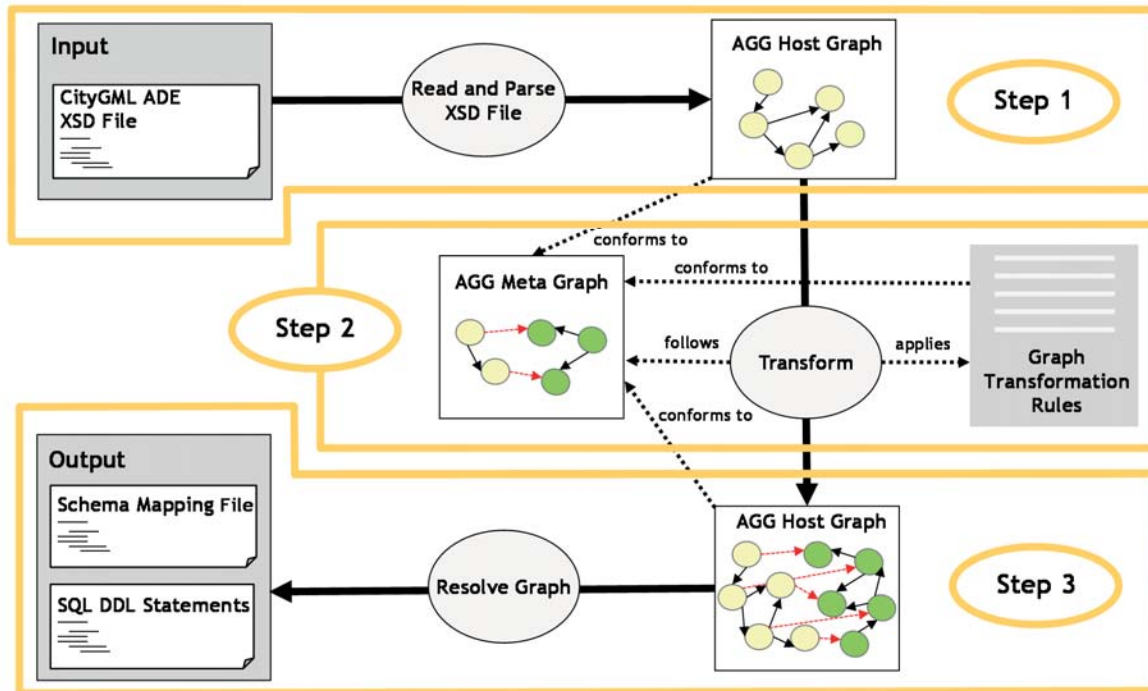


Fig. 9: Model Transformation using Graph Transformation System

In the first processing step, the XML schema definition (XSD) file of the input CityGML ADE is read into the converter tool using the Java library XSOM for reading and retrieving the underlying structural and semantic information from the XML elements. Subsequently, the parsed information is internally mapped onto Java objects forming an AGG-compliant graph representation of the object-oriented data model. In the next step, the created AGG graph will be used as the input graph for the AGG engine in order to apply the predefined graph transformation rules for performing the desired graph transformation process according to the mapping rules employed during the development of the 3DCityDB relational database schema. These graph transformation rules along with their supervisory meta-graph can be easily created by means of the graphical editor program delivered together with the AGG engine which allows for interactively defining and drawing the graph structures and saving them in XML format which can be reused by other AGG engines running on different platforms. After having completed the transformation process, the input graph together with the newly created nodes and edges will form the output graph representing the information of the object-oriented model and the derived relational database model as well as the mapping information between both modes. In the last step, the output graph will be resolved by the converter tool retrieving the target relational database schema which is then written to an SQL

² <https://github.com/yaozhihang/3DCityDB-Extensions-for-CityGML-ADEs>

document for a chosen database system like Oracle or PostgreSQL/PostGIS. Additionally, the mapping information can also be derived from the output graph in the form of an XML-based mapping schema that is used to facilitate the automatic data transformation between CityGML ADE instance documents and database systems via the standardized open interface WFS for web-based data access.

5 Conclusions and outlook

The 3D city models which are using CityGML together with its extension mechanism Application Domain Extension (ADE) are nowadays being broadly used as a common information hub for representing the most relevant urban objects along with their diverse graphical, topological, as well as semantic properties. As cities are highly complex systems in general whose underlying domain-specific information are usually very large in size and often also have complex data structures, the high-performance storage and management of such kinds of data information are usually required for a variety of simulations and analyses in the field of 3D GIS and pose a challenge to many researchers and software developers. The key intention of the presented work is to find and explore a new approach to dynamically extend the spatially-enhanced relational database management systems (SRDBMS) for efficiently handling the geospatial data elements of arbitrary CityGML ADEs.

A dynamically extendable database structure was first proposed which was designed based on the extension of the open source 3DCityDB database schema to provide an integrative database platform for handling multiple CityGML ADEs at the same time. Second, the automatic derivation of relational database schemas from the XML schema definition (XSD) files of CityGML ADEs was expressed as a special kind of model transformation process, in which the logical mapping rules between object-oriented models and relational database models were declaratively described as a set of user-defined graph transformation rules which can be carried out within a graph transformation system. Finally, based on the introduced conceptual solution, a Java-based desktop application has been developed to technically implement the automatic procedures ranging from the parsing and creation of a graph representation of the input CityGML ADE application schema via the execution of the dedicated graph transformation process up to the derivation of the target relational database schema in SQL format together with the respective mapping schema in terms of XML structure.

Future work will focus on the continuing development of the presented converter tool by improving and refining the existing graph transformation rules or adding new ones for automatically generating more compact and efficient relational database schemas. It is also planned to include this converter tool as well as the proposed extendable database structure into a future release of the 3DCityDB to provide the basis for handling arbitrary CityGML ADEs. For the practical use of this database solution, numerous challenging tasks are raised, including e.g. the development of a generic data access tool for interacting with the dynamically created database schema, since the current version of the 3DCityDB import/export tool doesn't support for reading and writing data elements of CityGML ADEs and hence must be conceptually as well as functionally extended. Moreover, an enhanced version of the WebGL-based 3D web client (Yao et al. 2016) being part of the 3DCityDB software package will be developed for the interactive 3D

visualization and exploration of spatial and thematic information of any CityGML ADEs on the web. In summary, our long-term development goal is to make the 3DCityDB software package to be an extensive open-source application platform with full support for the interoperable management, access, analysis, and visualization of domain-extendable 3D city models according to the CityGML standard.

6 Acknowledgments

This work has been carried out in large parts within the research project “*Fachlich erweiterbare 3D-Stadtmodelle – Management, Visualisierung und Interaktion*” funded by the company CADFEM GmbH and its subsidiary company virtualcitySYSTEMS GmbH, whom we would like to thank for their support.

7 References

- AGOUB, A., KUNDE, F. & KADA, M., 2016: Potential of Graph Databases in Representing and Enriching Standardized Geodata. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **25**, T. KERSTEN (Hrsg.), 208-216.
- CHATURVEDI, K. & KOLBE, T. H., 2016: Integrating Dynamic Data and Sensors with Semantic 3D City Models in the context of Smart Cities. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences **IV-2/W1**, 31-38.
- DEEGREE, 2016: deegree webservice 3.3.19 documentation. <https://wiki.snowflakesoftware.com/display/GLD/XML%2C+GML+and+Application+Schemas>, last access on 2016-12-13.
- EHRIG, H., PRANGE, U. & TAENTZER, G., 2004: Fundamental theory for typed attributed graph transformation. International conference on graph transformation. Graph Transformations, Lecture Notes in Computer Science, EHRIG, H., ENGELS, G., PARISI-PRESICCE, F. & ROZENBERG, G. (eds.), Springer, Berlin Heidelberg, 161-177.
- GRÖGER, G., KOLBE, T.H., NAGEL, C. & HÄFELE, K.-H., 2012: OGC City Geography Markup Language (CityGML) Encoding Standard Version 2.0.0. OGC Doc 12-019, Open Geospatial Consortium. Retrieved from <http://www.opengeospatial.org/standards/citygml>
- GA, D., DJURIC, D. & DEVED, V., 2006: Model Driven Architecture and Ontology Development. Springer Science & Business Media.
- GAMMA, E., HELM, R., JOHNSON, R. & VLISSIDES, J., 1995: Design Patterns: Elements of Reusable Object-oriented Software, Addison-Wesley, ISBN: 0-201-63361-2.
- ISO 19107:2003: Geographic Information – Spatial schema (GML) (ISO 19107:2003). International Organization for Standardization
- ISO 19109:2005: Geographic Information – Rules for application schemas (ISO 19109:2005). International Organization for Standardization
- ISO 19136:2007: Geographic Information – Geography Markup Language (ISO 19136:2007). International Organization for Standardization
- ISO 19118:2011: Geographic Information – Encoding (ISO 19118:2011). International Organization for Standardization

- KUTZNER, T. & KOLBE, T. H., 2016: Extending Semantic 3D City Models by Supply and Disposal Networks for Analysing the Urban Supply Situation. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **25**, T. KERSTEN (Hrsg.), 382-394.
- KOLBE, T. H., 2009: Representing and exchanging 3D city models with CityGML. 3D geo-information sciences, Lecture Notes in Geoinformation and Cartography, LEE, J. & ZLATANOVA, S. (eds.), Springer, Berlin Heidelberg, 15-31.
- KOLBE, T. H., YAO, Z., NAGEL, C., REDWEIK, R., WILLKOMM, P., HUDRA, G. & KUNDE, F., 2016: 3D City Database for CityGML Version 3.3.0. Documentation. Retrieved from <http://www.3dcitydb.org/3dcitydb/documentation/>
- MAO, B., HARRIE, L., CAO, J., WU, Z., & SHEN, J., 2014: NoSQL Based 3D City Model Management System. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **40**(4), 169.
- NOUVEL, R., BAHU, J.M., KADEN, R., KAEMPF, J., CIPRIANO, P., LAUSTER, M., HAEFELE, K.H., MUNOZ, E., TOURNAIRE, O. & CASPER, E., 2015: Development of the CityGML Application Domain Extension Energy for Urban Energy Simulation. Proceedings of 14th international conference of the international building performance simulation association (ibpsa) 2015.
- SHAPECHANGE, 2016: XML Schema – ShapeChange. <http://shapechange.net/targets/xsd/>, last access on 2016-12-13.
- STADLER, A., NAGEL, C., KÖNIG, G. & KOLBE, T. H., 2009: Making interoperability persistent: A 3D geo database based on CityGML. 3D geo-information sciences, Lecture Notes in Geoinformation and Cartography, Lee, J. & Zlatanova, S. (eds.), Springer, Berlin Heidelberg, 175-192.
- SNOWFLAKE, 2016: GO Loader Documentation – Snowflake Software Documentation Wiki. <https://wiki.snowflakesoftware.com/display/GLD/Home>, last access on 2016-12-13.
- TAENTZER, G., 2003: AGG: A Tool Environment for Algebraic Graph Transformation. International Workshop on Applications of Graph Transformations with Industrial Relevance, Springer, Berlin Heidelberg, 481-488.
- TAENTZER, G. & GARUGHI, G.T., 2006: A Graph-based Approach to Transform XML Documents. International Conference on Fundamental Approaches to Software Engineering, Springer, Berlin Heidelberg, 48-62.
- VAN DEN BRINK, L., STOTER, J. & ZLATANOVA, S., 2013: UML-Based Approach to Developing a CityGML Application Domain Extension. Transactions in GIS **17**(6), 920-942.
- YAO, Z., CHATURVEDI, C. & KOLBE, T.H., 2016: Browser-basierte Visualisierung großer 3D-Stadtmodelle durch Erweiterung des Cesium Web Globe. Geoinformationssysteme 2016.

Neue Dimensionen der Erdbeobachtung im Schulunterricht: Raumfahrt, 3-D und Augmented Reality

ANNETTE ORTWEIN¹, BENNET KREBS, JANNA MEHRING,
JOHANNES SCHULTZ & ANDREAS RIENOW

Zusammenfassung: ‚Columbus Eye - Live-Bilder von der ISS im Schulunterricht‘ ist ein vom Deutschen Zentrum für Luft- und Raumfahrt (DLR) gefördertes Projekt zur nachhaltigen Integration von Fernerkundung in den Schulunterricht. Als exklusiver europäischer Partner begleitet Columbus Eye das ‚High Definition Earth Viewing‘ (HDEV)-Experiment der NASA, bei dem die Erde rund um die Uhr von vier an der Internationalen Raumstation (ISS) installierten HD-Kameras beobachtet wird. Die hierbei entstehenden Videos werden in Bonn archiviert, auf dem Columbus Eye-Portal (www.columbuseye.uni-bonn.de) gestreamt und dienen als Grundlage für die Entwicklung interaktiver Unterrichtsmaterialien. Die für den gesamten MINT-Bereich konzipierten Materialien verknüpfen analoge mit digitalen und virtuellen Inhalten.

1 Die Mission: Lernen mit der ISS

Die Antwort auf die häufig gestellte Frage, wie Schülerinnen und Schüler (SuS) nachhaltig für Naturwissenschaften begeistert werden können, ist komplex. Zunächst müssen der Forschergeist und die Faszination für Wissenschaft geweckt werden, dann benötigen die SuS aber auch Instrumente und Methoden, mit denen sie sich neue Themenfelder erschließen können. Die Erforschung des Weltraums und die bemannte Raumfahrt - als ein wahrgewordener Menschheitstraum - eignen sich hervorragend dazu, diese Voraussetzungen zu schaffen. Die Möglichkeit, von oben auf unseren Planeten zu blicken, offenbart die Abhängigkeit zwischen den Menschen und ihrer Umgebung und führt so zu einem tieferen Verständnis des gekoppelten Mensch-Umwelt-Systems (VOß et al. 2010). Wie der griechische Philosoph Sokrates bereits in der Antike sagte: „Der Mensch muss sich über die Erde erheben, weil er nur dann die Welt verstehen wird, in der er lebt.“ Durch das Projekt ‘Columbus Eye – Live-Bilder von der ISS im Schulunterricht’ soll Erdbeobachtung nachhaltig als Methode in den Schulunterricht integriert werden. Professioneller Umgang mit der Methodik ist ebenso ein Lernziel wie die Förderung von Entscheidungskompetenz und Verantwortungsbewusstsein.

Faszinierende Bilder von der ISS, leicht verständliche und frei zugängliche Software für Lehrerinnen und Lehrer (LuL) und SuS sowie ein attraktives Online-Lernumfeld werden hier eingesetzt, um Fernerkundung für den Schulunterricht interessant zu machen (RIENOW et al. 2015a).





¹ Universität Bonn, Geographisches Institut, Meckenheimer Allee 166, D-53115 Bonn,
E-Mail: s3anortw@uni-bonn.de

2 Die Hardware: HDEV-Experiment

Im April 2014 wurden am Columbus-Labor der ISS vier kommerzielle off-the shelf (COTS) Kameras installiert und in Betrieb genommen. Die Kameras filmen die Erde rund um die Uhr aus drei unterschiedlichen Perspektiven. Zwei Kameras bedienen die Rückwärts-, eine die Vorwärts- und eine die Nadir-Perspektive. Die räumliche Auflösung in der Nadir-Perspektive beträgt etwa 500 m mit einer spektralen Auflösung von 390 bis 750 nm, die mittels eines CMOS-Sensor aufgenommen wird (RUNCO 2015). Die zeitliche Auflösung variiert zwischen 180 Minuten und drei Tagen, dabei sind Signalstörungsphasen und Nachtflüge eingeschlossen. Im oberen Teil von Tabelle 1 werden die spezifischen Daten der Kameras beschrieben. Auflösung und Winkel sind dabei abhängig von den allgemeinen (und den wechselhaften) Bedingungen der ISS, die sich in rund 400 km Höhe bewegt, jedoch Kursschwankungen und Höhenveränderungen unterliegt. Etwa einmal in 90 Tagen wiederholen sich die Belichtungsverhältnisse (RIENOW ET. AL. 2014).

Die Kameras sind verbunden mit der integrierten Bordelektronik zur Steuerung und Datenverarbeitung, sodass sie von außen durch eine TRek-Arbeitsstation gesteuert und eingeschaltet werden können. Obwohl die Taktung beeinflusst sein kann, bleiben Zoom, Linse und Lichtempfindlichkeit voreingestellt. Die Videos werden unbearbeitet über einen Kursverfolgungs- und Datenrelaissatellitensystem (TDRS) auf die Erde gestreamt (RUNCO 2015). Ursprünglich war das Ziel des High Definition Earth View (HDEV) – Experiments auf der ISS lediglich, die Beständigkeit und Nützlichkeit der CMOS-Kameras unter Weltraumbedingungen für weitere Missionen zu testen. Die hierbei entstandenen hochwertigen Videos von der Erdoberfläche könne jedoch hervorragend didaktisch in Wert gesetzt werden.

Tab. 1: HDEV-Kameras und Anwendungsmöglichkeiten im Schulunterricht

	Hitachi©	Panasonic©	Sony©	Toshiba©
Kamera				
Sicht	Vorwärts	Rückwärts	Rückwärts	Nadir
Beispielthema	Entstehung eines Zyklons	Bildkorrektur		Rayleigh-Streuung
Unterrichtseinheit	Das Auge des Zyklons	Mittelwertberechnung von der ISS		Streuung & Farben in der Atmosphäre
Typ	Augmented Reality	Lernmodul		Observatorium

3 Das Projekt: Columbus Eye

Als exklusiver europäischer Partner der NASA ist Columbus Eye zuständig für die Archivierung, Aufbereitung und Bereitstellung des Videomaterials von der ISS, das so für den Schulunterricht sowie die interessierte Öffentlichkeit zugänglich gemacht wird. Hierbei wird Columbus Eye vom Deutschen Zentrum für Luft- und Raumfahrt (DLR) unterstützt. Augenblicklich enthält das HDEV-Archiv in Bonn 14 Terrabyte Daten; darin enthalten sind alle Videos, die seit dem 23. September 2014 aufgenommen wurden.

3.1 Lernportal rund um die Live-Videos von der ISS

Eine Möglichkeit, die interessierte Öffentlichkeit an der bemannten Raumfahrt zu beteiligen, ist es, sie die Perspektive der Astronauten einnehmen zu lassen. Um dies zu ermöglichen, werden die Videos von der ISS live auf dem Columbus Eye-Portal gestreamt und kostenfrei zur Verfügung gestellt. Zusätzlich finden sich auf dem Portal ein Bereich mit aktuellen Nachrichten sowie Hintergrundinformationen zur ISS und Alexander Gersts Mission „Blue Dot – Shaping the future“ (DLR 2014). Außerdem werden ausgewählte Videos, z.B. mit spektakulären Sonnenuntergängen oder besonderen Wetterphänomenen, als Highlights in einem eigenen Bereich auf der Webseite präsentiert. Um die Videos nutzbar zu machen, werden die durch Rayleigh- und Mie-Streuung entstandenen Effekte minimiert sowie die Kontrastwerte und Farbintensität verbessert (RIENOW et al. 2015^B). Geotagging ermöglicht die Bestimmung des aktuellen Aufenthaltsortes der ISS und ist die Grundlage für die kartenbasierte Suchfunktion der Highlights. Mithilfe eines Suchalgorithmus können die Highlights entsprechend ihrer Zuordnung nach Thema, Ort oder Datum gefunden werden. Die auf dem Lernportal übersichtlich präsentierten Videos und Hintergrundinformationen sind verbunden mit zahlreichen Lerneinheiten, die im Schulunterricht eingesetzt werden können. Dieser intermediale Ansatz stärkt (1) naturwissenschaftliche Bildung - und damit einhergehend - (2) die Förderung des wissenschaftlichen Nachwuchses sowie (3) die öffentliche Unterstützung (bei zukünftigen) Weltraummissionen (RIENOW et al. 2014, 2015^B) Grundlage zur erfolgreichen Umsetzung des vorgestellten Ansatzes sind die interaktiven, benutzerfreundlichen und innovativen Lernmaterialien, erstellt nach einem moderaten konstruktivistischen Konzept. Die Ausarbeitung und Weiterentwicklung des Konzepts basiert auch auf den Erkenntnissen, die aus dem Projekt „Fernerkundung in Schulen“ (FIS) (www.fis.uni-bonn.de) gewonnen wurden. Weil die Stärkung der Medienkompetenz heute ein wesentliches Lernziel in der Schule ist, wird interaktives und intermediales Arbeiten mit Videos bei Columbus Eye in den Mittelpunkt gestellt (DGFG 2014).

3.2 Innovatives Lernen trifft Weltraumtechnik

Computergestütztes E-Learning ist kein Neuland in den Erziehungswissenschaften. Es ist bekannt, dass interaktive Lernmaterialien den SuS helfen, autodidaktische Kompetenzen zu entwickeln, ihre Fähigkeiten zur Selbstorganisation zu stärken sowie Entscheidungskompetenzen zu fördern. Das Konzept des E-Learnings nutzt u.a. virtuelle Klassenzimmer, computergestützte Lerneinheiten oder Online-Suchmaschinen (VOß et al. 2011, GRYL 2012). Ein vergleichsweise neuer Ansatz für den Einsatz virtueller Lernumgebungen ist die sogenannte Augmented Reality (AR), eine in erster Linie smartphonebasierte Technik, die analoge Lernmaterialien durch virtuelle Welten erweitert. Der Wandel vom „E-Learning“ zum „M-Learning“, bzw. von der Online- zur

Mobilen Lernumgebung bringt den Vorteil mit sich, dass mobile Endgeräte zu jeder Zeit und an jedem Ort verfügbar sein können – und damit auch die Lernmaterialien (CLARKE 2008, KORUCU & ALKAN 2011). Apps, kleine Smartphone-Anwendungen, verlinken Karten, Schaubilder oder reale Umgebungen mit passenden Informationen und ergänzen die Realität der Echtzeit-Bilder durch virtuellen Inhalt (DUNLEAVY et al. 2009; VUFORIA 2016). Informationen können so unabhängig von einem speziellen Lernort abgerufen werden. Zweidimensionales Kartenmaterial wird so zum interaktiven Spielplatz, bei dem Videos von der ISS oder Astronautenbilder wie das Lichtermeer der nächtlichen Erde in einer neuen Dimension entdeckt werden können. Diese Erfahrung lehrt zum einen die kritische Betrachtung des beobachteten Objekts, fördert aber auch die Auseinandersetzung mit der sinnvollen Nutzung des Smartphones hand (CLARKE et al. 2008, KORUCU & ALKAN 2011; VUFORIA 2016).

Hinter den Kulissen entsprechen die Apps den neuesten technischen Voraussetzungen, sind aber gleichzeitig auch mit älteren Android-Versionen kompatibel. Entwickelt wurde die App mit dem Android Developer© erweitert durch Vuforia© und wird verbreitet über das Columbus Eye-Portal und den Google Play Store©, kann die App von mehr als 85% aller Smartphone-Nutzer weltweit genutzt werden (STATISTA 2016).

4 Die Umsetzung: Interaktive Lernmaterialien

Die Erde wird rund um die Uhr aus drei verschiedenen Perspektiven beobachtet, sodass das produzierte Bildmaterial vor allem für forschungsorientierte Fächer wie Physik, Sozialwissenschaften und Geografie von großem Interesse ist. Durch die Beschäftigung mit ISS-Daten können somit einerseits die technischen Grundlagen der Bildaufnahme und Verarbeitung und andererseits die Prozesse auf der Erdoberfläche vermittelt werden. Dies kann aus unterschiedlichen Perspektiven geschehen -nicht in Bezug auf verschiedene Kamerawinkel, sondern auch durch unterschiedliche Formen von Lehr- und Lernmaterialien (Tab. 1). Im folgenden Kapitel sollen kurz die oben vorgenannten computerbasierten Lernmaterialien und Online-Tools sowie die neu explorierten 3D und M-Learning Techniken beschrieben werden.

4.1 Computerbasierte Lernmodule

Das wichtigste Element des Columbus Eye-Portals sind die Unterrichtsmaterialien, die sich jeweils einem Thema widmen, das dem Schulcurriculum im Bereich der MINT-Fächer entnommen ist. Die Materialien wurden so gestaltet, dass sie sofort einsetzbar sind und von den Schülern selbständig bearbeitet werden können (RIENOW et al. 2015a). Nach einer problemorientierten Einführung in das jeweilige Thema folgt ein interaktiver Teil, in dem die SuS Techniken der Fernerkundung anwenden, um die für die Beantwortung der Fragen relevanten Informationen zu erhalten. Mithilfe des Tests am Ende des Moduls können die SuS selbständig ihren Lernerfolg überprüfen. Damit die Materialien ohne aufwendige Vorbereitung im Unterricht eingesetzt werden können, wurde zu jeder Unterrichtseinheit ein Leitfaden für LuL entwickelt.

Ein Beispiel ist das mathematische Lernmodul „Mittelwertberechnung von der ISS“. In diesem Modul sollen die SuS mithilfe mathematischer Berechnungen Bildkorrekturen vornehmen (Abb. 1). Als Datenquelle dient ein von der ISS aufgenommenes Bild. Die SuS machen sich zum einen

mit Methoden der Statistik und Informatik vertraut und können zum anderen das theoretische Wissen gleich in die Praxis umsetzen (RIENOW et al. 2016).

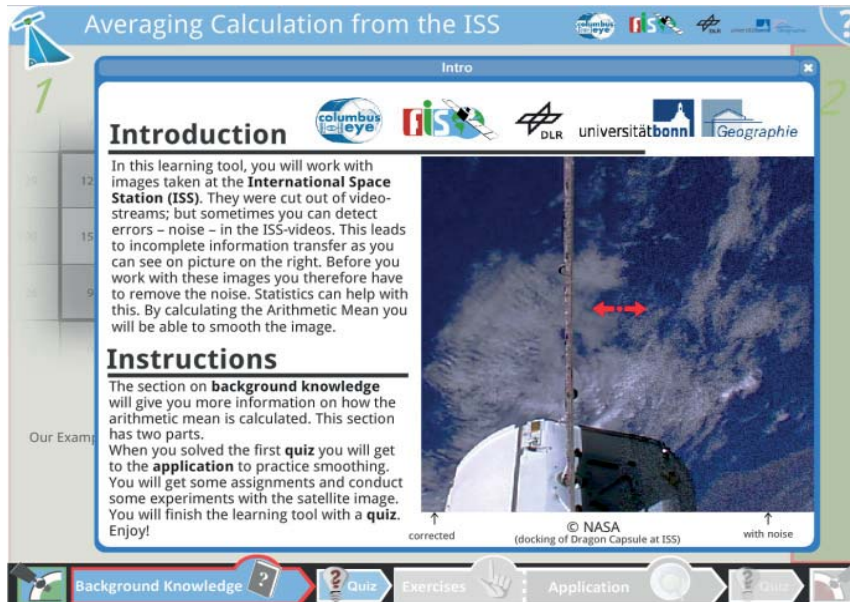


Abb. 1. SuS kalkulieren den Mittelwert, um Bildkorrekturen vorzunehmen

4.2 Online-Klassifizierungstool: Das Observatorium

Eine andere Möglichkeit, mit dem Bildmaterial der ISS zu arbeiten, bieten die aus den Videos entwickelten einzigartigen Panorama-Aufnahmen. Durch die Umwandlung von bewegten in statische Bilder können verschiedene Methoden der Fernerkundung erprobt werden (RIENOW et al. 2015b). Bisher befinden sich in dem sogenannten „Observatorium“ Panorama-Bilder von drei Regionen: Westafrika, Südamerika und Kanada. Imposante Eisdecken, die größte Wüstenregion der Welt und tropischer Regenwald stehen hier bereit zur Erforschung. Mithilfe von leicht zu bedienenden Online-Klassifizierungstools können die SuS anhand der Panoramabilder Informationen über Landoberflächen gewinnen und werden so mit Arbeitsabläufen in der Fernerkundung vertraut gemacht (RIENOW et al. 2015b; VOß et al. 2011). Die SuS können ihre Übungseinheit je nach Interesse auswählen und selbst eine Klassifizierung vornehmen. Dies geschieht auf Basis der Minimum-Distance-Klassifikation (Abb. 2). Bei diesem überwachten Klassifikationsverfahren werden mit Hilfe von Trainingsgebieten Klassen voneinander abgegrenzt. Das wesentliche Kriterium ist der Abstand der zu klassifizierenden Pixel zum Mittelwert der Farbeigenschaften, die durch die Trainingsklassen repräsentiert werden. Die Zuweisung eines Pixels zu einer bestimmten Klasse erfolgt über den geringsten euklidischen Abstand, d.h. ein Pixel wird der Klasse zugewiesen, zu deren Mittelwert er den kleinsten Abstand besitzt. (WACKER & LANGREBE 1972).

Im Onlinetool werden die Trainingsgebiete durch den Button „neue Oberfläche anlegen“ ausgewählt. Für jedes Panorama können gleichzeitig verschiedene Klassifizierungen angewendet werden, zum Beispiel können die Wolkendecken im Süden im Kontrast zu den nördlichen Wüsten hervorgehoben werden – beides im Westafrika-Panorama zu finden. Das Klassifizierungstool kann

sowohl als eigenständige Anwendung genutzt als auch in einen größeren Lernzusammenhang eingebettet werden. Erlern werden nicht nur Fernerkundungstechniken, sondern auch das Erstellen von thematischen Karten wie im Lehrplan für Geographie vorgesehen. Bei der abschließenden Interpretation der Daten helfen den SuS kurze Texte, die in die Umgebung des Lernmoduls integriert sind.



Abb 2. Westafrika im Fokus: Durchführung einer Minimum-Distance-Klassifikation anhand von Hintergrundinformationen zu einer bestimmten Landoberfläche

4.3 Mit Stift, Papier und Smartphone ins All abheben

Klassische Arbeitsblätter haben den Vorteil, dass sie klar strukturiert sind und den SuS erlauben, in ihrem eigenen Tempo zu arbeiten und das Erlernete jederzeit zu wiederholen. Gleichzeitig sollen herausfordernde Aufgaben selbständiges Denken fördern. Die Arbeitsblätter von Columbus Eye folgen diesen Prinzipien, werden aber ergänzt durch interaktive Lernmaterialien. Die Arbeitsblätter behandeln Themen aus verschiedenen naturwissenschaftlichen Fachrichtungen wie zum Beispiel Atmosphärische Streuung (Physik), Metropolregionen und ihr Naturraum (Biologie/Geografie) oder Wüsten und ihre Entstehung (Geografie). Durch die Einbindung von 3D-Umgebungen und Apps wurden die analogen Lernmaterialien in die nächste Generation überführt.

4.3.1 Lern-Apps erweitern die Realität

In diesem Jahr hat Columbus Eye die erste eigene App entwickelt und auf den Markt gebracht. Die App ist Teil der multimedialen Lerneinheit „Im Auge des Zyklons“, die auf der Struktur eines klassischen Arbeitsblattes basiert und zugleich AR-Techniken anwendet. Thematisiert werden die Entstehung und der Weg des philippinischen Zyklon Maysak. Analog zum multimedialen Ansatz können die analogen Arbeitsblätter zum Leben erweckt werden, wenn man sie durch die Kamera des Smartphones betrachtet. Z.B. wird ein Diagramm über die geheimnisvollen internen Mechanismen des Zyklons kombiniert mit einem Video, das am 31. März 2015 von der ISS aus

vom Zyklon Maysak aufgenommen wurde und einen Blick auf dieses besondere Wetterphänomen ermöglicht. Ein zweites Diagramm über Luftmassen bewegt sich plötzlich und zeigt, wie sich der Zyklon Maysak fortbewegt. (Abb. 3). Die Luftmassenkarte und der Aufriss des Zyklons dienen hierbei als reale Objektgrundlagen. Sie sind der Marker zur Bilderkennung für die AR-Inhalte der App. Das Motiv wird von der App gescannt und erweitert die Realität um animierte Luftmassenkarte und einen Überflug über den Tropensturm.

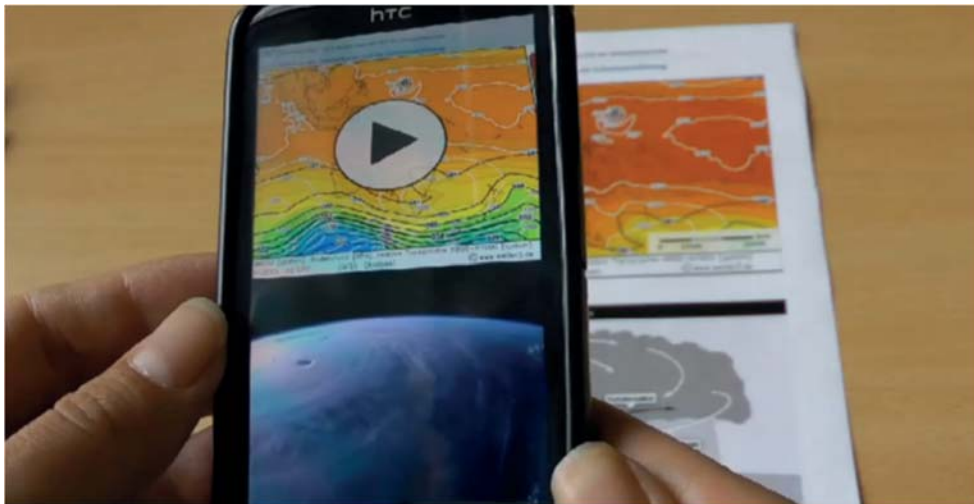


Abb. 3: Demonstration der Android App "Im Auge des Zyklons"

Bevor sie die interaktiven Elemente entdecken können, müssen sich die SuS auf analogem Wege wissenschaftliches Spezialwissen aneignen, z.B. über Häufigkeit, Formation und innere Struktur eines Zyklons. Die Lese- und Verständniskompetenzen der SuS werden im letzten Teil der Lerneinheit gefordert, wenn Aufgaben zum vorgestellten Thema gelöst werden müssen. Um die Aufgaben lösen zu können, ist es notwendig, sowohl die Informationen aus den Arbeitsblätter als auch diese aus der App zu kennen, denn während die analogen Materialien eher allgemeine Informationen zu Tropenstürmen enthalten, kann man nur über die Nutzung der App an konkrete Daten über den Tayfun Maysak gelangen, wie z.B. die Größe seines Durchschnitts.

Wenn es schließlich darum geht, das erlernte Wissen anzuwenden, kommen Stift und Papier ins Spiel. Der haptische Eindruck beim Notieren der Lösungen auf einem Arbeitsblatt sorgt dafür, dass Erlerntes nachhaltig im Gedächtnis bleibt.

4.3.2 Drei Dimensionen des Weltraums

Columbus Eye arbeitet neben der Weiterentwicklung der AR-Anwendungen auch daran, Arbeitsblätter durch 3D- und stereoskopischen Bildern zu erweitern. Mithilfe von stereoskopisch bearbeiteten Satellitenbildern, Bildern von der ISS und Bildern der Astronauten sowie 3D-Videos sollen SuS physikalischen Hintergründe und Zusammenhänge der Stereoskopie von der Absorption über Komplementärfarben, Farbe und Polarisationsfilter bis hin zu stereoskopischem Sehen erlernen. Mit diesem Wissen können die SuS dann ihre eigenen 3D-Bilder produzieren und so die physikalischen Grundlagen neuer Technologien erlernen, die beispielsweise beim 3D-Fernsehen oder in der Virtual Reality (VR) eingesetzt werden.



Abb. 4: Eintauchen in die Mojave-Wüste – 3D Erlebnisse mit der Rot-Cyan-Brille (prozessiertes Bild ISS048-E-68432, 30.08.2016, 20:21:00 GMT). Abrufbar unter: <https://www.facebook.com/columbuseyelite/photos/a.570448506427259.1073741828.570411923097584/688844064587702/?type=3&theater>

Auf dem Arbeitsblatt werden alle relevanten Techniken und Methoden vorgestellt, die für die Erstellung und das Anschauen von Stereoaufnahmen relevant sind, wie z.B. die Polarisierungsmethode oder Anaglyphenbilder, die aus zwei verschiedenen gefilterten Farbbildern bestehen, die jeweils von einem Auge durch „farbkodierte“ Gläser (rot und cyan, Abb. 4) betrachtet werden. Das Arbeitsblatt ist dazu in zwei Bereiche aufgeteilt. Im ersten Teil werden alle nötigen Hintergrundinformationen erstellt und im zweiten Teil werden Beispiele von 3D-Bildern gezeigt, die entweder Satellitenbilder oder Bilder von der ISS sind. Die Anaglyphenbilder von der ISS wurden anhand der HDEV-Videos mit MATLAB© errechnet (MICHEL 2013). Außerdem gibt es noch einen Fragenkatalog, der von den SuS beantwortet werden soll. Für die Beantwortung der Fragen sollte das Material aufmerksam gelesen und die physischen Grundlagen verstanden worden sein.

5 Ein Blick in die Zukunft: Technologie weiter denken

Fernerkundung und Erdbeobachtung eignen sich als Methode für die Vermittlung von Themen des interdisziplinären Schulunterrichts, indem sie die Lücke zwischen physikalischen und mathematischen Hintergrundinformationen und geographischer Anwendungsanalyse schließen. Das Projekt Columbus Eye ist im Jahr 2013 mit dem Ziel gestartet, ein Portal für die Live-Bilder von der ISS und die begleitenden Unterrichtsmaterialien einzurichten. Initiiert wurde das Projekt vor allem, um die Mission des deutschen Astronauten Alexander Gerst im Jahr 2014 zu begleiten und so die Themen Raumfahrt und Erdbeobachtung miteinander zu verknüpfen. Das Portal enthält leicht zugängliche Arbeitsblätter, verständliche Lernmodule, interaktives Lehrmaterial sowie 3D-Bilder und smartphone-basierte AR-Anwendungen.

Durch Projekttag in Schulen und Lehrerfortbildungen in Kooperation mit Bildungseinrichtungen in ganz Deutschland gelangten faszinierende Bilder von der ISS direkt in die Klassenzimmer. Bis jetzt konnten mehr als 1200 SuS und 200 LuL direkt erreicht werden. Über die Projekttag in den Schulen wurde im Fernsehen und Radio berichtet, wodurch das Interesse noch einmal verstärkt werden konnte (RTL NORD 2016).

Für die Zukunft des E-Learnings ist das Ausschöpfen des Potenzials neuer Technologien essentiell. Nur so kann die Lebenswirklichkeit der SuS sinnvoll in den Lehrplan integriert werden und gleichzeitig die Methodenkompetenz im Umgang mit den neuen Medien angesprochen und ausgebaut werden. Die zunehmende Vernetzung der SuS und LuL kann dabei auch für innovative Unterrichtsmaterialien und vor allem -Umgebungen genutzt werden. Hier könnten die an Hochschulen eingesetzten Massive Open Online Courses (MOOCs) für den Einsatz im Schulunterricht angepasst werden. MOOCs kombinieren traditionelle Formen der Wissensvermittlung mit online abrufbaren Videos, Lesematerial und Problemstellungen. Sie können wie reguläre Hochschullehrveranstaltungen auch mit Prüfungsleistungen abschließen. Während mehrwöchige MOOCs zum Thema Erdbeobachtung für den alltäglichen Schulunterricht eher ungeeignet scheinen, könnten Mini-MOOCs den Einsatz von Präsenzveranstaltungen, i.e. regulärer Unterricht im Klassenraum, mit Videokursen als bspw. wochenübergreifende Hausaufgaben kombiniert werden. Denkbar wären auch verstetigte Einsätze in Vertretungsstunden oder die Gründung von SuS-Arbeitsgruppen.

Mit dem Ende der HDEV-Mission in 2018 arbeitet die NASA zurzeit an einem Nachfolgeprojekt, um die Öffentlichkeit auch in Zukunft am Blick aus dem All teilhaben zu lassen. Die Herausforderung der Zukunft wird es sein, diese neuen Techniken möglichst zeitnah und barrierefrei mit curriculumsorientierten Vermittlungsmethoden einem breiten Publikum zur Verfügung zu stellen – denn nur so können faszinierende Lernumgebungen entstehen, die die SuS bei einer Reise durch das All schulpraktisches Wissen erfahren lassen.

6 Literaturverzeichnis

- CLARKE, K., DEDE, J.C. & DIETERLE, E., 2008: Emerging Technologies for Collaborative, Mediated, Immersive Learning. VOOGT, J. & KNEZEK, G. (Hrsg.), International Handbook of Information Technology in Primary and Secondary Education, 901-909.
- DEUTSCHES ZENTRUM FÜR LUFT- UND RAUMFAHRT (HRSG.), 2014: Blue Dot – Alexander Gerst shapes our future on the International Space Station, Bonn.
- DGFG (DEUTSCHE GESELLSCHAFT FÜR GEOGRAFIE) (HRSG.), 2014: Bildungsstandards im Fach Geographie für den Mittleren Schulabschluss mit Aufgabenbeispielen, Bonn.
- 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.
- GRYL, I., 2012: Reflexivity and Geomedia – Going Beyond Domain-specific Competence Development. *GI-Forum Geovisualization, Society and Learning*, Salzburg, Austria, 03.-06. July 2012, elektronischer Tagungsband.

- 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*. Belgien.
- 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., MENZ, G., WEPPLER, J. & RUNCO, S., 2014: *Columbus Eye – High Definition Earth Viewing from the ISS in Secondary Schools*. IAC-14-E1.1.8, 65th International Astronautical Congress, Toronto, Canada, 29 September - 03 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.
- RTL NORD, 2016, Projekttag an der Oberschule In der Sandwehen - RTL Nord. TV Sendung vom 20.04.2016. [http://rtlnord.de/nachrichten/projekttag-columbus-eye-an-der-oberschule-an-der-sandwehen.html](http://rtl nord.de/nachrichten/projekttag-columbus-eye-an-der-oberschule-an-der-sandwehen.html) (abgerufen am 13.01.2017).
- RUNCO, S., 2015: *International Space Station – High Definition Earth Viewing (HDEV)*, http://www.nasa.gov/mission_pages/station/research/experiments/917.html, letzter Zugriff am 13.01.2017.
- STATISTA (Hrsg.), 2016: *Prognose zu den Marktanteilen der Betriebssysteme am Absatz vom Smartphones weltweit in den Jahren 2016 und 2020*. <http://de.statista.com/statistik/daten/studie/182363/umfrage/prognostizierte-marktanteile-bei-smartphone-betriebssystemen/>, letzter Zugriff am 13.01.2017.
- VOB, 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.
- VOB, 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 13.01.2017.
- WACKER, A.G. & LANDGREBE, D.A., 1972: *Minimum Distance Classification in Remote Sensing*, LARS Technical Reports 25.

Nutzung eines Body of Knowledge für das Curriculum Design

WOLFGANG REINHARDT¹ & IMRAN MD. HOSSAIN¹

Zusammenfassung: Ein „Body of Knowledge“ (BoK) beschreibt Konzepte, Terme und Aktivitäten die ein Fachgebiet ausmachen. Er wird naturgemäß von Experten des Fachgebietes erarbeitet. Für das Gebiet „Geographic Information Science & Technology“ (GI S&T) wurde im Jahr 2006 vom UCGIS ein BoK vorgelegt. Im September 2016 wurde das Projekt GIN2K beendet, in dem eine europäische Version eines GI S&T BoK erstellt und eine Softwareumgebung geschaffen wurde, die es ermöglicht den BoK in „crowd-sourcing-Manier“ über Webfunktionen weiterzuentwickeln und zu nutzen. Außerdem wurde ein Softwaretool entwickelt, das die Erstellung eines Curriculums unter Nutzung des BoKs und anderer Quellen unterstützt. In dem Beitrag werden die Hintergründe und grundlegenden Konzepte der genannten Softwaretools beschrieben, sowie die Vorgehensweisen zur Erstellung eines Curriculum unter Nutzung des Softwaretools erläutert.

1 Einleitung

Ein „Body of Knowledge“ (BoK) beschreibt Konzepte, Terme und Aktivitäten die ein Fachgebiet ausmachen. Er wird naturgemäß von Experten des Fachgebietes erarbeitet. Für das Gebiet „Geographic Information Science & Technology“ (GI S&T) wurde im Jahr 2006 vom „University Consortium of Geographic Information Science“ (UCGIS) ein BoK in Form eines Fachbuches bzw. Leitfadens vorgelegt (UCGIS 2006), der hierarchisch in Wissensgebiete, Einheiten und Themen strukturiert ist. Dieser wurde in der Folge vielfach diskutiert, z.B. in REINHARDT & TOPPEN (2008), insbesondere auf seine Übereinstimmung mit der europäischen Sicht und seine Nutzbarkeit in der vorliegenden Form.

In Anlehnung an UCGIS (2006) soll der GI S&T BoK im Wesentlichen folgende Aufgabenstellungen unterstützen:

- Planung und Entwicklung des Curriculum / der Inhalte von Lehrveranstaltungen
Ein BoK unterstützt die Auswahl von Wissensgebieten / Themenbereichen sowie die Ausgestaltung von Inhalten einzelner Vorlesungen ebenso wie die Profil- bzw. Schwerpunktbildung von Studiengängen. Außerdem beinhaltet der BoK Kerngebiete die als absolut erforderlich gelten. Selbstverständlich kann die Planung und Entwicklung des Curriculum sowie der Inhalte von Lehrveranstaltungen nur durch Experten erfolgen, die die Anforderungen an Studierende des jeweiligen Studiengangs und die gewünschten Schwerpunkte kennen (vgl. Abschnitt 2.2). Ein BoK kann dabei als Grundlage für die Auswahl dienen.
- Studiengang Akkreditierung

¹ Universität der Bundeswehr München, Institut für Angewandte Informatik, D-85577 Neubiberg,
E-Mail: [Wolfgang.Reinhardt, Imran.Hossain]@unibw.de

Bei der Akkreditierung kann ein BoK aus den beim letzten Punkt genannten Gründen als Referenz für die Gutachter dienen.

- Studiengangs-Auswahl und Vergleich
Durch den Vergleich des Curriculum mit dem BoK wird das Erkennen der Schwerpunkte von Studiengängen erleichtert (siehe hierzu 2.2/2.3)
- Erstellung von Mitarbeiterprofilen und Auswahl von Bewerbern
Ein BoK ist gut geeignet als Basis für die Erstellung von Mitarbeiterprofilen (auch „Job-Profile“ genannt), in denen deren Kompetenzen deutlich werden. Ebenso können Firmen bei der Auswahl von Bewerbern deren Profil und Kompetenzschwerpunkte mit Hilfe des GI S&T BoK besser einordnen.

Für weitere Anwendungen sei auf den GI S&T BoK (UCGIS 2006) sowie auf Prager & Plewe (2013) verwiesen.

Die Defizite des BoK haben dazu geführt, dass eine Gruppe von in den Vereinigten Staaten tätigen GI Wissenschaftlern ein Projekt zur Konzeption und Entwicklung einer Ontologie-basierten Datenhaltung für den BoK und verschiedener Softwaretools für die Nutzung des BoK initiierten und bearbeiteten (AHEARN et al. 2013).

Nur geringfügig zeitversetzt wurde von einer größeren Gruppe von europäischen GI Wissenschaftlern unter der Leitung der KU Leuven ein Projektantrag ausgearbeitet, der zum Projekt GI-N2K (GI-N2K, 2013) geführt hat, das von der Europäischen Union unter dem „Life Long Learning Programme“ (LLP) von Oktober 2013 bis September 2016 gefördert wurde. Über das Projekt wurde bereits auf der gemeinsamen Jahrestagung in Hamburg berichtet (REINHARDT 2014).

Die Ziele des Projektes im Überblick:

1. Analyse der detaillierten Anforderungen an den BoK aus Sicht der Ausbildungsorganisationen wie Universitäten und Hochschulen sowie der Organisationen bei denen die Absolventen tätig werden können (Firmen, Behörden, etc.).
2. Konzeption und Entwicklung von Software-Werkzeugen zur Pflege der Inhalte des BoK sowie der Nutzung (in Kooperation mit den amerikanischen Kollegen). Als wesentliche Entwicklung zur Nutzung des BoK ist ein Softwaretool zur Unterstützung der Erstellung eines Curriculums zu nennen (siehe Abschnitt 2).
3. Überarbeitung / Erweiterung der Inhalte des BoK, auch hier in Kooperation mit den amerikanischen Kollegen.
4. Praktische Überprüfung der Nutzbarkeit des BoK.

Diese Ziele wurden in acht Arbeitspaketen bearbeitet. Der 1. Autor diese Beitrags war zusammen mit einem weiteren Kollegen Leiter des Arbeitspakets, das sich mit der Entwicklung und dem Test einer Softwareplattform zur Pflege der Inhalte des BoK, sowie zu deren Nutzung befasste. Außerdem leitete er eine Gruppe, die sich die Überarbeitung des Wissensgebietes Design Aspekte (DA) zum Ziel gesetzt hatte.

In dem Beitrag werden die Hintergründe und grundlegenden Konzepte der genannten Softwaretools beschrieben, sowie die Vorgehensweisen zur Weiterentwicklung des BoK und zur Erstellung eines Curriculum unter Nutzung des Softwaretools erläutert.

2 Softwaregestützte Erstellung eines Curriculum

Die Entwicklung von Geoinformatik-Curricula bzw. eines Geoinformatik-Kerncurriculum ist ein wichtiges Thema seit sich der Bereich an den Hochschulen etabliert hat. Als Vorreiter ist diesbezüglich sicherlich das NCGIA GIS Core Curriculum (NCGIA 1990) zu nennen, das in den 1980er entwickelt wurde. Auch in Deutschland gab es einige Ansätze hierzu (z.B. SCHIEWE 2009). Der in der Einleitung thematisierte GIS S&T BoK ist ein weiterer Meilenstein, da er auf eine Beschreibung der für das Fachgebiet relevanten Wissensgebiete und Themen abzielt. Daher ist er hervorragend als Basis für die Erstellung eines Curriculums geeignet. Die Aufgabe des Erstellers ist es jedoch, die Wissensgebiete und Themen auszuwählen und zu entscheiden in welcher Tiefe diese behandelt werden sollen bzw. können. Die ergibt sich aus der zur Verfügung stehenden Stundenzahl und daraus ob die Geoinformatik das Hauptthema – etwa in einem eigenständigen Bachelor- oder- Master Programm – oder nur ein unterstützendes Thema z.B. in einem Geodäsie- oder einem Geologie- Programm ist.

Ebenfalls erwähnt wurde der Fokus des Projektes GI-N2K, Softwaretools bereitzustellen, die es ermöglichen den BoK inhaltlich weiterzuentwickeln und auf komfortable Weise nutzen zu können, worauf im Weiteren einzugehen ist.



Abb. 1: Konzeptionelles Modell für die Entwicklung des GI S&T BoK (AHEARN et al. 2013)

2.1 GI S&T BoK – eine Grundlage zur Erstellung eines Curriculum

Wie mehrfach erwähnt ist die Computer- bzw. Softwaregestützte Haltung des GI S&T BoK eine wichtige Voraussetzung für die ebenfalls Computergestützte bzw. Softwaregestützte Entwicklung eines Curriculum. Dessen Entwicklung war das Ziel eines Projektes einer Gruppe von amerikanischen Geoinformatikern (AHEARN et al. 2013). Dessen konzeptionelles Modell zeigt Abb. 1. Im Zentrum der Entwicklung steht die Ontologiebasierte Haltung der Inhalte des BoK, bezeichnet als „BoKOnto“. Hierfür wurden Werkzeuge des semantischen Web, wie RDF und OWL eingesetzt.

Während die originäre Version des BoK die Inhalte hierarchisch gliederte wurde dies bei dieser Entwicklung aufgegeben und ontologische Prinzipien verwendet. Elementare Speichereinheit ist ein „Konzept“. Dieses kann auf unterschiedliche Weise mit anderen Konzepten verknüpft werden. Durch die Relation „subsequent“ kann eine hierarchische Relation realisiert werden, aber auch allgemeine Beziehungen (Is related to) oder Ähnlichkeiten (is similar to) können über Relationen formuliert werden. Abbildung 2 gibt hierzu ein Beispiel. Die Beschreibungen zu TIN und Voronoi werden als Konzepte gehalten und stehen in einer Beziehung zu einander, sind aber auch subsequente Konzepte von Tesselationen. Diese Relationen erleichtern das Navigieren durch den BoK und das Auffinden von ähnlichen oder in allgemeiner Beziehung stehenden Konzepten.

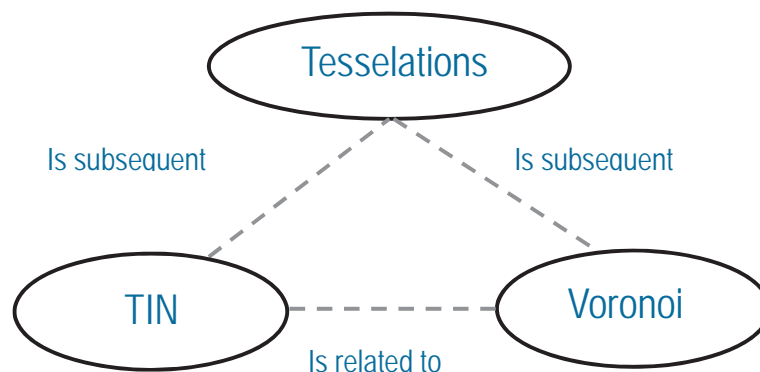


Abb. 2: Beispiele für Konzepte und Relationen im BoKOnto

Weiter wurden Komponenten konzipiert und entwickelt um Konzepte erstellen und abfragen zu können (VPE), diese zu visualisieren (BoKVis) oder um Konzepte mit Akteuren und Aktivitäten zu verbinden (BoKScorecard). Eingaben und Abfragen können auch über das BoK-Wiki erfolgen. Von Bedeutung ist auch, dass auf BoKOnto von externen Applikationen über Web-Services zugegriffen werden kann. Weitere Information zu dieser Entwicklung findet sich in der letztgenannten Literatur.

Diese Entwicklung wurde als technische Basis für die Entwicklung eines europäischen GI S&T BoK verwandt. Dazu wurde eine eigene Instanz des o.g. BokOnto erstellt. Inhaltlich wurde als Ausgangspunkt der UCGIS BoK verwendet.

Die Grundidee der Weiterentwicklung ist ein „Crowd Sourcing“ – Ansatz, bei dem registrierte Experten über die Plattform neue Konzepte und Änderungen vorschlagen können. Über die Annahme dieser Vorschläge entscheidet letztendlich ein Verantwortlicher für ein bestimmtes Wissensgebiet. Bei den Wissensgebieten wurde aus naheliegenden Gründen vom UCGIS BoK ausgegangen. Die Inhalte wurden dabei intensiv diskutiert, auf der Ebene der Wissensgebiete aber nur geringfügig verändert (s. Abbildung 3). Auf den nächsten Ebenen wurden allerdings deutliche Änderungen in der Struktur vorgenommen und es kamen viele neue Themen hinzu und andere wurden gestrichen. Dies ist nicht nur durch die „Europäische Sicht“ verursacht, sondern zumeist durch die in den letzten 10-15 Jahren hinzugekommenen Technologien und Konzepte, z.B. bei den Datenbanken oder in Verbindung mit dem Thema „Web-GIS“. Eine Auflistung der Einheiten und Themen ist im Rahmen dieses Beitrags nicht möglich. Ein Account für den Zugriff sowie die

Mitarbeit bei der Weiterentwicklung des BoK kann unter dem folgenden Link angefordert werden:
<http://mediongraf.init.uji.es/egist-bok/index.php/Special:RequestAccount>.

Analytical Methods	Analytical Methods
Conceptual Foundations	Conceptual Foundations
Cartography and Visualization	Cartography and Visualization
Design Aspects	Design and Setup of Geographic Information Systems
Data Modelling	Data Modeling, Storage and Exploitation
Data Manipulation	Data Manipulation
Geo Computation	Geo Computation
Geospatial Data	Geospatial Data
GI and Society	GI and Society
Organizational and Institutional Aspects	Organizational and Institutional Aspects
	Web based GI

Abb. 3: Wissensgebiete im UCGIS BoK (li) und im GI-N2K BoK (re)

2.2 Das Curriculum Design Tool - Konzept

Das Curriculum Design (CD)Tool dient zur Unterstützung der Erstellung oder der Weiterentwicklung eines Curriculum. Unter einem Curriculum wird dabei die systematische Beschreibung der Lernziele, Inhalte und des Studienplans eines, in diesem Fall, akademischen Studienprogramms. Das Studienprogramm beinhaltet die Pflicht und Wahlmodule des Programms, sowie die erforderlichen Beschreibungen und weitere Informationen. Module sind die Kerneinheiten des Programms und beinhalten in der Regel 2-5 Kurse (Lehrveranstaltungen wie Vorlesungen, Seminare etc.), in denen bestimmte individuelle Fachgebiete vermittelt werden. Kurse bestehen aus einer Reihe von einzelnen Vorlesungen und/oder Übungen.

Abbildung 4 beinhaltet die Struktur des Curriculums.

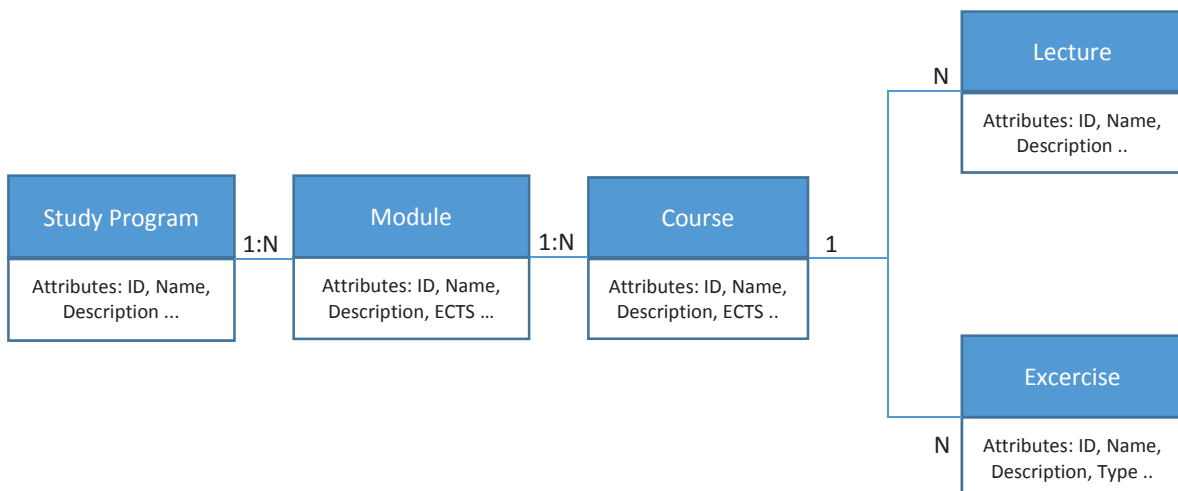


Abb. 4: Struktur des Curriculums im CD-Tool

Die Grundidee der Unterstützung durch das Tool liegt dabei in den folgenden Punkten:

- Die Struktur des Curriculums wird auf Datenbank-Tabellen abgebildet. Die Inhalte der einzelnen Elemente dieser Struktur werden in der CD-Tool-Datenbank gespeichert.
- Die Erstellung wird dadurch unterstützt, dass auf die im BoKOnto gespeicherten Konzepte zugegriffen und durch diese navigiert werden kann. Abbildung 5 zeigt die zu Grunde liegende Architektur. Die Implementierung des CD Tools wurde innerhalb des Projektes im Wesentlichen vom 2. Autor dieses Beitrags vorgenommen. Die Kommunikation mit dem BoK erfolgt über Web-Services.

Im Folgenden soll die Vorgehensweise bei der Erstellung oder Weiterentwicklung eines Curriculum skizziert werden:

- Grundsätzlich liegt die Gestaltung des Curriculum in der Obhut des Nutzers des Tools. Das CD-Tool unterstützt durch die Speicherung der Inhalte und durch die Möglichkeit des Navigierens durch die Inhalte des BoK und durch das Kopieren von Inhalten des BoK in die Struktur des Curriculums. Prinzipiell können auch andere Quellen als der BoK angebunden werden, dies ist allerdings noch nicht realisiert.

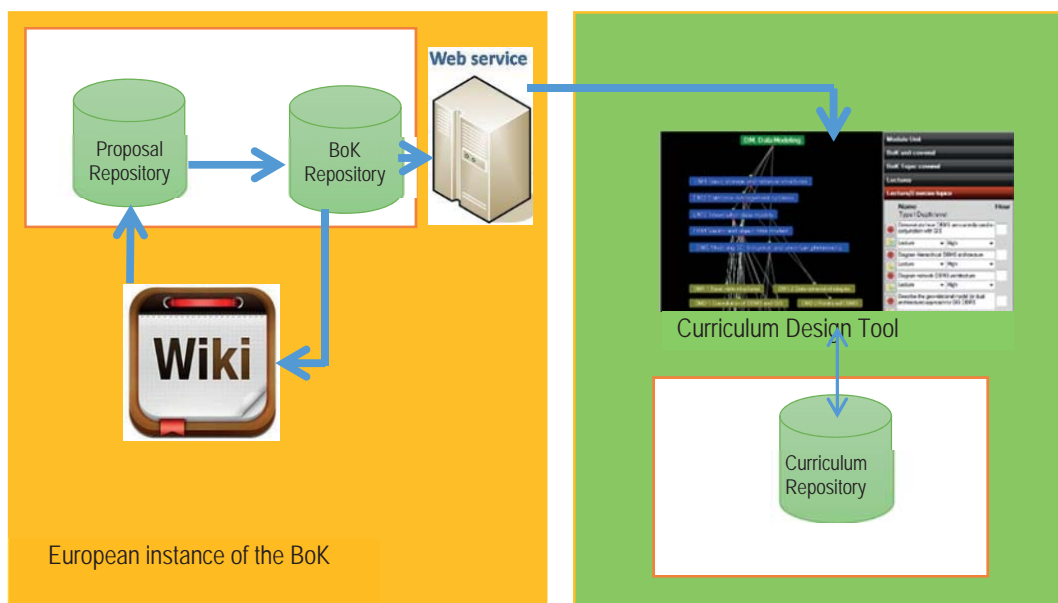


Abb. 5: Architektur des GI-N2K BoK und des Curriculum Design Tools

- Im ersten Schritt werden die Inhalte für das Studienprogramm, wie Name, Beschreibungen etc. erstellt und eingetragen sowie die allgemeinen Informationen zum bearbeiteten Modul und zum jeweiligen Kurs. Im Falle der Weiterentwicklung eines Curriculums sind diese Inhalte bereits vorhanden und können eingelesen werden.
- Auf der Ebene der einzelnen Vorlesungen und Übungen erfolgt dabei die Navigation durch die Konzepte des BoK und das optionale Kopieren von Texten. Wie erwähnt liegt die Gestaltung in der Obhut des Nutzers. Dies bedeutet, dass er/sie bereits Vorstellungen hat, welche Inhalte die einzelnen Vorlesungen und Übungen haben sollen. Mit den dadurch bereits vorhandenen Stichworten kann nun durch den BoK navigiert werden und andere relevante Konzepte gefunden werden. Was davon übernommen wird und wie viel Zeit dafür vorgesehen wird,

hängt im Wesentlichen von der Einschätzung des Benutzers, den vorgegebenen Zeitrahmen und dem Hintergrund der Studierenden ab. Beschreibungen der Konzepte können mit „copy and paste“ aus dem BoK in das CD-Tool übernommen werden.

Abbildung 6 zeigt einen Screenshot des Curriculum Tools, der diesen Vorgang illustriert. Auf der rechten Seite ist das Curriculum in Baumstruktur zu sehen. Die Wurzel stellt das Studienprogramm dar, auf der zweiten Ebene finden sich die Module, darunter die Kurse und darunter die einzelnen Vorlesungen und Übungen, die durch Knoten mit der Beschriftung des jeweiligen Namens repräsentiert werden. Bei vorhandenen Curricula kann durch Anklicken des Knotens z.B. eine Vorlesung ausgewählt werden und die Inhalte können editiert werden. Auf der linken Seite kann durch den BoK navigiert werden, und zwar graphisch (links oben) oder textlich (links unten).

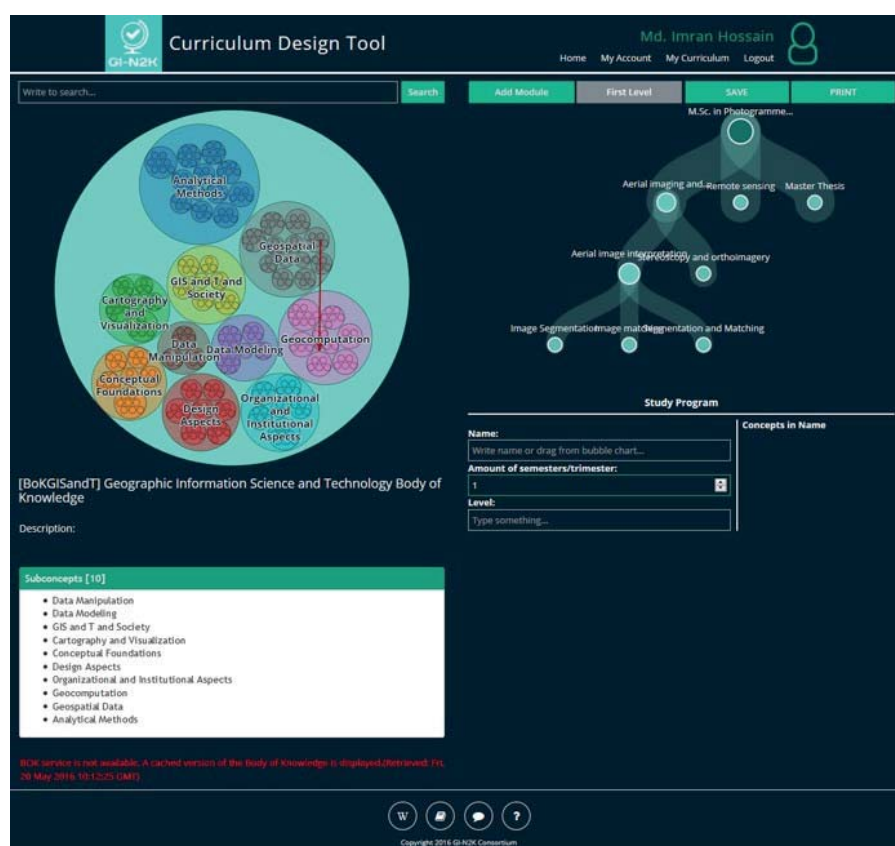


Abb. 6: Screenshot des CD-Tools

2.3 Weitere Anwendungsmöglichkeiten

Das CD-Tool kann nicht nur zur Unterstützung der Erstellung bzw. Weiterentwicklung eines Curriculum verwendet werden, sondern auch zum/zur:

- Vergleich von Curricula
- Erstellung von Job-Profilen

Ersteres geschieht dadurch, dass zwei verschiedene Curricula eingelesen werden und graphisch in Form einer Baumstruktur visualisiert werden. Der Benutzer ist dann in der Lage durch diese

Visualisierung die grundlegende Struktur der Curricula zu vergleichen. Durch Anklicken der Knoten kann auf die jeweilige Beschreibung zugegriffen werden, womit eine Vertiefung des Vergleichs ermöglicht wird.

Ein Jobprofil beschreibt detailliert die Anforderungen an einen beruflichen Tätigkeitsbereich. Dies kann ähnlich wie ein Curriculum strukturiert werden. Auf der untersten Ebene findet sich dann die Beschreibung der erforderlichen Kenntnisse und Fähigkeiten. Da der BoK alle wesentlichen Elemente des Fachgebiets enthält, kann dieser auch dazu benutzt werden ein Job-Profil zu erstellen.

3 Fazit & Ausblick

In diesem Beitrag haben wir beschrieben wie der GI S&T Body of Knowledge (BoK) für die Erstellung eines Curriculums und für andere Aufgaben genutzt werden kann und wir haben eine Softwareumgebung sowie ein Tool vorgestellt, durch das die Erstellung eines Curriculum unterstützt wird. Diese Software ist im Rahmen des im September 2016 beendeten EU-Projektes GI-N2K entstanden, kann aber auch nach Beendigung des Projektes weitergenutzt werden. Ein Link für die Anforderung des Zugangs zum BoK wurde in 2.1 gegeben, für den Zugriff auf das CD-Tool mögen sich Interessierte an die Autoren dieses Beitrags wenden.

4 Literaturverzeichnis

- AHEARN, S.C., ICKE, I., DATTA, R., DEMERS, M.N., PLEWE, B. & SKUPIN, A., 2013: Re-engineering the GIS&T Body of Knowledge. *International Journal of Geographical Information Science*
- GI-N2K, 2013: Projekt-Webseite des EU-Projektes Geographic Information: Need to Know. <http://www.gi-n2k.eu/>, letzter Zugriff am 12.01.2017
- NCGIA, 1990: Core Curriculum GIS for technical programs. <http://www.ncgia.ucsb.edu/cctp/ncgia.html>, letzter Zugriff am 12.01.2017.
- PRAGER, S.D. & PLEWE, B., 2013: Assessment and Evaluation of GIScience Curriculum using the Geographic Information Science and Technology Body of Knowledge. *Journal of Geography in Higher Education*, S 46-69
- REINHARDT, W., 2014: Zur Bedeutung eines „Body of Knowledge“ für die Geoinformatik. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation*, Band 23, Potsdam, SEYFERT, E., GÜLCH, E., HEIPKE, C., SCHIEWE, J., SESTER, M. (Hrsg.), Beitrag 133.
- REINHARDT, W. & TOPPEN, F., 2008: The UCGIS Geographic Information Science and Technology Body of Knowledge – Some thoughts from a European Perspective. *Proceedings of GIScience 2008 - Fifth International Conference on Geographic Information Science*. Park City, Utah, USA.
- SCHIEWE, J., 2009: Kerncurriculum Geoinformatik – Notwendige Grundlage für Studierende, Lehrende und Arbeitgeber. *GIScience*, (4), 137-141.
- UCGIS, 2006: UCGIS GI S&T Body of Knowledge. http://www.aag.org/galleries/publications-files/GIST_Body_of_Knowledge.pdf, letzter Zugriff am 12.01.2017.

Softwareunterstützte Lehre in Photogrammetrie mit PhoX

THOMAS LUHMANN¹

Zusammenfassung: Bei der Photogrammetrie handelt es sich um ein komplexes Fachgebiet in der Hochschullehre, besonders auf den Gebieten der Geodäsie, Geoinformation und Messtechnik, wo hochqualitative Ergebnisse gefordert werden. Vermehrt sind Blackbox-Lösungen für 3D-Bildbearbeitung und Punktwolkengenerierung verfügbar, die sehr einfach Ergebnisse erzeugen, z.B. durch Structure-from-Motion-Ansätze. In diesem Zusammenhang muss der klassische Ansatz Photogrammetrie zu lehren (z.B. mit Schwerpunkt auf Luftbild-Stereophotogrammetrie) reformiert werden, um Studierende und Anwender/Fachleute mit neuen Inhalten weiterzubilden und ihnen mehr Hintergrundwissen zu bieten. Seit etwa 20 Jahren wird in der Photogrammetriausbildung an der Jade Hochschule in Oldenburg Software für die digitale Photogrammetrie eingesetzt, die individuelle Übungen, Genauigkeitsanalysen der Ergebnisse und ein umfangreiches Tool-Angebot zur Visualisierung nahezu aller Standardfragen in der Photogrammetrie liefert. Während der letzten Jahre wurde daraus das Softwarepaket PhoX weiterentwickelt, das Teil eines neuen didaktischen Konzeptes in Photogrammetrie und verwandten Gebieten ist. Ferner dient es auch als Analysewerkzeug in aktuellen Forschungsprojekten. PhoX besteht aus einer projektbezogenen Datenstruktur für Bilder, Bilddaten, Messpunkten und -Merkmalen und 3D-Objekten. Es bietet Werkzeuge für nahezu alle wesentlichen photogrammetrischen Messungen, Bildbearbeitung, Transformationen, grafische Analysefunktionen, Simulationen und mehr.

Studierende nutzen das Programm mit vorgegebenen Übungen und Daten, wobei sie die Möglichkeit haben, die Ergebnisse ausführlich zu analysieren. Darin sind die Analyse von statistischen Qualitätsmerkmalen, aber auch die Bedeutung von Transformationsparametern, Rotationsmatrizen, Kalibrierungs- und Orientierungsparametern enthalten. Als ein spezifischer Vorteil erlaubt PhoX die interaktive Modifikation einzelner Parameter und die direkte Ansicht auf das Ergebnis im Bild- oder im Objektraum.

1 Einleitung

Der Autor lehrt die Fächer Photogrammetrie, digitale Bildverarbeitung, Fernerkundung und Computergrafik seit mehr als 20 Jahren. In dieser Zeit ist die technologische Entwicklung in der Photogrammetrie dramatisch fortgeschritten. Im Besonderen hat die Verfügbarkeit digitaler Bilddaten, der Automationsgrad in digitaler Bildmessung, Kalibrierung, Orientierung und Matching ein sehr hohes Niveau erreicht. Folglich werden heutige Studierende mit mehr oder weniger automatisierten Blackbox-Systemen konfrontiert, in denen die internen Prozessschritte dem Nutzer verborgen bleiben. Andererseits werden die modernen jungen Menschen überflutet von digitalen Daten, und sie erwarten in vielerlei Hinsicht voll-automatisierte Verfahren. Weltweit kann unter den Studenten eine Endverbraucher-Mentalität beobachtet werden.

Die Lehrerfahrungen des Autors in Photogrammetrie schließen Bachelor- und Master-Vorlesungen sowie spezielle Fortbildungsseminare für Berufstätige ein. Eine große Themenauswahl ist in internationalen Lehrbüchern zusammengefasst, z.B. LUHMANN et al. (2014). Der vor-

¹ Jade Hochschule Oldenburg, Institut für Angewandte Photogrammetrie und Geoinformatik, Ofener Straße 16/19, D-26121 Oldenburg, E-Mail: thomas.luhmann@jade-hs.de

liegende Beitrag ist eine aktualisierte und übersetzte Fassung der Publikation von LUHMANN (2016).

Die Ausbildung künftiger Ingenieure soll verschiedene gegensätzliche Qualitäten und Fähigkeiten berücksichtigen. Einerseits sollen Ingenieure in der Lage sein, neue Systeme zu bedienen und Werkzeuge zu kennen, um Projekte mit großen Datenmengen zu managen. Weiterhin sollen sie fähig sein, neue Technologien im Sinne ökonomischer und technologischer Leistungsfähigkeit angemessen einzusetzen. Andererseits sollen die Ingenieure ein profundes Verständnis für Algorithmen und Abhängigkeiten zwischen den verschiedenen Phasen in der Softwareverarbeitung haben, und sie sollen Fehler, Schwachstellen in den Projekten oder Potentiale für Verbesserungen identifizieren können. In diesem Konflikt, der nicht nur eine Herausforderung für den Lehrenden ist, neigen die Studierenden gewöhnlich dazu, sich auf automatische Lösungen zu konzentrieren, anstatt den mühsamen Weg des gründlichen Verstehens und der Analyse zu wählen. E-Learning-Werkzeuge in den Geoinformationswissenschaften finden hauptsächlich in Fernerkundung und GIS-Technologien Anwendung. In der Photogrammetrie existieren für Ausbildungszwecke nur wenige Software-Systeme. Beispiele sind ARPENTEUR (GUSSENMEYER et al. 2002), GRAPHOS (GONZALEZ-AGUILERA et al. 2015) oder E-Foto (RIBEIRO et al. 2014). In diesem Zusammenhang wurde das Software-Paket PhoX entwickelt, um ein Werkzeug für das Selbststudium, Tests und Übungen mit realen photogrammetrischen Daten zur Verfügung zu stellen. PhoX ermöglicht die Verarbeitung nahezu aller Bilder, vom gescannten oder digitalen Luftbild bis zu Fernerkundungsbildern. Der Hauptzweck des Programms liegt in der Datenanalyse, interaktiven Bildmessungen, Simulationen und Visualisierungen der Ergebnisse. Eine praxistaugliche Produktion ist nicht vorgesehen, obwohl es teilweise möglich wäre.

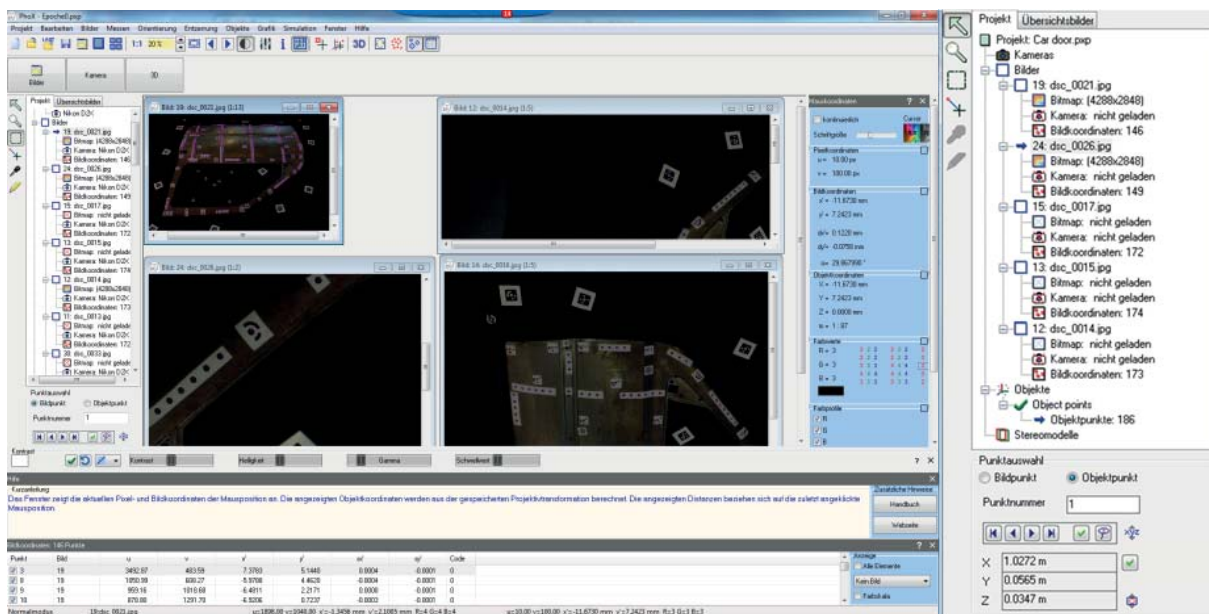


Abb. 1: Programmoberfläche von PhoX

Abb. 2: Projektbaum mit Bildern und Objekten

PhoX steht derzeit mit deutscher und englischer Benutzeroberfläche zur Verfügung (Abb. 1). Kurze Soforthilfeanleitungen können in jeder Phase des Programms kontextsensitiv abgerufen werden. Eine ca. 250seitige Bedienungsanleitung in Deutsch ist erhältlich. Individuelles Lehrmaterial (im PDF-Format) kann verknüpft und mit jeder Menü-Funktion dargestellt werden. Zusätzlich gibt es eine Anzahl vordefinierter Bedienungshilfen, die den Nutzer durch verschiedene Phasen der Verarbeitung leiten. Sie können durch den Nutzer modifiziert oder erweitert werden. Dieselbe Idee wird genutzt, um Übungsmodule für bestimmte Aufgabenstellungen der Photogrammetrie bereitzustellen.

2 Programmstruktur

2.1 Mehrbildansatz

PhoX basiert auf einem projektorientierten Mehrbild- und Multiobjektansatz für das Datenmanagement (Abb. 2). Die Bilder werden als logische Objekte organisiert, die Kalibrierungs- und Orientierungsdaten, Bildpunkte und Konturen, Bildsequenzen und Metadaten enthalten. Bitmaps sind nicht erforderlich, da die photogrammetrische Verarbeitung auch möglich ist, wenn keine Bitmaps gespeichert sind. Vorhandene Bitmaps können in Bildfenstern dargestellt werden, die jederzeit entfernt oder wiedergeladen werden können, um bei großen Bilddatensätzen Speicherplatz zu sparen. Neben Bildern können mit PhoX 3D-Objekte erzeugt und verwaltet werden. Ähnlich wie bei Bildobjekten kann eine unbeschränkte Zahl von 3D-Objekten eingeführt werden. Jedes 3D-Objekt wird als räumlicher Körper mit Koordinaten, räumlichen Transformationsparametern, Polygonen (topologische Verbindungen von Objektpunkten), Punktwolken und Triangulationsknoten betrachtet.

2.2 Import- und Exportfunktionen

Bitmaps können in allen gängigen Formaten geladen und gespeichert werden. Bilddaten wie innere und äußere Orientierungsdaten können in verschiedenen Formaten importiert oder exportiert werden, z.B. Leica LPS, INPHO, AICON, iWitness, PhotoScan, Pix4D und beliebige freie Textformate.

3 Bildmessung

3.1 Digitale Bildverarbeitung und -analyse

PhoX stellt eine große Auswahl an Bildanalyse- und Bildbearbeitungswerkzeugen bereit. Die Bildanalyse beinhaltet u.a. das interaktive Messen von Farbprofilen, Berechnung von Histogrammen oder die Darstellung von Pixel-, Bild- und Objektkoordinaten der Cursorposition. Die Bildbearbeitungsmethoden enthalten alle üblichen Kontrast- und Gammakorrekturen, Glättung, Schärfung oder Kantenfilter, Farbbearbeitung und mehr. Diese Funktionen verändern nicht die Bildgeometrie, können also auf kalibrierte Bilder in einem photogrammetrischen Prozess angewandt werden. Bildbearbeitungsketten können festgelegt werden, in denen mehrere einzelne Verarbeitungen sequenziell durchgeführt werden. Stapelverarbeitung ist möglich für eine frei wählbare Anzahl von Bildern (Abb. 3).

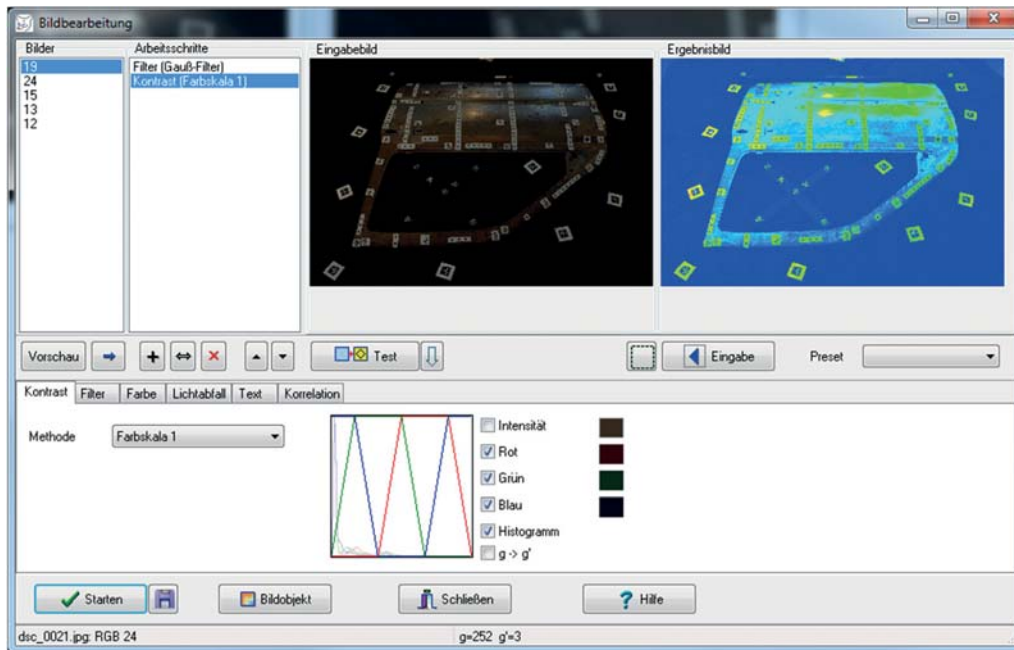


Abb. 2: Bildverarbeitungswerkzeug

3.2 Interaktive Messungen

Die interaktive Bildpunktmessung ist ein Hauptmerkmal von PhoX. Sie ermöglicht manuelle Messungen mit dem Maus-Cursor einschließlich eines Subpixel-Zooms (Abb. 4). Temporäre Gain, Offset und Gammakorrekturen können auf alle Bildfenster angewandt werden, um eine bessere Punkterkennung zu ermöglichen.

3.3 Automatische Bildmessungen

Die automatische Zielmarkenmessung wird durch die folgenden Verfahren ermöglicht:

- Schwerpunktoperator
- Ellipsen-Operator (Abb. 5)
- Template Matching
- Kreuzkorrelation

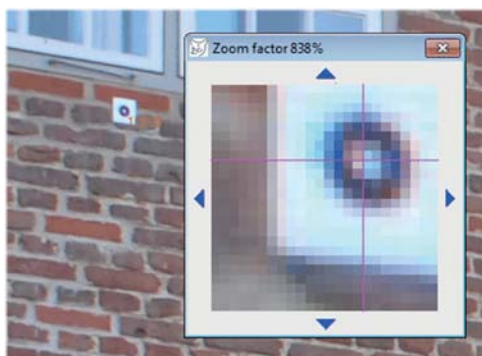


Abb. 3: Manuelle Bildmessung

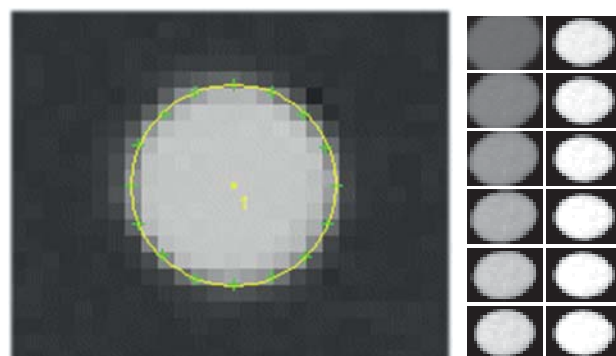


Abb. 4: Ergebnisse einer Ellipsenmessung (links) und eines Template Matchings (rechts)

Die Operatoren sind geeignet, um Markenzentren mit höchster Genauigkeit bis $1/20$ Pixel oder besser zu messen. Da sie Näherungsdaten der Markenpositionen benötigen, können diese durch Punkterkennung, Interest-Operatoren oder Rückprojektion von Objektkoordinaten bereitgestellt werden. Ein automatisches Bildmatching ist nicht enthalten, d.h. Punktnummern müssen manuell generiert werden.

4 Bildentzerrungen

4.1 Orthophotos

Die Entzerrung photogrammetrischer Bilder ist ein wesentlicher Teil in der Photogrammetrieausbildung. PhoX bietet ebene projektive Entzerrungen, Orthophotogenerierung mit Oberflächenmodellen und Entzerrung der Objektflächen mit äußeren Orientierungsdaten. Benachbarte oder überlappende Bilder können zu Bildmosaiken kombiniert werden (Abb. 6).

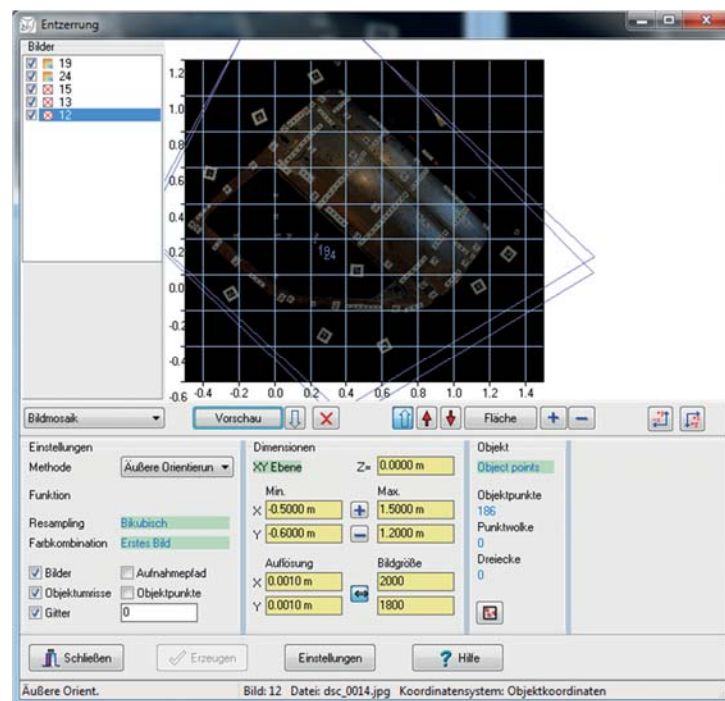


Abb. 5: Entzerrung und Mosaikbildung

4.2 Verzeichnungsfreie Bilder

Bilder mit gegebenen Parametern der inneren Orientierung können zu verzeichnungsfreien Bildern entzerrt werden. Sie werden mit derselben äußeren Orientierung wie das Originalbild, aber mit optimaler fehlerfreier innerer Orientierung erzeugt (Abb. 7).



Abb. 6: Verzeichnetes (links) und verzeichnungsfreies Bild (rechts)

4.3 Stereonormalbilder

Auf der Basis einer relativen oder absoluten Orientierung eines Stereopaars können normalisierte Bilder berechnet werden (Abb. 8). Diese Bilder sind wiederum verzeichnungsfrei mit derselben Translation der äußeren Orientierung wie die Eingabebilder. Auf der Grundlage von Normalbildern können auch Anaglyphenbilder für eine externe Stereobetrachtung erzeugt werden. Der Nutzer kann interaktiv x- und y-Parallaxen einstellen, um die Auswirkung und das Ergebnis dieser Modifikationen zu beobachten.

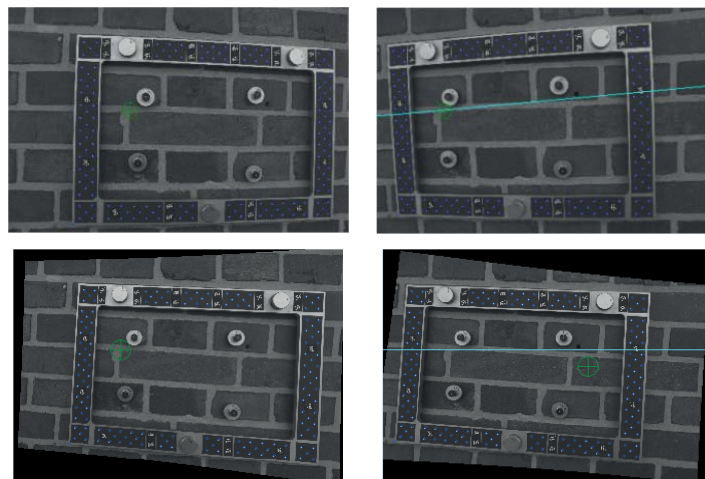


Abb. 7: Originalbildpaar (oben) und Stereonormalbild (unten) mit überlagerter Epipolarlinie

5 Kalibrierung und Orientierung

5.1 Optikberechnungen

Um ein besseres Verständnis für die physikalische Optik zu erzielen, können die Studierenden das Ergebnis einfacher geometrischer optischer Berechnungen beobachten. Parameter wie Schärfentiefe, Bildmaßstab, Bildauflösung und andere können berechnet und interaktiv verändert werden. Das Ergebnis ist unmittelbar in einer vereinfachten Linsendarstellung sichtbar (Abb.9).

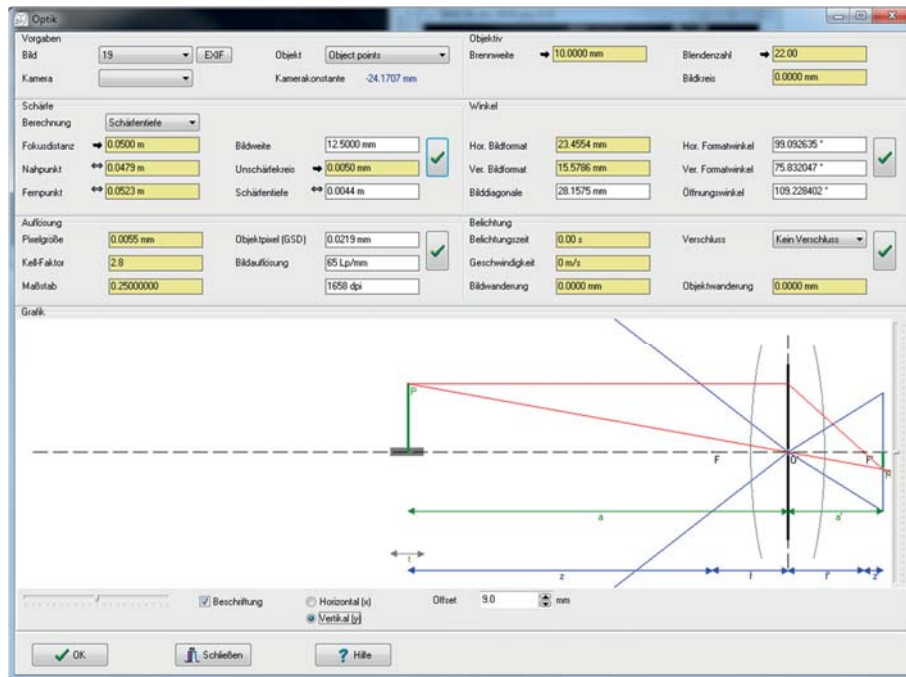


Abb. 8: Berechnung und Darstellung optischer Parameter

5.2 Kameraparameter

PhoX unterstützt den vollen Satz des inneren Orientierungsmodells (Brown-Parameter). Zusätzlich sind spezielle Modelle enthalten wie winkelabhängige Parameter, sphärische Modelle oder Lens-Map-Funktionen. Jeder Kameradatensatz wird in einer Kameradatei gespeichert, jedes einzelne Bild kann jedoch seine eigenen Parameter der inneren Orientierung besitzen.

5.3 Kalibrierverfahren

Bisher ist kein Bündelausgleichungsprogramm mit Selbstkalibrierung implementiert. Näherungsberechnungen von Parameter der inneren Orientierung werden durch erweiterten Rückwärtsschnitt oder durch DLT bereitgestellt. Es sind jedoch Importfunktionen für Ergebnisdateien verschiedener Bündelausgleichungsprogramme implementiert.

5.4 Orientierungsverfahren

PhoX bietet relative Orientierung, absolute Orientierung, Rückwärtsschnitt, DLT und 3D-Ähnlichkeitstransformationen. Jede Funktion basiert auf einer Kleinsten-Quadrate-Ausgleichung. Detaillierte Informationen zu den Ergebnissen werden jeweils in einer Protokolldatei gespeichert, die dem Nutzer zur Verfügung steht (Beispiel in Abb.10).

5.5 3D-Messungen

Gewöhnlich werden 3D-Punkte in PhoX durch räumlichen Vorwärtsschnitt generiert. Die implementierte Funktion arbeitet mit einer unbegrenzten Anzahl an Bildern in jedem beliebigen 3D-Koordinatensystem. Auch hier stehen alle Ausgleichungsparameter in einer Log-Datei.

Ist ein Oberflächenmodell gegeben (z.B. Dreiecksvermaschung), ist auch Monoplotting für 3D-Objektmessungen möglich.

Für Demonstrationszwecke kann eine räumliche 3D-Messmarke genutzt werden. Sie arbeitet nach dem Prinzip des analytischen Plotters, wobei eine gegebene 3D-Position in die Bilder zurückprojiziert wird. Der Nutzer beobachtet Parallaxen oder Abweichungen zwischen Bildern und korrigiert die 3D-Position, bis eine gute Übereinstimmung zwischen allen Bildern (auch mehr als zwei) erreicht ist.

Gemessene 3D-Punkte können zu einer 2½ Dreiecksvermaschung oder Polygonen verarbeitet werden. Einfache 3D-Modelle (z.B. Gebäudemodelle) können mit Texturen überlagert werden.

```

Iteration: 5
Zuschlag : 0.00000
Sigma 0 : 0.00233

X0 [1] : 0.00000 +/- 5.64339
Y0 [2] : 0.00000 +/- 9.95409
Z0 [3] : 0.00000 +/- 2.51305
om [4] : -0.00000 +/- 0.00032
ph [5] : 0.00000 +/- 0.00016
ka [6] : 0.00000 +/- 0.00013

Korrelationen:
X0 1.00000
Y0 0.22827 1.00000
Z0 0.73580 0.16100 1.00000
om -0.22166 -0.99810 -0.15312 1.00000
ph 0.99395 0.23051 0.72500 -0.22394 1.00000
ka 0.12566 0.89263 0.08409 -0.89277 0.13001 1.00000
      X0      Y0      Z0      om      ph      ka

Anzahl Iterationen: 5
Sigma 0: 0.0023
Anzahl Beobachtungen: 12

Transformationsparameter:
X0 =9615.29625
Y0 = 772.80942
Z0 =11882.00717
omega= 1.175859 ( 67.37174*)
phi= -1.977121 (-113.28067*)
kappa= -0.400405 (-22.94153*)
R =
-0.36397369 -0.15405947 -0.91857979
-0.93077686 0.02382638 0.36481056
-0.03431609 0.98777426 -0.15206716
    
```

Abb. 9: Protokolldatei eines räumlichen Rückwärtsschnittes



Abb. 10: Interaktive Manipulation von Orientierungswerten

6 Simulation

6.1 Interaktive Datenmanipulationen

Alle Daten in PhoX sind dem Anwender verfügbar. Konsequenterweise können einzelne Datenwerte oder Parameter interaktiv verändert werden, um deren endgültige Auswirkung zu analysieren.

Ein wichtiges Werkzeug ist die Modifikation von Parametern der inneren und äußeren Orientierung, um die dreidimensionale Auswirkung im Raum zu beobachten und die Projektion von Objektpunkten ins Bild direkt mit einem grafischen 3D-Viewer anzuzeigen. Das Werkzeug ermöglicht unerfahrenen Studierenden mit Rotationen und Daten im 3D-Raum vertraut zu werden und die Wirkung der Kollinearitätsgleichungen zu verstehen (Abb. 11).

6.2 Monte-Carlo-Simulationen

Verschiedene Werkzeuge zur statistischen Fehlerfortpflanzung nach der Monte-Carlo-Methode sind implementiert. Sie erlauben Berechnungen für die Genauigkeitsanalyse von Rückwärtschnitts, Mehrbildvorwärtsschnitt, 3D-Ähnlichkeitstransformation und 6DOF-Berechnungen. Besonders für diese komplexen 3D-Berechnungen ist die Fehlerfortpflanzung schwierig, somit eröffnen die implementierten Werkzeuge einen einfachen Weg, eine Abschätzung für die erreichbare Genauigkeit zu erzielen (Beispiel in Abb. 12).

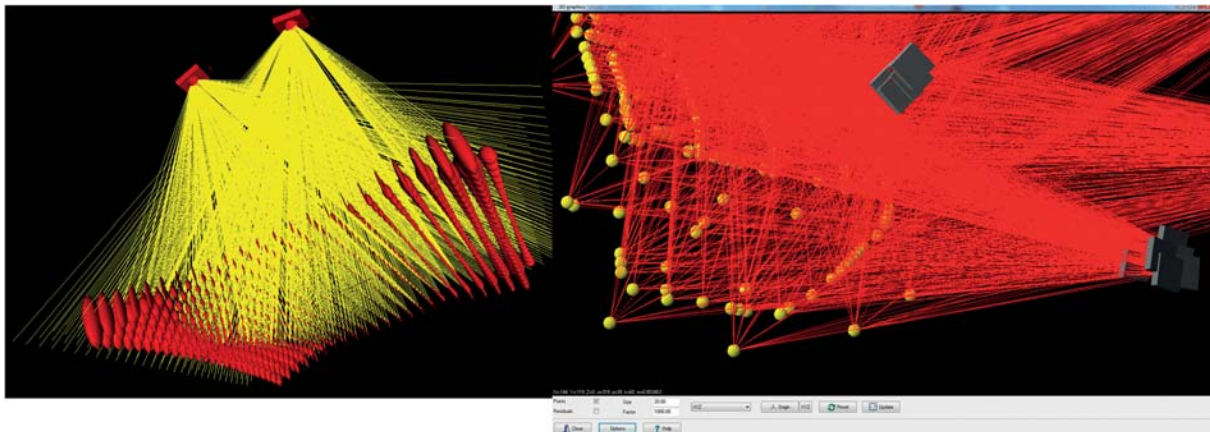


Abb. 11: Berechnung der stereoskopischen Messgenauigkeit im Objektraum durch Monte-Carlo-Simulation

Abb. 12: Interaktiver 3D-Viewer

Ergänzend kann den Bildkoordinaten und den Bildfarbwerten statistisches Rauschen hinzugefügt werden, um den Einfluss auf die Bildmessgenauigkeit zu studieren.

6.3 Synthetische Bilder

Synthetische Bilder können generiert werden, wenn Orientierungsparameter, ein 3D-Oberflächenmodell und Texturbild gegeben sind. Durch Resampling im Objektraum wird das Bild mit Farbwerten aufgefüllt. Synthetische Bilder können genutzt werden, um Referenz- oder fehlerfreie Bilder herzustellen, von denen alle Parameter bekannt sind. Ergänzend bietet PhoX eine große Auswahl synthetischer Bildmuster, darunter Siemensstern, Schachbrettmuster, Gitter, Farbkeile usw.

7 Visualisierungen

PhoX ermöglicht die Darstellung einer unbegrenzten Anzahl von Bildfenstern. Neben der Bild-darstellung werden verschiedene graphische Ausgaben ermöglicht. Für die Darstellung von 3D-Szenen mit Kamerapositionen, Objektpunkten, Oberflächen etc. ist ein integrierter 3D-Viewer verfügbar (Abb. 13). Zusätzliche können Daten im VRML-Format exportiert werden.

Verzeichnungskurven und flächenhafte Auswirkungen können dargestellt werden. Einzelne Parameter (z.B. radiale oder tangentielle Verzeichnung) können ein- und ausgeschaltet werden, um ihren individuellen Einfluss in dem Bild zu beobachten (Abb. 14).

Footprints von Bildern, die Verteilung von Bild- und Objektpunkten, Histogramme und Verteilung von Bild- und Objektresiduen, Längenmessfehler und mehr können in 2D-Diagrammen (Beispiel in Abb. 15) visualisiert und bei Bedarf nach Excel exportiert werden.

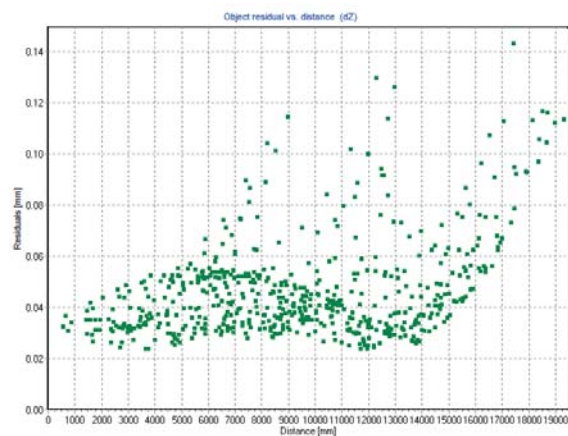


Abb. 13: Verteilung der Residuen von Objektkoordinaten als Funktion des Punktabstandes

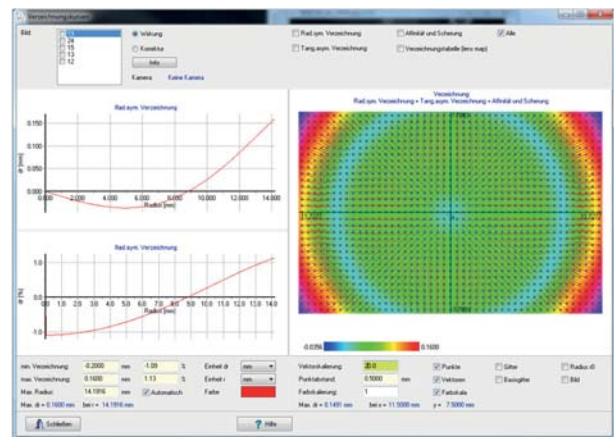


Abb. 14: Verzeichnungskurven und -diagramme

8 Didaktisches Konzept

8.1 Modulbeschreibungen

In der studentischen Ausbildung ist Photogrammetrie ein Pflichtfach. Es wird in zwei Semester-vorlesungen mit jeweils vier Semesterwochenstunden gelehrt. In Ergänzung zu den Vorlesungen müssen die Studierenden an zwei Hauptübungen teilnehmen (Abb. 16). Weiterhin müssen PhoX-Module mit zunehmender Komplexität innerhalb bestimmter Zeitvorgaben während des Semesters erfolgreich absolviert werden.

Unter anderen werden zusätzliche photogrammetrienähe Module wie „Digitale Bildverarbeitung“, „Nahbereichsphotogrammetrie“, „Ingenieurgeodäsie“, „Fernerkundung“ und „Computergrafik“ angeboten. Die Masterstudierenden können „Photogrammetrische Informationsverarbeitung“ wählen, um vertieftes Wissen in der 3D-Bildverarbeitung zu erlangen.

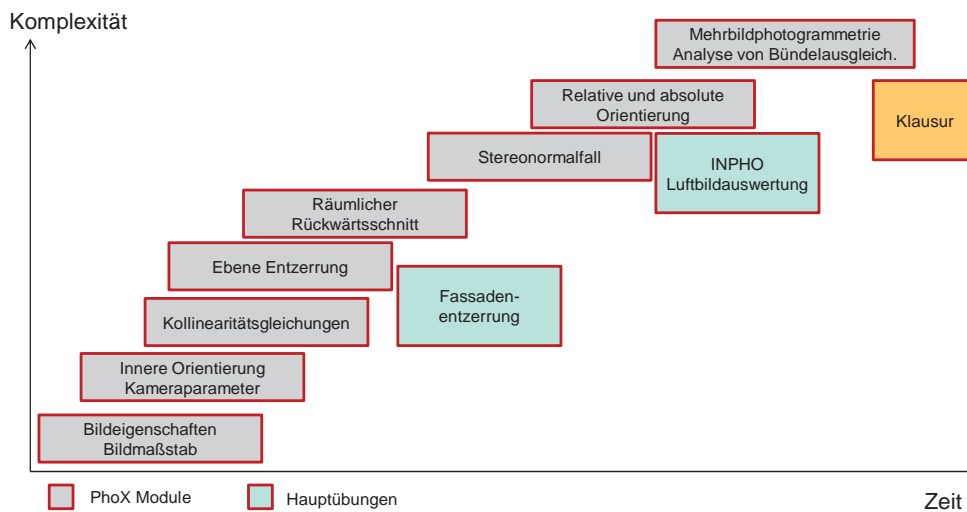


Abb. 15: Übungen im Bachelorkurs Photogrammetrie

8.2 Arbeitsablauf

Während der langen Lehrerfahrung hat der Autor ein abnehmendes Interesse der Studierenden beobachtet, Ergebnisse kritisch zu hinterfragen, eine gründliche Datenanalyse zu betreiben und sich generell mit photogrammetrischen Aufgaben zu beschäftigen.

Vor etwa drei Jahren wurde daher das Konzept in der Photogrammetrieausbildung verändert. Jetzt muss jeder Bachelor-Studierende in beiden Semestern der photogrammetrischen Lehrveranstaltungen etwa acht verschiedene Übungen mit PhoX erfolgreich abliefern (Abb. 16). Es beginnt mit einfachen Analysen von Inhalt und Eigenschaften digitaler Bilder, enthält derzeit die klassische innere Orientierung analoger Bilder, relative und absolute Orientierung, Umgang mit Kollinearitätsgleichungen, Epipolargeometrie sowie Rückwärts- und Vorwärtsschnitt.

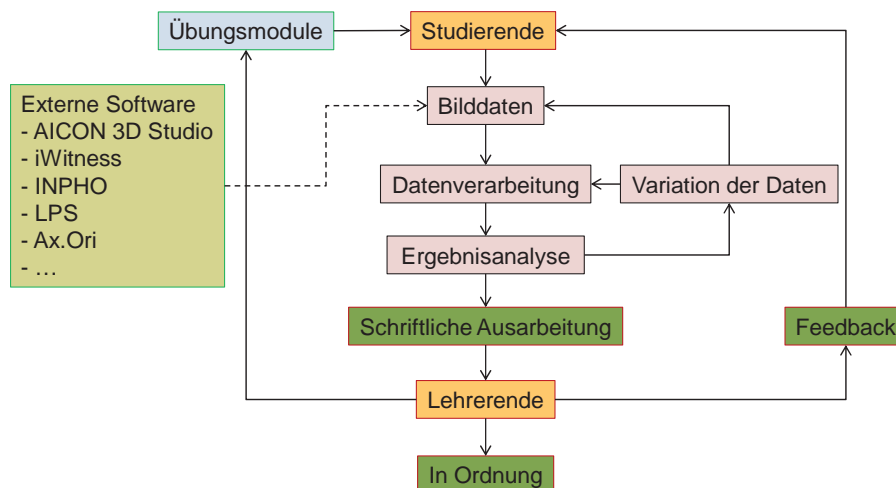


Abb. 16: Übungen im Bachelorkurs Photogrammetrie

In jeder Übung muss eine Standardaufgabe gelöst werden (Abb. 17). Anschließend müssen spezifische Versuche wie die manuelle Änderung der Eingangsdaten oder Wiederholungsmessungen

durchgeführt werden. Jeder Studierende muss eine schriftliche Übungsausarbeitung abliefern, die korrigiert und zurückgegeben wird. Ohne diese direkte Rückmeldung wäre der didaktische Erfolg gering.

Dieses neue Konzept hat dazu geführt, dass die Zahl der Studierenden, die ihre Prüfungen erfolgreich im ersten Anlauf bestehen, signifikant zugenommen hat. Als positiver Nebeneffekt kann eine bessere Beziehung zum Fach Photogrammetrie beobachtet werden.

8.3 Übungsbeispiele

Bis heute haben die Studierenden Zugang zu folgenden Übungen:

- Innere Orientierung und Pixelkoordinatentransformation
- Charakteristika von Luftbildern und Kameras
- Bildmaßstab von Luftbildern
- Ebene projektive Entzerrung
- Relative Orientierung
- Absolute Orientierung
- Kollinearitätsgleichungen
- Rückwärtsschnitt und DLT
- Stereobildverarbeitung und Epipolargeometrie
- Vorwärtsschnitt
- Analyse von Ergebnissen der Bündelausgleichung

8.4 Lehrmaterial

Für jeden Schritt verfügt PhoX über themenrelevante Hilfeinformationen. Kurzanleitungen werden im Wesentlichen für interaktive Bildoperationen bereitgestellt, während die Bedienungsanleitung eine komplette Beschreibung aller Funktionen bietet.

Vom Anwender (oder Lehrenden) können Übungsskripte editiert werden, die aus einer Sequenz vordefinierter Programmfunktionen bestehen, die der Student befolgen kann (aber nicht muss). Die gleiche Skriptsprache wird für sogenannte Assistenten genutzt, die den Anwender durch die Standardaufgaben der Photogrammetrie führen, z.B. den Ablauf einer Entzerrung.

Wenn Lehrende ihr eigenes PDF-Lehrmaterial nutzen wollen, kann dieses durch eine mit den entsprechenden Menüfunktionen verknüpfte Liste berücksichtigt werden. So können Studierende direkt innerhalb einer PhoX-Funktion auf das begleitende Lehrmaterial zugreifen.

9 Zusammenfassung

Bei der Photogrammetrie handelt es sich um ein komplexes Fachgebiet, das spezielle Anstrengungen in der Lehre erfordert. Da Studierende (und andere Anwender) sich mit Mathematik, Physik, Ausgleichsrechnung, Sensortechnologien, Bildverarbeitung, geodätischen Netzen, Genauigkeitsdefinitionen usw. auseinandersetzen müssen, sollten Lehrkonzepte die Integration verschiedenster Disziplinen in ein gemeinsames Fachgebiet vorsehen. Da Standardsoftwarepakete keine tiefgreifende Analyse von Einflüssen und Ergebnissen erlauben, versucht PhoX diese Lücke zu schließen. Die zentrale Idee von PhoX basiert auf interaktiven Prozessen und Analysen der meisten wichtigen photogrammetrischen Verfahren. Es ist weder als Werkzeug für den pro-

fessionellen Produktionseinsatz gedacht noch für eine vollständig automatisierte 3D-Objektrekonstruktion.

Allerdings ist der Einsatz einer Lernsoftware nicht sinnvoll, wenn das Lehrkonzept nicht inhaltlich darauf abgestimmt ist. Im Übrigen müssen Studierende häufig zur Durchführung von Übungen gezwungen werden, da freiwillige Angebote meistens nicht angenommen werden.

PhoX wird seit mehreren Jahren an der Jade Hochschule in Oldenburg angewandt. Es wird nicht nur Studierenden der Bachelor- und Master-Studiengänge zur Verfügung gestellt, sondern dient auch als leistungsstarkes Analysetool in Forschungsprojekten. Ferner wird es zur Ausbildung an der Kiev National University for Construction and Architecture im Rahmen einer Kooperation in Forschung und Lehre eingesetzt (KRAVCHENKO et al. 2016). Das Programm ist für akademische Zwecke kostenlos unter <https://iapg.jade-hs.de/phox/> erhältlich und wird mittlerweile in mehr als 20 Ländern eingesetzt. Es ist in deutscher und englischer Sprache erhältlich.

10 Literaturverzeichnis

- GONZALEZ-AGUILERA, D., LOPEZ, L., RODRIGUEZ-GONZALVEZ, P., GUERRERO, D., HERNANDEZ-LOPEZ, D., REMONDINO, F., MENNA, F., NOCERINO, E., TOSCHI, I., BALLABENI, A. & GAIANI, M., 2015: Development of an all-purpose free photogrammetric tool. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41(B6)**, 31-38.
- GUSSENMEYER, P., DRAP, P. & GAILLARD, G., 2002: ARPENTEUR 3.0: Recent developments in web based photogrammetry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **34(6)**, CVI.
- KRAVCHENKO, J., LUHMANN, T. & SHULTS, R., 2016: Concept and practice of teaching technical university students to modern technologies of 3D data acquisition and processing: A case study of close-range photogrammetry and terrestrial laserscanning. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41(B6)**, 65-69.
- LUHMANN, T., ROBSON, S., KYLE, S. & BOEHM, J., 2014: *Close-Range Photogrammetry and 3D Imaging*. 2nd edition, de Gruyter, Berlin, 684 p.
- LUHMANN, T., 2016: Learning photogrammetry with interactive software PhoX. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41(B6)**, 39-44.
- RIBEIRO, J.A., PIERROT-DESEILLIGNY, M., BRITO, J.L., BERNARDO FILHO, O. & MOTA, G.L., 2014: The Synergy of Open-Source Educational Software Development: The integration of the E-Foto and MicMac Software tools for Digital Photogrammetry. *Proceedings of the 3rd Open Source Geospatial Research & Education Symposium OGRS 2014*, Jolma, Ari, Sarkola, Pekka, Lehto & Lassi (eds.), 47-54.
- VYAS, A., KOENIG, G., 2014: Computer aided teaching in photogrammetry, remote sensing and geomatics – A status review. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40(6)**, 113-118.

Transformationen der antiken Stätte Bubastis/Tell Basta (Ägypten) im letzten Jahrhundert: Analyse moderner und historischer kartographischer und fernerkundlicher Quellen

TOBIAS ULLMANN¹, EVA LANGE² & ROLAND BAUMHAUER¹

Zusammenfassung: Von der prädynastischen Zeit bis zur römischen Herrschaft in Ägypten war Bubastis (Tell Basta) eine bedeutende Metropole im südöstlichen Nildelta; bekannt durch teils einzigartige antike Bauwerke. Durch die Urbanisierung sind große Teile der Stätte jedoch im Laufe des letzten Jahrhunderts überbaut worden und es liegen für viele dieser Bereiche keine, bzw. nur unvollständige Dokumentationen über das Gelände und früher Grabungen vor. Um die Transformation von Tell Basta nachzuvollziehen, wurde auf Grundlage moderner und historischer kartographischer und fernerkundlicher Quellen eine Geodatenbasis erstellt. Die Ergebnisse zeigen, dass ein Großteil des nordwestlichen Tells ist im Zuge der Ausdehnung von Zagazig verlorengegangen ist. Dementgegen zeigte die Analyse von Wilkinsons Karte (1822-1827), dass im südöstlichen und nördlichen Bereich noch heute potentiell archäologische relevante Areale zugänglich sind.

1 Einleitung

Die antike Stätte Bubastis (Tell Basta) im südöstlichen Nildelta war von der prädynastischen Zeit (ca. 3200 v. Chr.) bis zur römischen Herrschaft in Ägypten eine der wichtigsten Metropolen des Nildeltas; bedingt durch ihre vorteilhafte geographische Lage. Den Gipfelpunkt ihrer Bedeutung erlangte Bubastis als Residenzstadt der libyschen Könige der 22. Dynastie (Beginn des 1. Jt. v. Chr.). Besonders erwähnenswert ist die Existenz teils einzigartiger antiker Bauwerke, wie z. Bsp. der Große Tempel der felidengestaltigen Göttin Bastet und die Palastgebäude und Nekropolen des Alten und Mittleren Reiches (LANGE 2008; BIETAK & LANGE 2014). Heute liegt das (Grabungs-)Gelände am südöstlichen Stadtrand von Zagazig, der Hauptstadt der ägyptischen Provinz Sharqiyya. Durch die moderne Urbanisierung sind bereits große Teile der nördlichen und südlichen antiken Stätte überbaut worden und somit heute unzugänglich. Für viele dieser Bereiche liegen keine, bzw. nur unvollständige Dokumentationen durch frühe Grabungen vor. Zudem sind große Bereiche des noch zugänglichen südöstlichen Tells bisher kaum erforscht. Nach den Berichten des Herodot (Hist. II, 138; ca. 450 v. Chr.) lagen hier der Tempel des Merkur, der Dromos und das Stadtzentrum selbst (LANGE & ULLMANN 2015). Im Rahmen eines am Lehrstuhl für Ägyptologie an der Universität Würzburg angesiedelten deutsch-ägyptischen Grabungsprojektes wird seit mehreren Jahren die archäologische Erforschung des antiken Bubastis betrieben. Zur Unterstützung dieser Arbeiten wurde eine Geodatenbasis zum antiken Ausmaß und der funktionalen Organisation der Stadt, unter Nutzung von historischen Karten und Luftbildern, aufgebaut. Ziel der Analyse ist es die neuzeitlichen Transformation von Bubastis

¹ Lehrstuhl für Physische Geographie, Institut für Geographie und Geologie, Universität Würzburg, Am Hubland, D-97074 Würzburg; E-Mail: tobias.ullmann@uni-wuerzburg.de

² Lehrstuhl für Ägyptologie, Institut für Altertumswissenschaften, Universität Würzburg, Residenzplatz 2, Tor A, D-97070 Würzburg; E-Mail: eva.lange@uni-wuerzburg.de

durch die Integration aller verfügbaren Daten in ein geographisches Informationssystem (GIS) nachzuvollziehen und zudem mögliche Hinweise auf noch nicht bekannte Siedlungsbereiche zu sammeln (LANGE et al. 2016).

2 Datenbasis und Datenverarbeitung

Zu diesem Zweck wurde eine Recherche zu verfügbaren modernen und historischen kartographischen und photographischen, bzw. fernerkundlichen, Quellen durchgeführt. Zudem wurden Skizzen und Pläne der Ausgrabungen unter Leitung von El-Sawi und Bakr aus den 1970er und 1980er Jahren aufgearbeitet (EL-SAWI 1977; EL-SAWI 1979; BAKR 1982; BAKR 1992). Tab. 1 zeigt die für die Analyse zur Verfügung stehenden Daten. Die jüngste Quelle zu Tell Basta stand durch ein hochaufgelöstes Satellitenbild von WorldView-2 über die ESRI Basemap von ArcMap 10.3 zur Verfügung. Der Datensatz zeigt das gesamte Stadtgebiet von Zagazig mit einer räumlichen Auflösung von ca. 0,5 m. Ein Abgleich mit im Feld aufgenommenen Passpunkten, bzw. *Ground Control Points* (GCP), zeigte eine hohe räumliche Lagegenauigkeit des Datennetzes und seine Eignung für weitere räumliche Analysen. Mit Hilfe der Grabungsdokumentation von El-Sawi und Bakr wurde für das Jahr 1980 eine detaillierte Fundkarte des nördlichen Tells erstellt. Für den Zeitschnitt 1969 standen hochaufgelöste Satellitenbilder der US-amerikanischen CORONA Mission zur Verfügung. Der panchromatische Datensatz besitzt etwa eine Auflösung von 4,0 m und war für die Abgrenzung des damaligen Geländes von Tell Basta geeignet.

Tab. 1: Fernerkundliche und kartographische Quellen

Name	Zeitpunkt der Aufnahme	Typus	Räumliche Auflösung/ Maßstab	Quelle
ESRI Basemap (WorldView-2)	2012	Digitales RGB-Satellitenbild	ca. 0,5 m	ESRI Basemap Imagery in ArcMap 10.3
Grabungsdokumentation von El-Sawi und Bakr	ca. 1980	Handzeichnung	n/a	BAKR 1992; EL-SAWI, 1979
CORONA Imagery	1969	Panchromatisches Satellitenbild	ca. 4,0 m	U.S. National Archives in 2015
WW II Imagery	1942-1943	Panchromatisches Luftbild	ca. 1,0 m	Luftbilddatenbank Estenfeld 2015
Survey of Egypt Map (SoE)	ca. 1935	Topographische Karte	1:25.000	Survey of Egypt 1935; Zur Verfügung gestellt von Dr. R. Schiestl (LMU) and Prof. Dr. S. Seidlmayer (DAI) in 2015
Wilkinson's Map of Bubastis	1822-1827	Handzeichnung/ Skizze	n/a	WILKINSON 1843; LÉCLERE 2008

Durch eine Recherche der *Luftbilddatenbank Estenfeld* (Deutschland) unter Leitung von Dipl.-Geogr. Christian Büdel (Universität Würzburg, Deutschland) standen für das Jahr 1943 hochaufgelöste panchromatische Luftbilder vom südlichen Stadtrand von Zagazig und von Tell Basta zur Verfügung. Diese Aufnahmen wurden von deutschen Aufklärern während des Zweiten Weltkriegs aufgezeichnet. Der Datensatz besitzt in etwa eine räumliche Auflösung von 1,0 m und eignete sich für eine detaillierte Analyse der sichtbaren Oberflächenstrukturen. Für das Jahr 1935 wurden durch Dr. Robert Schiestl (LMU - Ludwig-Maximilians Universität München) und Prof. Dr. Stephan Seidlmayer (DAI - Deutsches Archäologisches Institut) topographische Karten der Landvermessung (SoE – Survey of Egypt) im Maßstab 1:25.000 zur Verfügung gestellt. Die älteste kartographische Darstellung von Bubastis stand durch die Aufzeichnungen von Sir J. G. Wilkinson zur Verfügung (WILKINSON 1843; LÉCLERE 2008). Diese Skizze wurde zwischen den Jahren 1822 und 1827 angefertigt und zeigt den damaligen Umfang von Bubastis, sowie weitere Anmerkung Wilkinsons zur damaligen Gestalt des Geländes. Diese Daten wurden digitalisiert, in das GIS integriert und manuell auf Basis von eindeutig zuordenbaren Passpunkten georeferenziert. Als Grundlage für die Georeferenzierung diente für alle anderen Daten dabei das Satellitenbild von WorldView-2. Sicher zuordenbare Passpunkte (GCPs) fanden sich hierbei vor allem entlang von Straßenzügen, Feldgrenzen und Kanälen, welche in vielen Fällen eine hohe zeitliche Persistenz zeigten. Für alle Datensätze (1935-1980) wurde so eine ausreichend hohe räumliche Lagegenauigkeit von unter 15 m (RMSE – Root Mean Square Error) erzielt. Die Georeferenzierung von Wilkinsons Karte zeigt sich dabei als problematisch denn die wenigen von Wilkinson eingezeichneten Strukturen erschwerten eine genaue räumliche Zuordnung. Mit Hilfe des von Wilkinson verfassten Berichts sowie der dokumentierten Längenangaben und der verzeichneten Lokalität und Orientierung des Großen Tempels der Bastet gelang dennoch eine belastbare Georeferenzierung. Nach dem Aufbau dieser Geodatenbasis wurden anschließend erkennbare Oberflächenstrukturen und der Umfang des Tells im zeitlichen Verlauf digitalisiert und im Kontext der bisher bekannten Besiedelungsgeschichte interpretiert.

3 Ergebnisse

Das Ergebnis dieser Arbeitsschritte ist in Abb.1 dargestellt. Diese zeigt den Umfang von Tell Basta in den Jahren 1935, 1943, 1969 und 2012, sowie die oberflächennahe Strukturen und Siedlungsrelikte für das 2012. Erkennbar ist, dass insbesondere im nördlichen und südlichen Teil große Flächen des ursprünglichen Geländes durch die Urbanisierung verlorengegangen sind. Zwischen 1935 und 1943 entstand im südlichen Bereich von Tell Basta eine erste moderne Überbauung, die auch heute noch existiert (Industrieanlagen und Deponie). Der Großteil des nördlichen Tells wurde dagegen erst später, zwischen den Jahren 1943 bis 1969 und 1969 bis 2012 überbaut. Hier befanden sich nach den Grabungsergebnissen von El-Sawi (1968-1970) der Katzenfriedhof und der Teti-Tempel. Dementgegen zeigten der östliche Bereich des Tells und der Bereich in der Nähe der Palastanlagen im neuzeitlichen Verlauf nur geringen Änderungen. In allen fernerkundlichen und kartographischen Quellen ist der an den Tell im Osten angrenzende Bereich einer landwirtschaftlichen Nutzung zuzuordnen. Die Kartierung der im WorldView-2 Datensatz erkennbaren Strukturen zeigt neben den Planquadraten der laufenden Ausgrabungen auch oberflächliche Siedlungsrelikte.

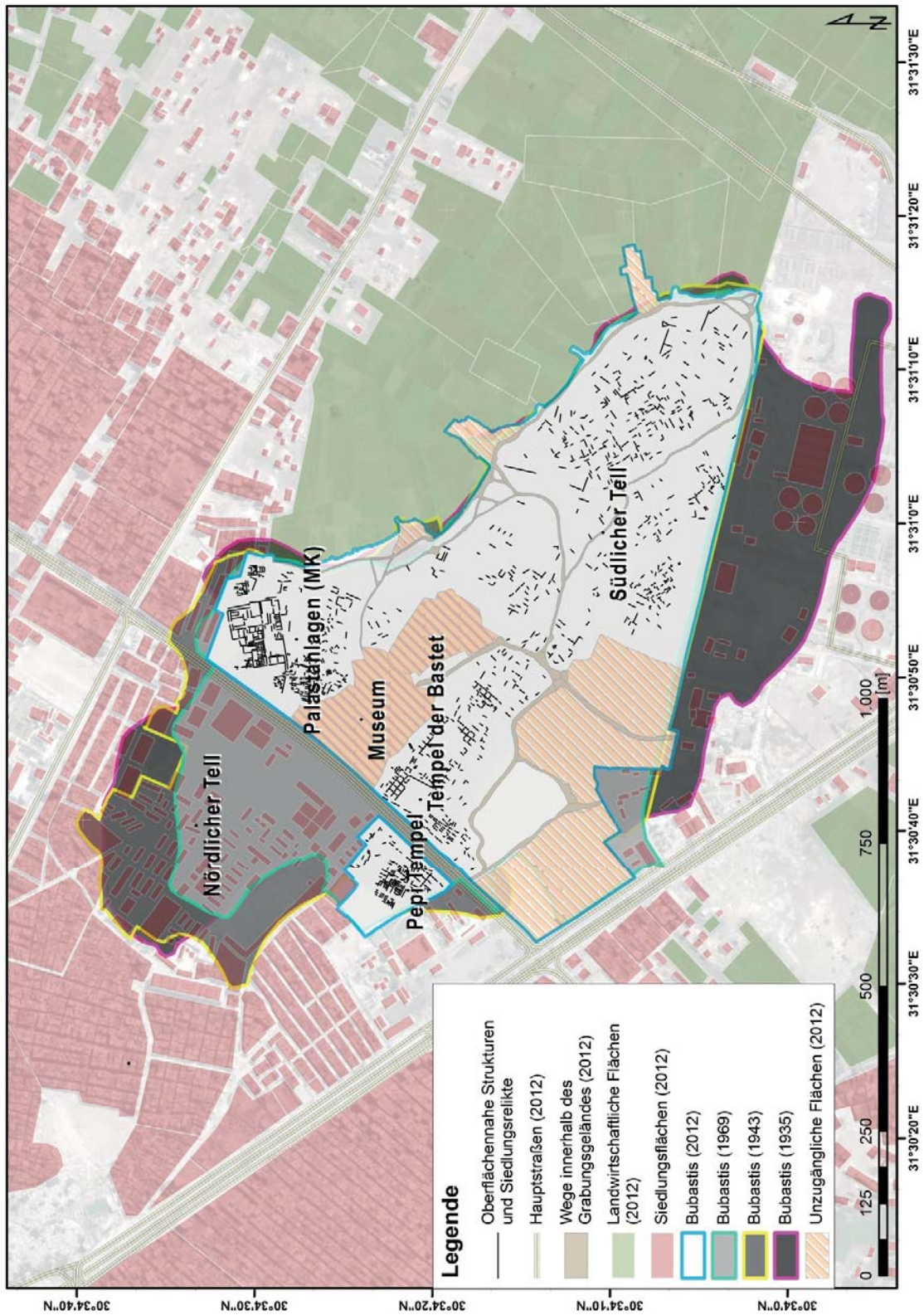


Abb. 1: Umfang von Tell Basta in den Jahren 1935, 1943, 1969 und 2012 und oberflächennahe Strukturen und Siedlungsrelikte (2012).

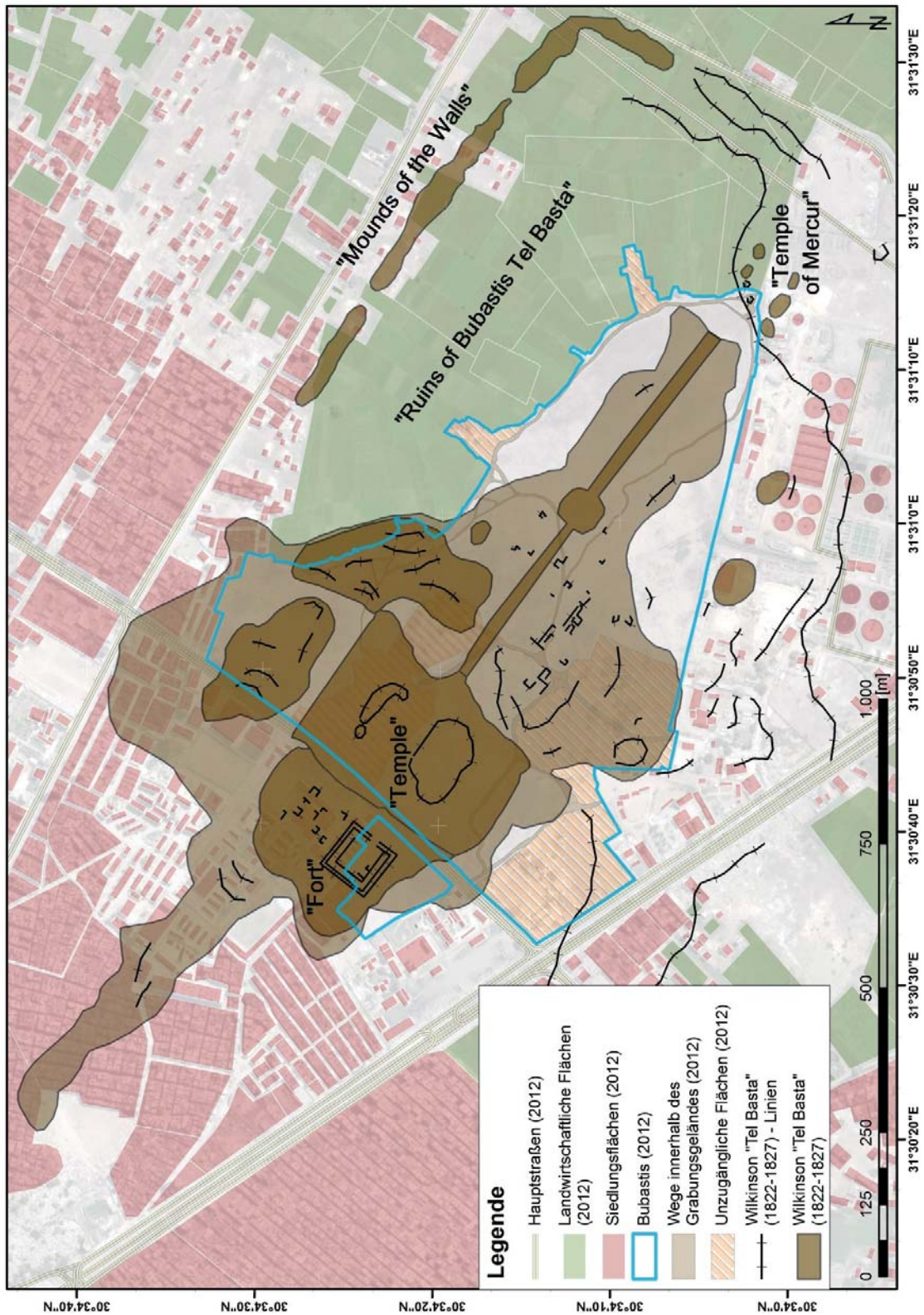


Abb. 2: Überlagerung der digitalisierten und georeferenzierten Karte von Wilkinson (1822-1827) mit dem heutigen Umfang von Tell Basta und den Siedlungsflächen von Zagazig (2012).

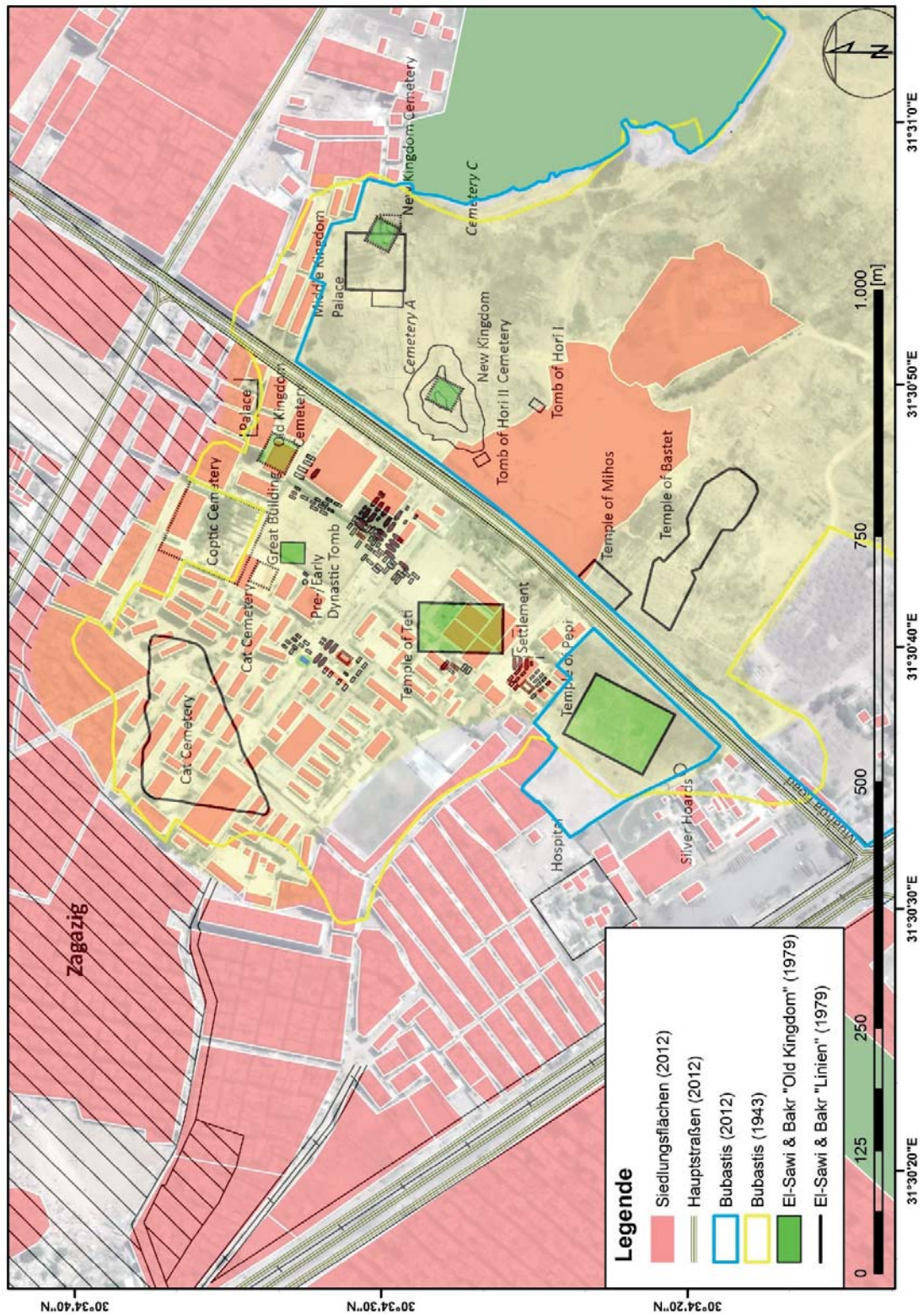


Abb. 3: Siedlungsflächen von Zagazig (2012), Umfang des nördlichen Teils (1943 & 2012) und digitalisierte und georeferenzierte Dokumentation der Grabungen von El-Sawi und Bakr (1979).

Insbesondere im weitgehend unerforschten südlichen Tell wurde so eine vom Palastbereich und Tempelbereich abweichende Orientierung der Straßenzüge und Bauwerke sichtbar; höchstwahrscheinlich handelt es sich hierbei um Relikte der römischen Besiedelung.

Die Überlagerung der digitalisierten und georeferenzierten Karte von Wilkinson (1822-1827) mit dem heutigen Umfang von Tell Basta und den Siedlungsflächen von Zagazig im Jahr 2012 ist in Abb. 2 dargestellt. Es zeigt sich, dass sich der Tell zur Zeit von Wilkinsons Aufenthalt in Bubastis wesentlich weiter nach Norden und Süden erstreckte. Die Zeichnungen lassen zudem darauf schließen, dass der an den Osten des Tells angrenzende Bereich von einer wallartigen Erhebungen begrenzt war; d.h. bereits zur Zeit von Wilkinson scheint das Gelände im Osten des Tells tiefer gelegen zu haben – der Sakral- und der Nekropolenbereich war vermutlich topographisch höherliegend und somit vom Profanbereich getrennt. Der Vermerk von Wilkinson („Ruins of Bubastis Tel Basta“) deutet des Weiteren daraufhin, dass hier zum Erstellungszeitpunkt der Skizze deutlich sichtbare Strukturen der früheren antiken Stätte oberflächennah existierten. In den fernerkundlichen Bilddaten (1943 - 2012) sind jedoch keine Hinweise auf diese Strukturen erkennbar und das Gelände unterliegt heute der landwirtschaftlichen Nutzung, bzw. in zunehmendem Maß der Urbanisierung von Zagazig. Weiter von Bedeutung ist zudem die Einzeichnung von Wilkinson im südlichen Bereich des Tells; die eingezeichnete Linie stimmt ungefähr mit der südlichen Grenze des Tells im Jahr 1935 überein und markiert somit möglicherweise die südliche Grenze der antiken Stätte. Die von Wilkinson gezeichnete Hauptachse (der „Dromos“), der Marktplatz und der Tempel des Merkur im südlichen Bereich des Tells sind nach heutigem Kenntnisstand in ihrer genauen Lage nicht bekannt; der Großteil des südlichen Geländes ist noch nicht systematisch erforscht. Diese Strukturen sind ebenfalls in keiner der anderen kartographischen oder fernerkundlichen Quellen (1935 - 2012) ersichtlich. Auch die im Norden des Tells verzeichneten Strukturen und das „Fort“ sind nach heutigem Kenntnisstand ihrer genauen Lage nach nicht bekannt; möglicherweise deutete Wilkinson die Umfassung des Pepi-Tempel als das „Fort“.

Die Siedlungsflächen von Zagazig im Jahr 2012, der Umfang des nördlichen Tells in den Jahren 1943 und 2012, sowie die digitalisierte und georeferenzierte Dokumentation der Grabungen von El-Sawi und Bakr aus dem Jahr 1979 ist in Abb. 3 abgebildet. Diese Überlagerung zeigt deutlich dass ein großer Teil von bedeutenden Fundplätzen durch die moderne Urbanisierung verlorengegangen ist: Insbesondere der Katzenfriedhof, der Teti-Tempel und der Friedhof des Alten Reiches sind heute unzugänglich.

4 Fazit & Ausblick

Die Analysen der fernerkundlichen und kartographischen Quellen zeigen auf, dass ein Großteil der antiken Stadt Bubastis im Zuge der Urbanisierung von Zagazig verlorengegangen ist; die noch im Jahr 1935 zugängliche Fläche war bedeutend größer, insbesondere im nordwestlichen Teil des Tells. Die von Bakr und El-Sawi ergrabenen Siedlungsrelikte (insbesondere der Katzenfriedhof und der Teti-Tempel) im Bereich des nördlichen Tells sind heute aufgrund der modernen Bebauung nicht mehr erschließbar; lediglich der Pepi-Tempel ist heute noch zugänglich. Durch den Übertrag in das GIS lassen sich die früheren Lokalitäten jedoch räumlich zuordnen und sie sind somit, zumindest bezüglich ihrer Lage, dokumentiert. Dies wird für

künftige Studien zur räumlichen Organisation der verschiedenen antiken Stadtbereiche wertvolle Beiträge leisten. Die Analysen zeigten zudem, dass der in Palastnähe befindliche Bereich von Bubastis im hier untersuchten neuzeitlichen Verlauf nur geringen Änderungen unterworfen war. Nach Wilkinsons Karte (1822-1827) sind der südöstliche Bereich (von Wilkinson als „Ruins of Bubastis“ bezeichnet) und die nördliche Umfassung („Mounds of the Walls“) heute noch zugänglich und somit potentielles Ziel zukünftiger Grabungen. Diese Lokalitäten unterliegen heutzutage, ebenso wie der von Wilkinson als „Temple of Mercur“ bezeichnete Bereich, einer landwirtschaftlichen Nutzung und sind daher größtenteils zumindest noch nicht überbaut.

Die Anlage der antiken Siedlung scheint zudem auf die heutige Stadtstruktur nachzuwirken; die Zufahrtsstraße und der östliche Kanal zeigen die gleiche Orientierung wie die frühere Umfassung von Bubastis nach Wilkinsons Karte. In allen Zeitschnitten zeigte sich zudem, dass die Form des Siedlungsareals von Bubastis länglich gestreckt ist und in nordwestlicher/südöstlicher Richtung orientiert ist. Dies könnte ein Hinweis auf die geomorphologische Position des Siedlungsplatzes sein; eine Gezira, welche über dem damaligen Hochwasserniveau lag.

5 Literaturverzeichnis

- BAKR, M. I., 1982: New excavations of Zagazig University, 153-167. *L'Égyptologie en 1979: axes prioritaires de recherches* **1**, Anonymous (ed.), Paris.
- BAKR, M. I., 1992: Tell Basta I: Tombs and Burial Customs at Bubastis. The Area of the so-called Western Cemetery, Cairo.
- BIETAK, M. & LANGE, E., 2014: Tell Basta: The palace of the Middle Kingdom. *Egyptian Archaeology* **44**, 4-7.
- EL-SAWI, A., 1977: Preliminary report on Tell Basta excavations. Seasons 1969, 1970, 1971, *ZÄS* **104**, 127-131.
- EL-SAWI, A., 1979: Excavations at Tell Basta. Reports of Seasons 1967–1971 and Catalogue of Finds. Charles University of Prague.
- LANGE, E., 2008: Legitimation und Herrschaft in der Libyerzeit: eine neue Inschrift Osorkons I. aus Bubastis (Tell Basta). *ZÄS* **135**, 131-141.
- LANGE, E. & ULLMANN, T., 2015: Goddess on the water: the sacred landscape of Bubastis. *Egyptian Archaeology* **47**, 17-19.
- LANGE, E., ULLMANN, T. & BAUMHAUER, R., 2016: Remote Sensing in the Nile Delta: Spatio-Temporal Analysis of Bubastis / Tell Basta. *Egypt and the Levant* **26**, International Journal for Egyptian Archaeology and Related Disciplines, M. BIETAK (Hrsg.), 377-392.
- LECLERE, F., 2008: Les villes de Basse Égypte au Ier millénaire av. J.-C.: analyse archéologique et historique de la topographie urbaine. *Publications de l'Institut français d'archéologie orientale du Caire* **2**, BdÉ 144/2.
- WILKINSON, J.G., 1843: *Modern Egypt and Thebes: being a description of Egypt, including the information required for travelers in that country*, London.

Concept of a Method for Change Detection in Topobathymetric Point Clouds

RICHARD BOERNER¹, LUDWIG HOEGNER¹ & UWE STILLA¹

Abstract: Change detection of 3D point clouds is addressed in multiple research projects. This paper introduces a concept for change detection in topobathymetric point clouds based on topobathymetric full waveform laser scanner data. Airborne laser data with a green laser not only includes hard targets or volumes, given e.g. by vegetation, but also points from the water surface, underwater inhomogeneity and river ground. Challenge of the intended work is to detect changes in waterbeds, in the water level or calculate the morphodynamic of rivers. The change detection in topobathymetric point clouds requires pre-processing like the co-registration of multi temporal point clouds, segmentation and classification. A concept for the processing chain is proposed. First results of the segmentation by region growing are shown of the test site of the river Mangfall in Bavaria, Germany.

1 Introduction

Bathymetry is the measuring of underwater depth of seas, oceans or other large bodies of water. In general acoustic and optical sensors can be used in which latter can be distinguished into active and passive techniques. A passive way to determine the water depth is by colour bathymetry. The water depth calculation depends on the water colour in an image (DELAI et al. 2014). An active case of bathymetry is the airborne laser bathymetry (ALB). For that, a green laser mounted on an airplane, is used to measure points under the water surface. This was first used to detect submarines from the air in the 1960's. Additionally to the green laser, infrared lasers are used in combination (PFENNIGBAUER et al. 2010; STEINBACHER et al. 2012). While the infrared laser reflects on the water surface and measures the surface, the green laser penetrates the water and measures the water ground. To ensure the water penetration, the laser beam have an incidence angle of approximately 20 degrees to the water surface. Measuring morphodynamic changes after flood events and evaluating the impact of this changes on instream habitats based on topobathymetric point clouds from a combined infrared and green laser (MANDLBURGER et al. 2015). This kind of river monitoring is needed to satisfy the aims of the European Directive 2000/60/EC (EUROPEAN UNION 2000) and create flood risk maps, required by the European Directive 2007/60/EC (EUROPEAN UNION 2007).

A topobathymetric point cloud is a set of points mapping the water surface, the river ground and the river waterfront. Changes of water levels and the river beds can be captured by point clouds taken from sub sequent flights. The search for new or missing points is the task of the change detection. New points can be produced by new buildings or deposits on the ground or waterbed. Also a growing tree produces new points in another scan time. In contrast to that, the drift of water ground points, landslides, demolished buildings or fallen trees are determined by missing points

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

in the second scan time. So it can be used to recognise movements in the river ground points or to determine and control the water level. The challenge from such a change detection is, that the discrete position measuring of the laser scanner typically differ between two times. Furthermore, occlusions have to be considered and change detection of vegetation is a challenging task due to the volumetric reflections and changing canopy properties. Therefore, an accurate co-registration of the two point clouds to the two time stamps is needed for the change detection.

The co-registration of point clouds has the aim to create transformation parameters from one point cloud to the other. With this transformation it is possible to adjust the offset and orientation misalignments between the point clouds, caused by the measuring inaccuracy. The adjusted transformation parameters are further used to assign one point in the first point cloud to the corresponding point in the other point cloud. For the registration it is important to know which class the points belong to. By knowing the class of potential matches, mismatches, like ground to building points, are avoided. Furthermore, the class information can be used for the change detection by defining probabilities of changes. For example, vegetation points change with a higher probability as ground points.

Chapter 2 show related works in the fields of full waveform analysis, point cloud classification, registration and change detection. The proposed concept and its processing chain is shown in chapter 3. An overview of the used test site is given in chapter 4 and the last chapter shows the first results.

2 Related work

2.1 Full Waveform

Laser scanners rare divided into the ones, which recognises only the time stamps to each returning impulse and the ones which recognises the whole returning signal. This so-called full waveform must be analysed to derive discrete points. JUTZI & STILLA (2006) show how to separate the full waveform signal into multiple returning pulses along the laser beam. For this, the sent waveform is used to find reflections of scatters. Additional attributes, like the number of returns are derived from the entire signal and assigned to each point. This information can be used to determine multiple classes, including single, first, last and middle pulses (REITBERGER et al. 2008). Furthermore, the returning signal can be used for a vegetation classification (RUTZINGER et al. 2008) or to segment tree crowns and tree stems (REITBERGER et al. 2009).

2.2 Classification

The point cloud classification can be divide into two main parts. The first part is the segmentation and the second one the matching of class labels to the segmented regions. The segmentation aims for the creation of regions of connected points. Most of these segmentation approaches are point based (RUTZINGER et al. 2008; SHAPOVALOV et al. 2010; ZHANG et al. 2016). For that, k nearest neighbours are searched and connected to a graph. To build the segments from this graph, there are methods for a region growing, which combine points, for example on the base of their echo width (RUTZINGER et al. 2008). Or a graph cut at the local maxima of the points height values is used (ZHANG et al. 2016). Another base for such segmentation are volume clusters, so called voxels. These voxels have a defined neighbourhood and are segmented together, if they are close

in position and their intensity or colour difference is below a threshold (AIJAZI et al. 2013). To match a segment to a specific class, a descriptor has to be calculated. These descriptors are used to define the properties of a specific class. In the case of voxels, the descriptor can consist of the surface normal or the geometry of the segment. Other additional descriptor attributes are the height differences between the geometrical centre and the barycentre. For non-geometric attributes, intensity or colour values can be chosen (AIJAZI et al. 2013). In the point wise classification, machine learning approaches were used. ZHANG et al. (2016) uses trained SCLDA-based features as their descriptors. Furthermore, they apply an AdaBoost classifier to match the features to a class. Another machine learning approach is to use Markov random fields, like in SHAPOVLAV et al. (2010). For this Markov random field, they need a unary potential, which is the descriptor for each single point and a pairwise potential, which is the descriptor for an edge in the graph. For the unary potential spectral and directional features, spin images, angular spin images and distributions of heights are used. The angle between normals, the difference in altitudes and the difference in positions are their pairwise potential. NIEMEYER et al. (2014) uses conditional random fields instead of Markov random fields. They use linear models and the Random Forest algorithm to train and obtain their unary and pairwise potentials. These potentials are calculated on the base of intensity, echo numbers, height above a digital terrain model (DTM), point densities and some other feature attributes. RUTZINGER et al. (2008) uses a classification tree based on the echo width, echo ratio and dense ratio of the segments. But they have only the aim to detect vegetation points, so the classification tree isn't very deep. A detailed segmentation and tree species classification is shown in REITBERGER et al. (2008) and a segmentation based on graph cuts in REITBERGER et al. (2009).

2.3 Registration

The extraction of transformation parameters from one point cloud to another, is the aim of the registration task. Existing approaches are summarized in a review by SALVI et al. (2006) and TAM et al. (2013). The registration methods can be divided into coarse and fine registration (SALVI et al. 2006). A simple method to register surfaces are point signatures, which are calculated from the intersection of a sphere, centred at point P, to the local surface of P of each point (CHUA & JARVIS 1997). Another way is to use a principal component analysis (PCA) based approach, which looks for corresponding principal axis of the covariance matrix of neighbouring points (CHUNG et al. 1997). JOHNSON (1997) introduces spin images, which accumulate each point of a surface into a tangent plane to a specific surface point. Therefore, the distance from the normal vector of the specific surface point to the other surface points and the distance from these surface points to the tangent plane, are used as the axis of the spin image. TAREL et al. (1998) uses a polynomial representation of a surface function to model surfaces and matches these functions. BRUNNSTRÖM & STODDARD (1996) have shown, how to build a genetic algorithm for the registration task. This chooses points on the surface randomly and defines the best match by a fitness function. This fitness function depends on angles between normal vectors and the distance between two points. The most common fine registration algorithm is the iterative closest point (ICP) algorithm (BESL & MCKAY 1992). This algorithm starts with a seed estimation of the transformation parameters. After this initial transformation the algorithm searches for the best transformation parameters by minimizing the distances between the point correspondences iteratively. Another method to model

surfaces is to use signed distance fields (MASUDA 2001). The distance from a point to the surface and the surface normal approximates a so called signed distance field (SDF) of the surface to the closest point. Two surfaces are matched together by minimizing the distance between two SDFs. With the focus on the surface type, the registration methods can be divide into rigid and non-rigid (TAM et al. 2013). In the special case of airborne laser scanning data, a rigid PCA based approach creates good results (HEBEL & STILLA 2009; HEBEL & STILLA 2012). Here, roofs were used as rigid surfaces and enables also a calibration of the airborne laser scanning (ALS) system.

2.4 Change Detection

A change detection can be implemented either point based or voxel based. The idea for the point wise change detection is to subtract meshes. This can be for example a digital surface model (DSM) (MURAKAMI et al. 1999). Here, the difference between the elevation of two DSMs is projected to a change image. This is used to detect building changes in ALS data. Another approach which subtract digital elevation models (DEMs) is shown by MANDLBURGER et al. (2015). To differentiate true changes from measuring noise, a statistical test is used to detect significantly changes. This is used to model the morphodynamic of a river. A voxel structure and laser beam reconstruction can be used for change detection as well (HEBEL et al. 2013). New points are detected, if the laser beam traverses occupied voxels in the reference data. If the laser beam comes to lie in an empty voxel in the reference data, the traversed voxels of the beam are marked as free areas. Another method is to compare distances and plane intersection angles between two voxels of two point clouds (GIRADEAU-MONTAUT et al. 2015). This detects not only changed points, it also calculates how far the points have moved. A Hausdorff distance performs best results in this context. The Hausdorff distance is the distance from one point to its closest point. To get this GIRADEAU-MONTAUT et al. (2015) calculate the distance from one point to all others in neighbouring voxels and search for the shortest distance.

3 Concept

This paper proposes a concept for a change detection. This consist mainly of three parts. The first part is a classification of the point cloud data. In the second part two point clouds from different times will be registered, using the results of the classification. And the third part is the change detection itself, using the results from the classification. All three parts are implement on the base of a voxel structure. Since the data is very big (up to 200 Mio points) it is not suitable to build a graph, based on all points. Instead of this, a voxel space is used as a generalisation of the point cloud. The concept scheme is shown in Fig. 1.

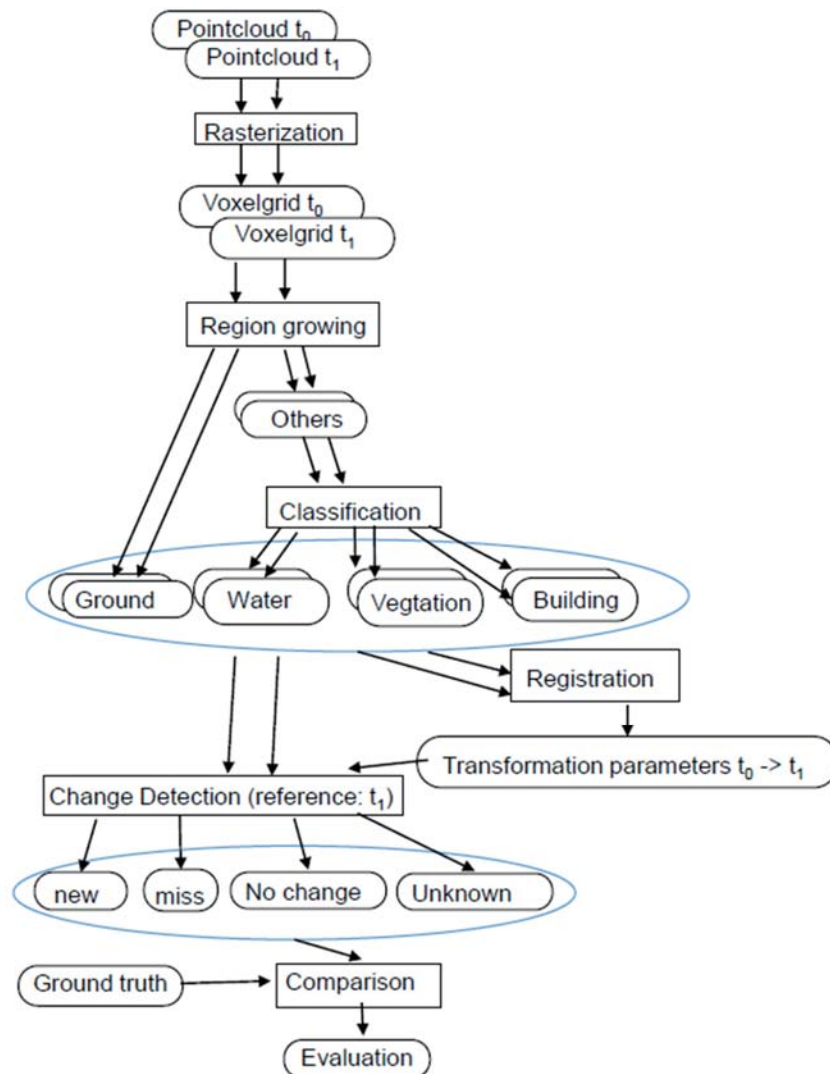


Fig. 1: The proposed concept. The rectangles show the processing steps which create the output in the ellipses. These outputs are used for further processing, which is shown by the arrays. These arrays point from one output to the processing step which needs these output.

The classification segments the point cloud into ground, water, vegetation and building points. These labels are further used in the registration to avoid mismatches, for example from ground points to building points. Furthermore, a change in vegetation points is much more probable as a change in ground points. So the class labels must also be used in the change detection part. To classify the point cloud, the cloud has to be segmented first. For this segmentation, neighbouring voxels can be put together on the basis of attributes, like the intensity difference (AJAZI et al. 2013). In this way a region growing algorithm on the basis of voxels creates segments of points with same conditions. The voxel classification presented in AJAZI et al. (2013) uses a linear match of the calculated descriptor to the class labels. The proposed method will use a machine learning algorithm for this

task. The advantage of the machine learning algorithm is, that they can be easily adjusted to new classes. So the classification method can be adjusted to another classification application in the future. Possible machine learning algorithm for a match of a single descriptor to a class label are Random Forest or support vector machines.

The registration of natural scenes is a challenging task as plane structures which are easy to co register mainly appear only in urban areas. In the case of voxel structures the concept introduces a combined coarse and fine registration. The coarse registration uses bigger voxels, so that they don't change between two times (if the landscape doesn't change drastically). An iterative algorithm, which uses iteratively smaller voxels, solves the fine registration. In addition to the geometrical shape of the voxel structure, the class information of each segment, and therefore the voxel, is used to roughly filter match-able and not match-able voxels. Therefore, only voxels of the same class are considered in the pairwise matching process. Furthermore, a hybrid data structure is used to model ground points as a DTM and register two DTMs. Volumetric parts like water and vegetation, are modelled as a voxel structure and handled in a different way for the change detection.

The change detection, proposed by HEBEL et al. (2013) is suitable for voxel structures and detects missing points as well as new points. But changes inside one voxel are not detected, so there has to be some fine adjustments. Dependent on the voxel size the ground points can change in a sub voxel manner. Therefore, a conversion from the voxel space to the point space is needed in the change detection, like in GIRADEAU-MONTAUT et al. (2015). But the calculation of distances from one point to all other inside a voxel is very computational expensive. To fuse the voxel and point wise change detection approaches, there is a first change detection on voxel base. This is done by a ray tracing of the laser Beam (HEBEL et al. 2013) or by defining attributes, which are checked for a change. A significant change in density for example can mark therefore changed voxels. These changed voxels are further compared with a point wise change detection. The implementation of the point wise change detection is dependent on the point type. Therefore, solid points can be compared with another method as water or vegetation points. The calculation of differences in DEMs performs a suitable change detection for solid points. So if the changed voxel consists of solid points, the next step is to calculate the difference of the DEMs for this points. For water points, the point wise change can be calculated with the distance of two fitted planes to the water surface. Vegetation points can further be checked of the whole bounding box difference. So that the point wise change detection calculates volume changes. Thresholds, defining a significant change, depend on two parameters. One parameter for each class, so that the threshold for vegetation points is in general higher, than for others. The second parameter for the threshold is a global one, which considers measurement accuracy and misalignments from the registration before.

4 Experiments

First algorithmic tests of the proposed method are performed on a dataset of the Mangfall area in Bavaria, Germany (Fig. 2), recorded with a VQ 820G Scanner. The dataset reaches from the Tegernsee to the highway A8. The whole scan for the first flight consist of nearly 200 Mio points and maps a flight strip of nearly 17 km length and 500 m width. On the ground, the point distance in a single flight strip is 4 dm. A water penetration up to 5 m was reached with the laser beam. The

laser scanner measures with a wavelength of about 532 nm. By a flight height of 500m, the laser beam gets a 50 cm footprint. Furthermore, the scanner is able to create a high spatial resolution of 20 to 50 points/m². A rotating multi facet mirror, used as the scan mechanism, ensures an incidence angle of 20° with respect to the flight direction, with an accuracy up to 1° (STEINBACHER et al. 2012). The scanning system uses a full waveform approach, which recognises up to 4 pulses inside a single laser beam. Recognised point attributes are the intensity, the number of returns, the return number, the scan angle rank, the GPS time and two flags which determine the scan direction and flight line edge points.

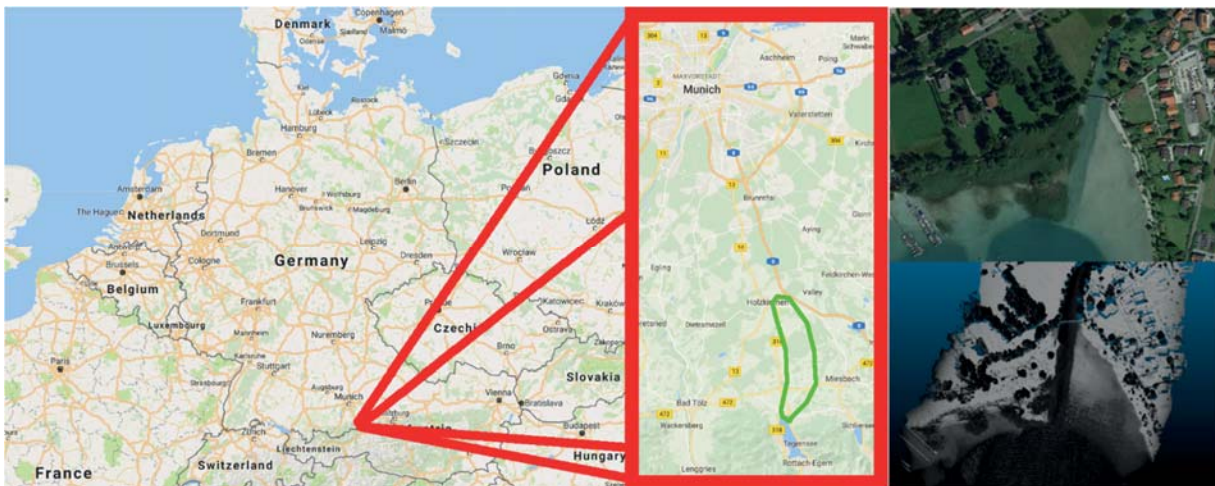


Fig. 2: The scanned region (Photos from Google maps). Left: In the Background: whole Germany and parts of Europe. In the zoomed part (red rectangle): the scanned Mangfall area (marked in green) from the Tegernsee in the south to the highway in the north. Top right: clip from the Tegernsee of the scanned area. Bottom right: scanned data with the point's intensity as grey value.

5 Discussion

This paper has summarized currently existing methods for point cloud processing. This paper gives a short overview of existing methods for point cloud processing. A more detailed overview can be found in WEINMANN (2016). These methods include full waveform analysis, classification, registration and change detection tasks. In the special task of river surveying, bathymetric applications and some laser scanner devices are also proposed. Since the European directive 2007/60/ EC demands the creation of change maps, the change detection task with bathymetric data becomes more important. Currently used methods for point cloud classification and registration mostly use a point wise algorithm. Such a point wise processing can become unnecessary computational expensive with very big point clouds (up to 200 Mio points). Therefore, a voxel structure should be used to generalize such big point clouds in a better way. A very difficult task is the registration in non-urban areas. There exist no rigid structures or shapes which can be used for the registration task. Furthermore, also the registration should be handled with a voxel structure to get big point cloud processable.

In this context, the paper introduced the idea of defining a voxel structure and use this structure for all processing tasks. In future work, the authors will introduce these voxel structure and approaches for classification, registration and change detection on the base of voxels. The special interest will be the change detection of topobathymetric point clouds. The classification must be able to differentiate between dry and wet ground points.

The first task of point cloud classification is currently under construction but the authors reaches first good results for the segmentation task. Fig. 3 show the first result of a voxel based ground detection.

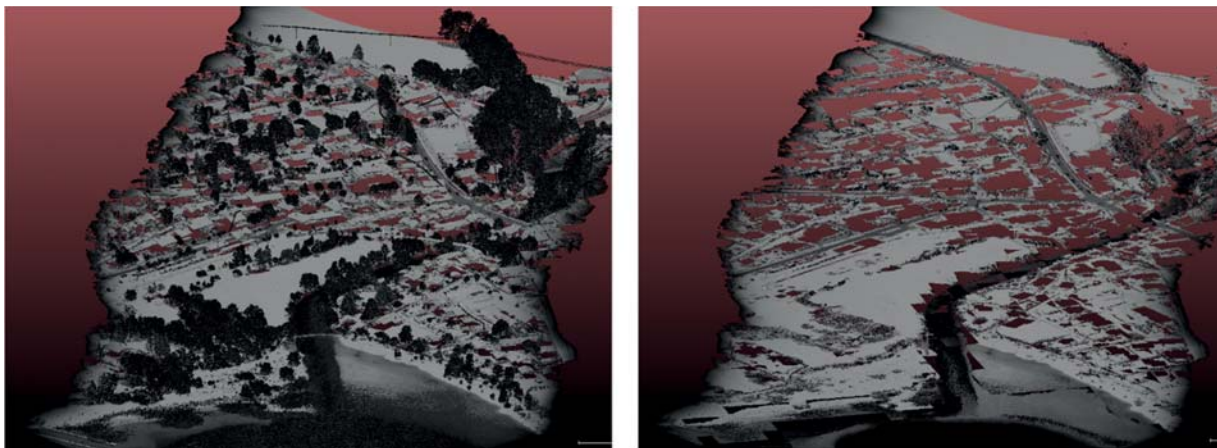


Fig. 3: Result of a voxel based ground detection. Left: original data with intensity as grey value. Right: ground points with intensity as grey value.

These ground segments are handled as the first class and the non-ground points will be segmented and matched to the other classes in the future. The voxel based ground detection will be introduced in a future paper when the parameter adjustment is done correctly. This ground detection is a fast version, which processes the shown 5 Mio points in Fig. 3, in less than 1 second on a single thread and is highly parallelizable. For the processing an intel core i7-6820HQ CPU with 2.7 GHz was used. In short, the used method is to use the position of voxels to detect the lowest ones in a local area. Because of the footprints of roofs, roof points which are marked as ground need to be filtered. The used filter looks for big distances of the height coordinate of neighboured voxels. Since steep slopes create big distances in height too, they are marked as non-ground points as well as roof points (this is called overcut). If there is noise between the roof and the ground, or the building is very flat, some roof points are still marked as ground points (this is called undercut). The undercut and overcut are shown in Fig. 4 on the example of a part of Gmund (a Village north to the Tegernsee, which can be seen in Fig. 3).

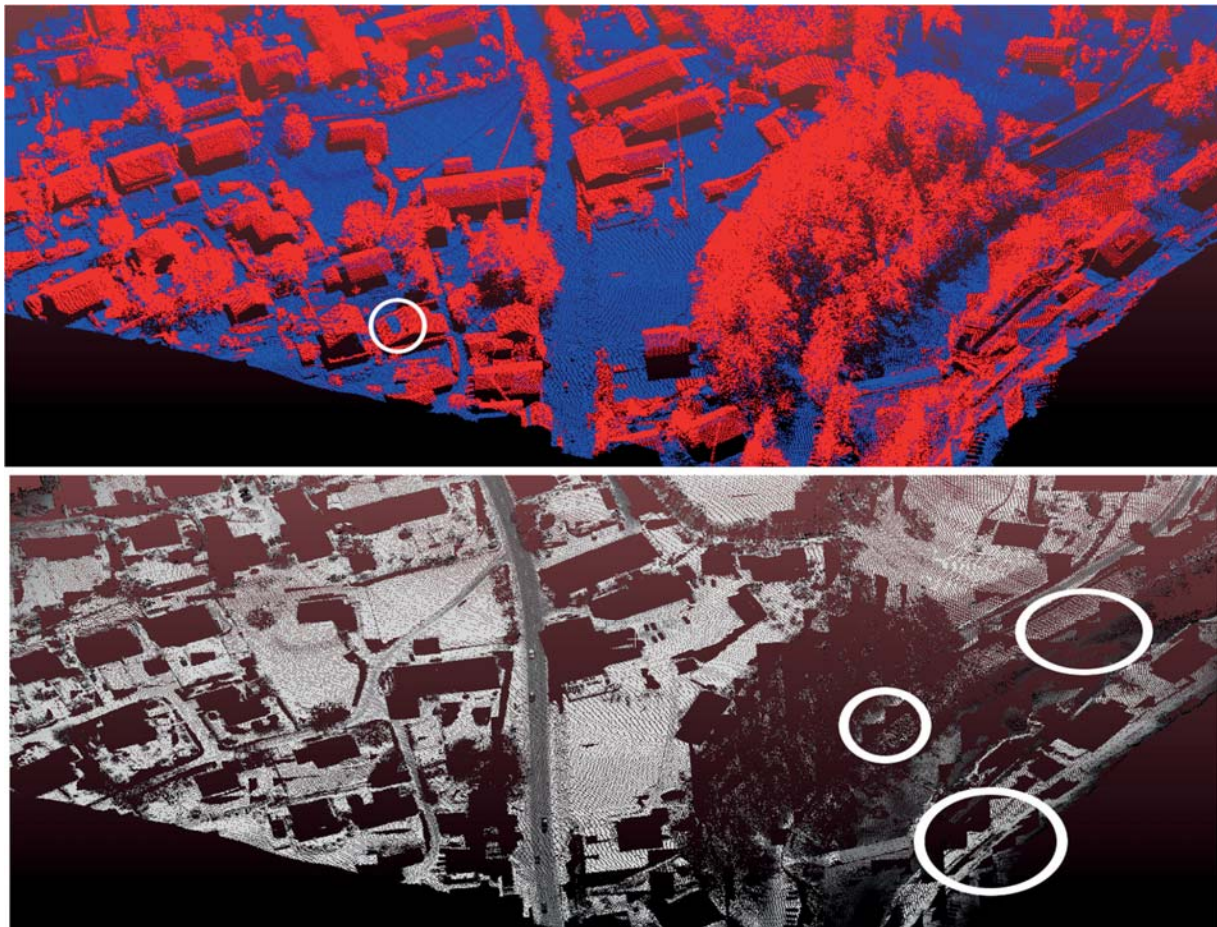


Fig. 4: Some problems with the ground detection. Top: coloured point types blue = ground, red = others. Marked area: undercut, some roof points detected as ground. Bottom: ground points with intensity as grey value. Marked areas: overcut, some points of steep slopes are not marked as ground points.

Future work will pursue the idea to create a filter for wrong classifications by looking at the neighbourhood. An undercut like in Fig. 4 will result in a small ground area inside a roof area. Therefore, the ground area should be marked as roof area. The same filter should be able to decimate the overcuts. This are areas of another class inside a ground area and therefore marked as ground.

In summary, the special interest of future work will be the introduction of methods for automatic point cloud classification, registration and change detection, which can handle big point clouds in a suitable time. Therefore, these methods will fully depend on a voxel structure to make the processing time independent of the point cloud's number of points. These methods should also be able to process non-urban and non-rigid areas.

6 Acknowledgments

This research is funded by Bayerische Forschungsförderung, project “Schritthaltende 3D-Rekonstruktion und –Analyse (AZ-1184-15)”, sub project „Änderungsdetektion in Punktwolken“. The dataset is provided by our project partner SteinbacherConsult.

7 References

- AIJAZI, A. K., CHECCHIN, P. & TRASSOUDAIN, L., 2013: Segmentation Based Classification of 3D Urban Point Clouds: A Super-Voxel Based Approach with Evaluation. *Remote Sensing* **5**(4), 1624-1650.
- BESL, P. J., MCKAY, D., 1992: A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(2), 239-256.
- BRUNNSTRÖM, K. STODDART, A. J., 1996: Genetic Algorithms for Free-Form Surface Matching. *Pattern Recognition, Proceedings of the 13th International Conference* **4**, 689-693.
- CHUA, C. S. & JARVIS, R., 1997: Point Signatures: A New Representation for 3D Object Recognition. *International Journal of Computer Vision* **25**(1), 63-85.
- CHUNG, D. H., YUN, I. D. & LEE, S. U., 1997: Registration of multiple range views using the reverse calibration technique. *Pattern Recognition* **31**(4), 457-464.
- DELAI, F., MORETTO, J., PICCO, L., RIGON, E., RAVAZZOLO, D. & LENZI, A.M., 2014: Analysis of Morphological Processes in a Disturbed Gravel-Bed River (Piave River): Integration of LiDAR Data and Colour Bathymetry. *Journal of Civil Engineering and Architecture* **8**(5), 639-648.
- EUROPEAN UNION, 2000: Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. <http://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A32000L0060>, last access on November 29, 2016.
- EUROPEAN UNION, 2007: Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32007L0060>, last access on November 29, 2016.
- GIRADEAU-MONTAUT, D., ROUX, M., MARC, R. & THIBAUT, G., 2005: Change detection on point cloud data acquired with a ground laser scanner. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**(3), W19.
- HEBEL, M. & STILLA, U., 2009: Automatische Koregistrierung von ALS-Daten aus mehreren Schrägansichten städtischer Quartiere. *Photogrammetrie Fernerkundung Geoinformation* **3**, 261-275.
- HEBEL, M. & STILLA, U., 2012: Simultaneous Calibration of ALS Systems and Alignment of Multiview LiDAR Scans of Urban Areas. *IEEE Transactions on Geoscience and Remote Sensing* **50**(6), 2364-2379.
- 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.

- JOHNSON, A. E., 1997: Spin-Images: A Representation for 3-D Surface Matching. Ph. D. thesis, Carnegie Mellon University, USA.
- JUTZI, B. & STILLA, U., 2006: Range determination with waveform recording laser systems using a Wiener Filter. *ISPRS Journal of Photogrammetry & Remote Sensing* **61**, 95-107.
- MANDLBURGER, G., PFENNIGBAUER, M., STEINBACHER, F. & PFEIFER, N., 2011: Airborne Hydrographic LiDAR Mapping – Potential of a new technique for capturing shallow water bodies. *Proceedings of the 19th International Congress on Modelling and Simulation*, Perth, Australia, 12-16.
- MANDLBURGER, G., HAUER, C., WIESER, M. & PFEIFER, N., 2015: Topo-Bathymetric LiDAR for Monitoring River Morphodynamics and Instream Habitats - A Case Study at the Pielach River. *Remote Sensing* **7**(5), 6160-6195.
- MASUDA, T., 2001: Generation of Geometric Model by Registration and Integration of Multiple Range Images. *Third International Conference on 3-D Digital Imaging and Modeling*, 254-261.
- MURAKAMI, H., NAKAGAWA, K., HASEGAWA, H., SHIBATA, T. & IWANAMI, E., 1999: Change detection of buildings using an airborne laser scanner. *ISPRS Journal of Photogrammetry & Remote Sensing* **54**, 148-152.
- NIEMEYER, J., ROTTENSTEINER, F. & SOERGEL, U., 2014: Contextual classification of lidar data and building object detection in urban areas. *ISPRS Journal of Photogrammetry & Remote Sensing* **87**, 152-165.
- PFENNIGBAUER, M., STEINBACHER, F., ULLRICH, A. & AUFLEGER, M., 2010: A Novel Approach to Laser-Based Hydrographic Data Acquisition. *European LiDAR Mapping Forum*, Salzburg, Austria.
- REITBERGER, J., KRZYSZEK, PL., AND STILLA, U., 2008: Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing* **29**(5), 1407-1431.
- REITBERGER, J., SCHNÖRR, CL., KRZYSZEK, PL., AND STILLA, U., 2009: 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS Journal of Photogrammetry & Remote Sensing* **64**(6), 561-574.
- RUTZINGER, M., HÖFLE, B., HOLLAUS, M. & PFEIFER, N., 2008: Object-Based Point Cloud Analysis of Full-Waveform Airborne Laser Scanning Data for Urban Vegetation Classification. *Sensors* **8**(8), 4505-4528.
- SALVI, J., MATABOSCH, C., FOFL, D. & FOREST, J., 2006: A review of recent range image registration methods with accuracy evaluation. *Image and Vision computing* **25**(5), 578-596.
- 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).
- STEINBACHER, F., PFENNIGBAUER, M., AUFLEGER, M. & ULLRICH, A., 2012: High resolution airborne shallow water mapping. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39**, 55-60.
- TAM, G. K. L., CHENG, Z. Q., LAI, Y. K., LANGBEIN, F. C., LIU, Y. MARSHALL, D., MARTIN, R. R. SUN, X. F. & ROSIN, P. L., 2013: Registration of 3D Point Clouds and Meshes: A Survey

- From Rigid to Non-Rigid. *IEEE transactions on visualization and computer graphics* **19**(7), 1199-1217.
- TAREL, J. P., CIVI, H. & COOPER, D. B., 1998: Pose estimation of free-form 3D objects without point matching using algebraic surface models. *Proceedings of IEEE Workshop Model Based 3D Image Analysis*, 13-21.
- WEINMANN, M., 2016: *Reconstruction and Analysis of 3D Scenes*. Springer International Publishing, Cham. doi 10, 978-3.
- WU, B., YU, B., YUE, W., SHU, S., TAN, W., HU, C., HUANG, Y., WU, J. & LIU, H., 2013: A Voxel-Based Method for Automated Identification and Morphological Parameters Estimation of Individual Street Trees from Mobile Laser Scanning Data. *Remote Sensing* **5**(2), 584-611.
- ZHANG, Z., ZHANG, L., TONG, X., MATHIOPOULOS, P. T., GUO, B., HUANG, X., WANG, Z. & WANG, Y., 2016: A Multilevel Point-Cluster-Based Discriminative Feature for ALS Point Cloud Classification. *IEEE Transactions on Geoscience and Remote Sensing* **54**(6), 3309-3321.

Untersuchungen des Einflusses von Satellitenjitter auf die Bildqualität von hochauflösenden optischen Sensoren

ANGELO CONJO¹ & RALF REULKE²

Zusammenfassung: Jitter werden in der Raumfahrt als ungewollte, stochastische Bewegungen definiert, die durch aktive Instrumente, z.B. zur Lageänderung im Raum, zur Energiegewinnung oder für die Kühlung des Systems verursacht werden. Jede einzelne Komponente hat eine eigene typische Jittercharakteristik. Durch das Zusammenwirken dieser Systeme entsteht ein spezifisches Spektrum.

Ein vorhandener Messplatz wurde so erweitert, dass quer zur Bewegungsrichtung die Störungen durch Jitter simuliert werden können. Es wurde ein geeignetes Auswertungsverfahren entwickelt, um die Änderung der Abbildungsqualität zu erfassen. Dazu wurde die durch Jitter verursachte Verschmierung ist an der resultierenden Point-Spread-Function (PSF) anhand ihrer Verbreiterung und ihrer vertikalen Pixelverschiebung beschrieben.

1 Einleitung

Um die Effizienz von CCD-Zeilensensoren zu verbessern, werden seit den 90er Jahren sogenannte Time-Delay Integration (TDI)-Zeilensensoren verwendet. Sie ermöglichen durch Mehrfachbelichtung und Aufintegration eine Verlängerung der Belichtungszeit und dadurch eine wesentliche Verbesserung des Signal-Rausch-Verhältnisses. Gleichzeitig muss man sich intensiver mit Jitter-Effekten beschäftigen, da Jitter die Bildqualität entscheidend beeinflussen kann (siehe zum Beispiel (RAN 2008)).

Heutzutage ist es ohne weiteres möglich, diese Jittereffekte durch Einsatz von Lageregelungs- und Stabilisierungssystemen zu minimieren und durch Bildverarbeitungsalgorithmen den Jitter zu korrigieren. Offen bleibt jedoch, wie im Detail die Bildqualität durch Jitter in Form und Größe beeinträchtigt wird. Dazu wurde ein vorhandener Messplatz so erweitert, dass quer zur Bewegungsrichtung die Störungen durch Jitter simuliert werden können. Es wurden für bekannte Satellitenplattformen die Größe und Spektren des Jitter festgestellt, um einen geeigneten Aktor zu spezifizieren. Dieser Antrieb wurde adäquat charakterisiert um definierte Jitterfunktionen und ihre Reproduzierbarkeit zu gewährleisten. Anschließend wurde er auf dem Messplatz integriert.

Um stochastische Störbewegungen ausführen zu können, wurde ein Generator für die Jittermodulation programmiert. Ziel ist eine experimentelle Bestimmung der Bildgüte bei verschiedenen Jittersimulationen auf dem vorhandenen Messplatz. Es soll herausgefunden werden, welche Parameter welchen Einfluss auf die Abbildung des gegebenen optischen Systems haben.

¹ Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Institut für Optische Sensorsysteme, Rutherfordstrasse 2, D-12489 Berlin, E-Mail: Angelo.Conjo@dlr.de

² Humboldt-Universität zu Berlin, Institut für Informatik, D-10099 Berlin, E-Mail: reulke@informatik.hu-berlin.de

Nach einigen grundlegenden Betrachtungen zu den Grundlagen von TDI und Jitter, wird der Messplatz vorgestellt. Das dritte Kapitel beschreibt das Experiment und die Datenauswertung. Im vierten Kapitel werden die Ergebnisse zusammengefasst. Der Beitrag endet mit einem Fazit und einem Ausblick.

2 Grundlagen

2.1 TDI Sensor und Jitter

Der TDI Sensor ist eine Sonderform eines Zeilensensors. Durch Mehrfachbelichtung von ein und demselben Objekt lässt sich eine SNR Verbesserung erzielen. Im Gegensatz zu konventionellen Zeilensensoren steigt bei dieser Technologie der Kalibrierungsaufwand, da Sensor und Auslesegeschwindigkeit synchronisiert werden müssen. Wenn die Synchronisierung nicht gewährleistet ist, kommt es zur Verschlechterung der Bilddaten.

Allgemein wird Jitter eingeteilt in Niedrigfrequenz- und Hochfrequenzjitter. Betrachtet man einen langen Zeitraum, sind die resultierenden Änderungen stochastisch. Bei Zerlegung in kleinere Zeitintervalle lassen sich wiederholende Zyklen erkennen. Ist die Integrationszeit kleiner oder gleich der Periode einer Sinusschwingung, handelt es sich um ein Niedrigfrequenzjitter. Ist die Integrationszeit größer spricht man von Hochfrequenzjitter. Niedrigfrequenzjitter führt zu strukturellen Veränderungen des aufgenommenen Objekts im Bild, Hochfrequenzjitter führt zu Verschmierungen im Bild. Bei Erdbeobachtungssatelliten wie: Alos, Aster, ZY-3, Beijing-1 treten Amplituden zwischen 0,2 - 3 Pixel auf. Die Frequenzen können sich in einem Bereich von 0 bis über 1000 Hz bewegen (TONG et al. 2014).

2.2 Messplatz

Um den Einfluss der Winkelfehlausrichtung und der Synchronisationsfehlpassung auf die optische Abbildungsqualität zu untersuchen, wurde ein spezieller experimenteller Aufbau und ein Verfahren entwickelt (**Fehler! Verweisquelle konnte nicht gefunden werden.**).

Zur Analyse der Bildqualität wird die Modulation Transfer Function (MTF) entlang der Flugrichtung (TDI-Richtung) analysiert. Das Target, ein Pinhole von 10 μm Durchmesser, wurde auf einer schnell beweglichen linearen Plattform montiert, um die Bodenbahngeschwindigkeit des Satelliten zu simulieren. Die Plattform ist die Ultra-Präzisions-Linearmotorstufe XMS50 von Newport (Linearstage). Das Target wird durch eine Quarz Wolfram Halogen (QTH) Lampe über eine Glasfaser mit einer Fokussierlinse beleuchtet. Die Wellenlängenauswahl erfolgt durch Interferenzfilter mit 10 nm Bandbreite. Die Geschwindigkeit der schnellen Plattform muss mit der Zeilensynchronisationsfrequenz des TDI-Regimes synchronisiert werden.

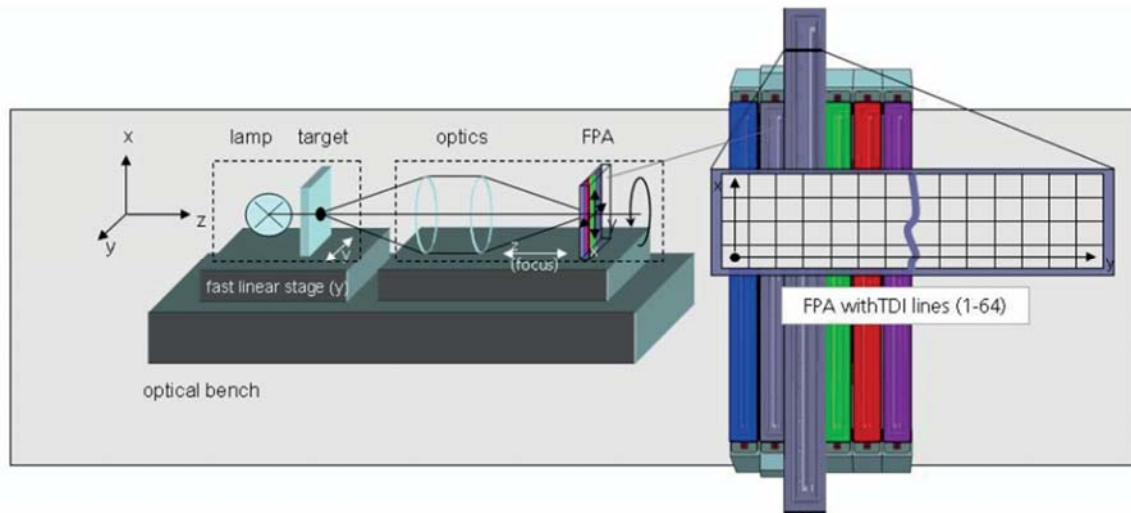


Abb. 1: Dynamischer Testaufbau zur Vermessung von TDI-Sensoren (aus SCHWARZER (2008))

Eine 1:1-Relaisoptik transformiert das Bild des Targets in die Fokalebene (FPA). Dazu wird das telezentrische Objektiv S5LPJ7210 von Sill Optics mit einer Brennweite von 72 mm, einem Arbeitsabstand von 180 mm und einem Sichtfeld von $70 \times 70 \text{ mm}^2$ verwendet.

Der TDI-Sensor, ist der CCD8091 von Fairchild. Es ist ein $9216 \text{ Pixel} \times 128 \text{ Zeilen}$ Hochgeschwindigkeits-TDI-Sensor mit einer Pixelgröße von $8,75 \times 8,75 \mu\text{m}^2$ und einer 14-Bit-A / D-Wandlung. Die elektronisch wählbare Anzahl von TDI-Stufen beträgt 4, 8, 16, 32, 64, 96 und 128. Weil wir eine 685 km Umlaufbahn simulieren wollen, müssen wir für diesen Detektor eine Zeilensynchronisationsfrequenz von 9,26 kHz verwenden. Unter Berücksichtigung der Abbildungsbeziehungen der Optik benötigen wir eine Geschwindigkeit der linearen Stufe von 81 mm / s , um die Geschwindigkeit des Satelliten zu simulieren.

2.3 Messplatzerweiterung

Der XMS 50 ist ein linearer Präzisionspositionierer der Firma Newport. Auf ihm werden Justierelemente, Optiken, die notwendigen Erweiterungen wie Adapter und Jitterantrieb montiert. Er realisiert die synchrone Bewegung des Objektpunktes mit der Ladungsverschiebung der TDI Kamera.

Wegen der Erweiterung des Schlittenaufbaus muss der XMS 50 an das zu bewegende Gewicht angeglichen werden. Dazu sind die Parameter „ScalingAcceleration“ (passt die ausgehende Motorspannung des Controllers an für die maximale Beschleunigung) und „Acceleration-Limit“ (gibt die vom beladenen Schlitten maximal mögliche Beschleunigung an) in der Initialisierungsdatei des Motorcontrollers anzupassen. Das Ziel ist die Vorbeugung und Minimierung von Störgrößen, die neben dem zu untersuchenden Jitter auch Einfluss auf die PSF-Messung haben.

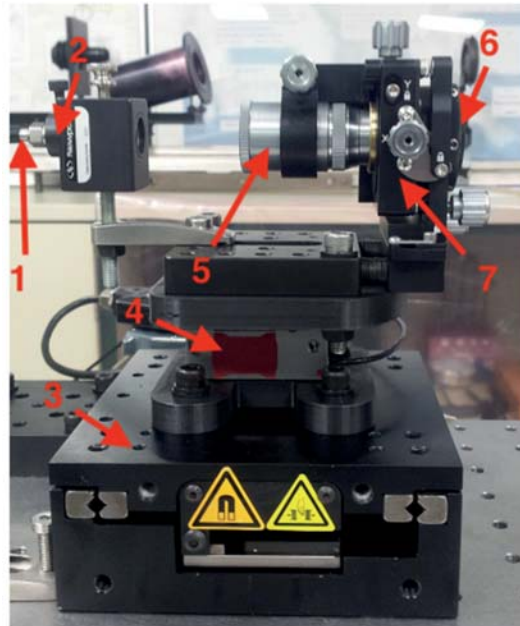


Abb. 2: Erweiterung der linearen Plattform mit Piezo

In der oben gezeigten Darstellung ist die erweiterte Konstruktion zu sehen (siehe Abb. 2). Mit Hilfe der Glasfaseroptik (1) ist es möglich, die Beleuchtung auf den Schlitten zu bringen. Die Faser besitzt einen Durchmesser von $200\ \mu\text{m}$. Am Ende dieser Glasfaser befindet sich ein Kollimator (2). Dieser wird an den beweglichen Teil des Versuchsaufbaus, der Linearstage, gekoppelt. Der XMS 50 (3) simuliert die Bahngeschwindigkeit des Satelliten, diese ist mit der Auslesegeschwindigkeit der TDI Kamera synchronisiert. Auf der Linearstage ist der PU 100 HL (4) montiert. Er simuliert die Jitter-Bewegungen quer zur Bahnrichtung des Schlittens. Unter und über dem Piezo sind Adapterplatten angeflanscht, die zum einen den Jitterantrieb mit der Linearstage verbinden und zum anderen den Piezo-Aktor mit dem optischen Aufbau fixieren. Von der Austrittsöffnung des Kollimators passiert der Lichtstrahl ein Mikroskopobjektiv (5). Mit ihm ist es möglich, durch eine mechanische Blende die Belichtung zu steuern. Das Objektiv fokussiert den ankommenden Lichtstrahl auf das Pinhole oder Lochblende (6). Der Durchmesser sollte um $10\ \mu\text{m}$ betragen. Gemessen wurde ein Durchmesser von $12,1\ \mu\text{m}$. Pinhole und Mikroskopobjektiv sind an einer 5-Achsen-Justierung (7) montiert. Damit wird ein optimales Einstellen des Signaldurchlaufs gewährleistet.

2.4 Untersuchungen zum XMS 50

Um PSF Messungen mit Jitter messen zu können, muss eine synchrone Geschwindigkeit des Positionierers gewährleistet sein. Weiterhin soll ein Einblick auf das Verhalten des Schlittens bei Jitter-Erregung gegeben werden.

Zur Parameterfindung für eine konstante Geschwindigkeit fährt der Schlitten eine festgelegte Strecke von $20\ \text{mm}$ mit jeweils wechselnden Beschleunigungsparametern ab. Die erste Messung wird mit einer Beschleunigung von $500\ \text{mm/s}^2$ durchgeführt. Anschließend wird der Parameter um $500\ \text{mm/s}^2$ erhöht und die Messung wiederholt. Die Beschleunigung wird bis zu einem Wert von 2000

mm/s² untersucht. Anhand der Position des Schlittens und der jeweiligen Geschwindigkeit kann eine Grafik erstellt werden, die den Einfluss der Beschleunigung auf die konstante Bewegung darstellt. Anschließend kann aus der grafischen Darstellung abgelesen werden, bei welcher Beschleunigung für dieses Gewicht ein konstanter Verfahrweg gewährleistet wird.

Die Auswertung belegt, dass mit einem Beschleunigungswert von 1500 mm/s² die stabilsten Verfahrwege zu realisieren sind.

Diese Untersuchung macht auch deutlich, dass der XMS 50 keine synchrone Bewegung ausführen kann. Durch Gegenmaßnahmen wie optimale Anpassung an das Gewicht wird der Fehler minimiert. Werden dem Schlitten Jitter hinzugefügt, haben diese auch einen deutlichen Einfluss auf den Speed Ratio. Der Speed Ratio beschreibt das Verhältnis zwischen Zeilenfrequenz und Geschwindigkeit des Schlittens. Bei diesem Test schwanken die aufgenommenen Positionier-Geschwindigkeiten ca. 0,6 mm/s um den Idealwert von 81 mm/s.

2.5 Jitter-Simulationsprogramm

Für die Voruntersuchung wurde ein vorhandenes Programm erweitert. Das Programm setzt sich aus den folgenden Elementen zusammen: TDI Matrix, Ausleseregister zum Auslesen der integrierten Stage Information, punktförmige Lichtquelle, Jitter Störbewegung, Benutzeroberfläche für das Einstellen der gewünschten Simulationsparameter und einer zweidimensionalen grafischen Darstellung nach der Simulation.

Um vergleichbare Werte mit dem Messaufbau zu erhalten, sind die Parameter an die Messwertverhältnisse angepasst. Ein Zeilenelement besitzt die Größe eines Pixels von 8,75µm. Da der verwendete Spot nicht ruckartig, sondern fließend von Stage zu Stage wandert, werden die Pixel in kleinere feine Schritte eingeteilt. Die Anzahl der Integrationsstufen ist wie auf dem Messplatz variabel einstellbar.

Bei jeder Ladungsverschiebung übergibt die letzte Stufe ihre Daten an das Ausleseregister. In ihr wird das Ergebnis der integrierten Signale zeilenweise gesammelt. Anschließend werden die Ausleseregisterdaten genutzt für die grafische Darstellung.

Um eine punktförmige Lichtquelle zu erzeugen, wird eine zweidimensionale Standard- Gauß-Verteilung verwendet. Im Zentrum des Lichtspots ist der Signalwert am höchsten. Entfernt man sich von dem Zentrum, sinkt die Leuchtkraft. Diese Funktion wird wie folgt beschrieben:

$$f(x, y) = \frac{1}{2 \cdot \pi \cdot \sigma_x \cdot \sigma_y} \cdot e^{-\frac{1}{2} \left(\left(\frac{x - \mu_x}{\sigma_x} \right)^2 + \left(\frac{y - \mu_y}{\sigma_y} \right)^2 \right)}$$

Das Ergebnis von $f(x, y)$ ist ein Zahlenwert, der als Lichtintensität im erstellten Programm interpretiert wird. Die Kennzahlen σ_x und σ_y beschreiben die Größe des zu betrachtenden Lichtflecks. Die Störbewegung erfolgt in Y-Richtung. Die Jitteramplitude wird aus Pixelanzahl multipliziert mit der Pixeldimension berechnet.

Die Jitterfrequenz wird als Konstante vom Benutzer eingegeben. Im Programm können die Auswirkungen auf die PSF von Sinusfunktionen als auch von Rauschfunktionen simuliert und dargestellt werden.

3 Experiment

3.1 Messung

Um qualitative Aussagen des abzubildenden Spots treffen zu können, wird ein Versuch ohne Jitter durchgeführt. Die Messergebnisse dienen als Referenz für die mit Jitter behafteten Messungen. Um reproduzierbare Daten zu erhalten, ist es nötig, feste Parameter zu setzen. Der Lichtspot wird für jeden Versuch auf das gleiche Pixel der 9k TDI Zeile fokussiert. Die Untersuchungen beziehen sich auf sinusförmige Jitter. Die Amplituden betragen 10 μm und 30 μm . Es wurde eine Frequenz von 1 Hz und 200 Hz verwendet.

Die Schlittengeschwindigkeit beträgt 81 mm/s und ist synchron mit der Ladungsträgerverschiebung von 9259 Hz der TDI-Kamera. Für jede Jitter-Untersuchung werden 30 Wiederholungen durchgeführt. Die Ansteuerung des Schlittens erfolgt mit LabVIEW. Die Kameraaufnahme erfolgt über einen separaten Computer.

Ist eine Messung beendet, wird die Ausrichtung des Lichtspots überprüft und gegebenenfalls korrigiert. Für die spätere Messauswertung werden die Daten des Schlittens, die sogenannte gathering-Datei, und die Daten der 9kTDI Kamera in sogenannte spot-Dateien abgespeichert.

3.2 Auswertung

Wesentliche Parameter zur Bewertung sind FWHM und Pixelverschiebung. Die Verschmierung des Spots wird über die FWHM berechnet. Die Abkürzung steht für „Full Width at Half Maximum“ und beschreibt die Breite der PSF an dem Ort, an dem die Intensität nur noch die Hälfte des ursprünglichen Peaks beträgt. Mit der Halbwertsbreite ist es außerdem möglich, die Intensitätsverschiebung zu ermitteln. Sie wird berechnet mit

$$FWHM = |x_{halfmax_1} - x_{halfmax_2}|$$

Dabei ist $x_{halfmax_1}$ der erste Wert, an dem die Funktion auf die Hälfte abgesunken ist, $x_{halfmax_2}$ der zweite Wert, an dem sich der Spitzenwert halbiert. Je kleiner die FWHM, desto geringer ist die Verschmierung. Um die Halbwertsbreite berechnen zu können, muss eine lineare Interpolation

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1}$$

durchgeführt werden. Dabei wird ausgenutzt, dass die Orte der Pixel x_1 , x_2 , die zugehörige Intensität y_1 , y_2 und ein gegebener Parameter zwischen den bekannten Pixeln, das halbe Helligkeitsmaximum, bekannt sind. Durch die Anwendung des Verfahrens kann ein Schätzwert für die gesuchte Größe der Position an der gegebenen y -Koordinate des halben Maximums bestimmt werden. Nach Umstellen auf die gesuchte Variable nach

$$x_{halfmax} = \frac{y - y_1}{y_2 - y_1} \cdot (x_2 - x_1) + x_1$$

kann $x_{halfmax}$ in die FWHM Formel eingesetzt und die Halbwertsbreite bestimmt werden.

Für die Berechnung der oben genannten Größen werden nur Daten der Kamera benötigt. In der Spot-Datei befinden sich je Messdurchlauf ungefähr 200.000 Messwerte. Über ein Ausleseprogramm wird die benötigte Helligkeitsverteilung mit den entsprechenden Intensitätswerten herausgefiltert und als txt-Datei mit einer Größe von 32x32 Pixeln abgespeichert.

Um eine aussagekräftigere PSF zu erhalten, werden die Daten durch eine Gaußschen Normalverteilung angepasst.

4 Ergebnisse

In Abb. 3 werden einzelne PSF-Kurven dargestellt, die sich bei diesem Experiment ergeben haben. Gezeigt wird jeweils ein Messdurchlauf je Versuchsspezifikation.

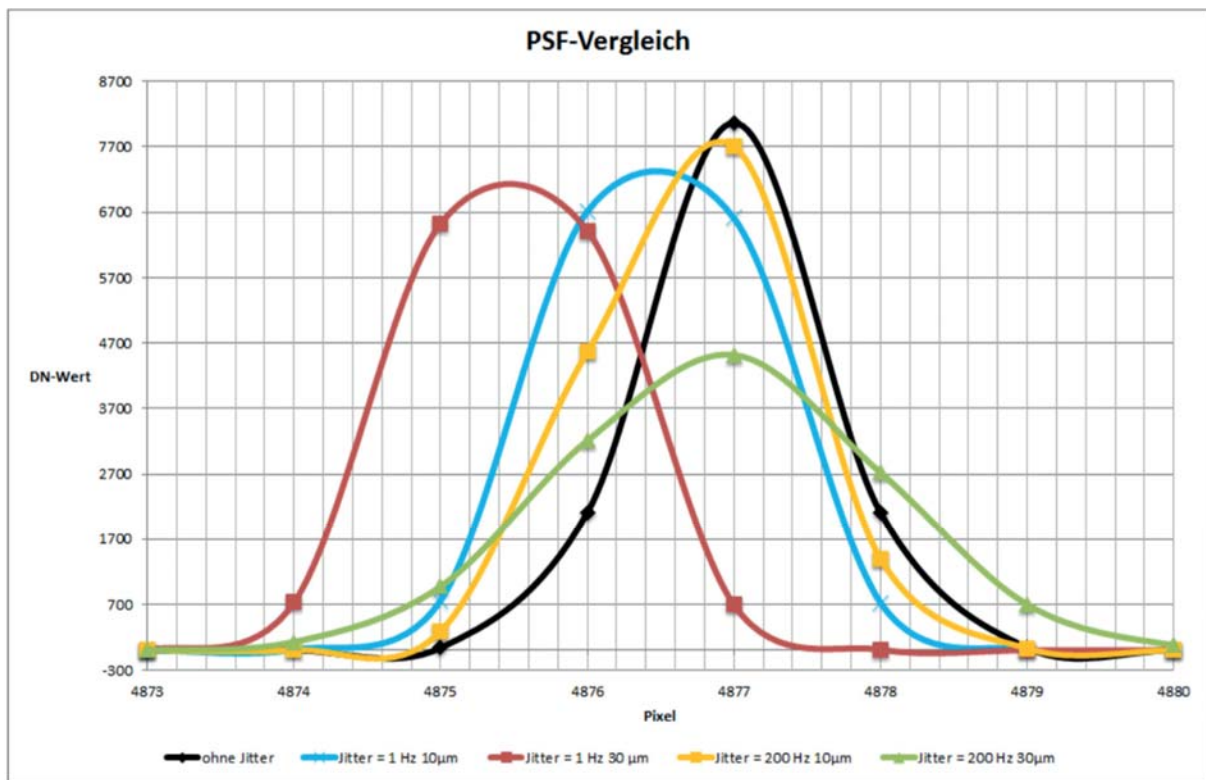


Abb. 3: Ergebnis der PSF-Messungen für verschiedene Frequenzen und Amplituden

Die schwarz gekennzeichnete Funktion stellt die optimale PSF dar. Die 1 Herz Jitter-Messungen weisen eine deutlich breitere PSF auf. Im Gegensatz zu der Referenzkurve ist diese verschoben. Bei den kleinen Frequenzen haben hohe Amplituden große Auswirkung auf die Verschiebung der kompletten PSF. Je höher die Amplitude des Jitter, desto größer ist der Abstand zum Referenz Pixel. Bei 200 Hz und kleiner Amplitude wird ebenfalls die PSF breiter. Die Verteilungskurve ist jedoch nur geringfügig verschoben. Werden bei solchen Frequenzen die Amplituden erhöht, hat

dies große Auswirkungen auf die Verbreiterung der PSF. Je höher die Amplitude des Jitter, desto breiter wird die PSF.

Die Untersuchung zeigt, dass kleine Frequenzen sich auf die komplette Verschiebung der PSF bemerkbar machen. Die Verschmierung ist bei dieser Art Jitter wenig ausgeprägt. Dies erfolgt dadurch, dass bei langsamen Frequenzen die Wahrscheinlichkeit höher ist, nur eine Pixelzeile während des Durchlaufs zu beleuchten. Außerdem reichen bei kleinen Frequenzen die Samplingrate der Kamera noch aus um den Lichtspot noch Abbildungsgetreu, jedoch verschoben abzutasten. Hohe Frequenzen dagegen verschmieren die gesamte PSF. Dies erfolgt dadurch, dass bei schnellen Frequenzen innerhalb eines TDI-Durchlaufs mehrere Pixelzeilen beleuchtet werden können und die Samplingrate nicht mehr ausreicht um schnelle Bewegungsabläufe zu detektieren. Die Verschiebung ist bei dieser Art Jitter weniger ausgeprägt. Hohe Frequenzen bei kleinen Amplituden scheinen daher auch weniger Einfluss auf die Verschlechterung der Bildqualität zu haben.

5 Fazit & Ausblick

Im Rahmen dieser Arbeit wurden die Herangehensweise, Methodik und Problemlösung zur experimentellen Bestimmung der Bildgüte von hochauflösenden optischen Systemen unter Einfluss verschiedener Jittersimulationen betrachtet.

Die Untersuchung zeigte, dass hohe Frequenzen mit großen Amplituden zu einer starken Verschmierung im Bild führen. Bei geringen Amplituden und hohen Frequenzen, nähert sich diese FWHM dem Wert ohne Jitter sehr gut an. Die strukturellen Veränderungen (gemessen durch Verschiebung der PSF) sind gering.

Bei niedrigen Frequenzen ist die Verschmierung bedeutend kleiner als bei hohen Frequenzen. Die Pixelverschiebung ist allerdings stark ausgeprägt und abhängig von der jeweils verwendeten Amplitude. Für weitere Messungen könnten die Jitter in Frequenz und Amplitude weitere Variationen erfahren um eventuell ein Parameter für den Übergang vom Verschmieren zum Verschieben der PSF zu finden. Für eine bessere Interpretation der Ergebnisse ist es unverzichtbar, den Verlauf des Jitter während eines Messdurchlaufs zu detektieren, weil der Antrieb in seiner Bewegung aktuell noch eine Blackbox darstellt.

6 Literaturverzeichnis

- RAN, Q., CHI, Y. & WANG, Z., 2008: Property and Removal of Jitter In Beijing-1 Small Satellite Panchromatic Images, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **37**(B1).
- TONG, X., YE, Z., XU, Y., TANG, X., LIU, S., LI, L., XIE, H., WANG, F., LI, T., & HONG, Z., 2014: Framework of jitter detection and compensation for high resolution satellites. *Remote Sensing* **6**(5), 3944-3964.
- SCHWARZER, H., BOERNER, A., DEGEN, K.-H., ECKARDT, A. & SCHERBAUM, P., 2008: Dynamic PSF and MTF measurements on a 9k TDI CCD. *Proc. of SPIE* **7106**, 71061F-1.

Vergleich der optischen Eigenschaften von hyperspektralen Kameras

CHRISTOPHER SPIESS¹ & RALF REULKE²

Zusammenfassung: Der Beitrag beschäftigt sich mit dem Vergleich zweier hyperspektraler Kameras: ein System mit einem Offner-Spektrographen und eine Kamera mit direkt auf dem Sensorchip aufgebrachten spektralen Filtern. Dabei sollte geklärt werden, ob und wie sich beide Systeme in ihren optischen Eigenschaften unterscheiden. Außerdem sollen entsprechende Maße für die Beschreibung untersucht und dafür angewendet werden. Im Deutschen Zentrum für Luft- und Raumfahrt (DLR) wird zurzeit das Projekt DLR Earth Sensing Imaging Spectrometer (DEISIS) bearbeitet (ECKARDT et al. 2015). Dabei handelt es sich um ein hyperspektrales System mit einem optischen Gitter im Offner-Design. Dieses wird künftig an der International Space Station (ISS) installiert, um von dort die Ozeane, die Erdoberfläche und auch die Atmosphäre zu beobachten. Die Idee zu diesem System entstand bereits im Jahr 1998. Im Rahmen einer Vorstudie um 2002 entstand das Offner-Spektrometer (Bread-board), welches nun für den Vergleich genutzt wird. Im Verlaufe der Arbeit wird untersucht, wie sich die räumlichen und spektralen Eigenschaften (z.B. Auflösung) der klassischen high-end und neuartigen low-cost Variante verhalten.

1 Einleitung

Hyperspektrale Kameras nehmen ein Objekt mit hoher spektraler und moderater räumlicher Auflösung auf. Im Allgemeinen geschieht dies in 100-200 Bändern, mit Bandbreiten zwischen 5 und 10 nm (LIANG 2005). Für die Aufnahme solcher hyperspektralen Bilder gibt es eine Vielzahl von technischen oder physikalischen Möglichkeiten. Klassische Systeme (z.B. ENMAP) verwenden Gitter oder Prismenspektrographen. Hier wird durch eine abbildende Optik eine Abbildung auf den Spalt realisiert. Die räumliche Abtastung erfolgt in Spaltrichtung. Der Spektrograph registriert für jedes Bildpixel senkrecht zum Spalt ein Spektrum. Alternativen sind zum Beispiel spektrale Filter, die direkt auf den Kamerachip aufgebracht werden können. Ziel dieser Arbeit ist der Vergleich von zwei hyperspektralen Systemen, unterschiedlicher Technologie:

1. Ein optisches Gitter im Offner-Design
2. Eine Kamera mit Fabry-Pérot Filtern variabler Wellenlänge, direkt auf dem Detektor

Die zweite Variante ist sehr kompakt, mit geringen Anschaffungskosten verbunden und relativ neu. Im Jahr 2012 wurde von IMEC ein erster Prototyp dieses Filter-Spektrometers veröffentlicht (TACK et al. 2012).

¹ Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Institut für Optische Sensorsysteme, Rutherfordstrasse 2, D-12489 Berlin, E-Mail: C.Spiess@dlr.de

² Humboldt-Universität zu Berlin, Institut für Informatik, D-10099 Berlin, E-Mail: reulke@informatik.hu-berlin.de

Der Aufbau der Arbeit ist wie folgt: Das nächste Kapitel gibt einen Überblick über die aktuellen Verfahren zur Ableitung von spektralen Signaturen. Anschließend wird das Offner-Spektrometer im Detail beschrieben. Um einen direkten Vergleich zwischen Filter- und Offner-Spektrometer zu ermöglichen, wurde für den Detektor des Offner-Spektrometer dieselbe Kamera wie für das Filter-Spektrometer gewählt. Die Kalibrierung des Detektors wird beschrieben und die PRNU-Korrektur eingeführt. Hier wird auch über die Messung der spektralen Empfindlichkeit des Detektors berichtet.

Das Filter-Spektrometer wird genauer im 4. Kapitel betrachtet. Es werden die Ergebnisse der Messung der Transmission der Filter, der spektralen Auflösung und PRNU Korrektur präsentiert. Der Beitrag schließt mit einem Fazit.

2 Hyperspektrale Sensoren

Im Folgenden sollen der Detektor und die beiden Spektrometer kurz vorgestellt werden.

2.1 Detektor

Es wird der CMOSIS CMV2000 mit folgenden Eigenschaften verwendet:

Tab. 1: Detektorparameter

SensorgroÙe	2048 x 1088 Pixel
PixelgröÙe	5,5 x 5,5 μm^2
Sensorformat	2/3“
Bittiefe	10 bit & 12 bit
Auslesegeschwindigkeit	340 bilder/s (für 10 bit)
Full well capacity	13.5 Ke-

Der Detektor wird vom der Firma XIMEA mit einem USB3 Interface zur Verfügung gestellt.

2.2 Filter-Spektrometer

Fabry-Pérot Filter mit einer schmalen Transmissionsfunktion werden direkt auf den Detektorchip aufgetragen. Ein solcher Filter, der räumlich mehrere Zeilen breit ist, wird durch zwei hochreflektive Bragg-Reflektoren im Abstand d gebildet. Zwischen den beiden Elementen existiert ein transparentes Medium der Brechzahl n_g .

Ein hyperspektrales Bild wird durch die Bewegung der Kamera senkrecht zur räumlichen Zeilenrichtung erhalten. Werden die Einzelbilder räumlich koregistriert, lässt sich ein hyperspektrales Bild der gesamten Szene erstellen.

Die Größe des Filter-Spektrometers liegt bei ca. 90 x 25 x 25 mm.

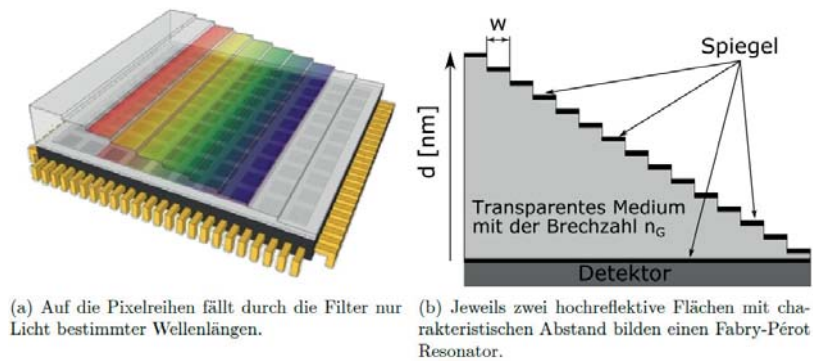


Abb. 1: Fabry-Pérot Filter auf einem lichtempfindlichen Sensor (aus IMAC: Hyper Spectral Imaging)

2.3 Gitter-Spektrometer

Das hyperspektrale Instrument arbeitet wie ein Zeilenscanner. Die Szene wird durch das Objektiv in den Spalt abgebildet. Dieser definiert den Objektausschnitt, welcher über das Offner-Design 1:1 auf den Detektor abgebildet wird – mit der spektralen Zerlegung des Lichtes senkrecht zum Spalt (Abb. 2). Charakteristisch ist, dass das im Spalt befindliche Licht durch den Spiegel kollimiert auf das konvexe Reflexionsgitter trifft, wo die spektrale Aufspaltung erfolgt. Das spektral zerlegte Licht wird anschließend durch den Offner-Spiegel auf den Detektor fokussiert. In vertikaler Richtung erfolgt die Gitterdispersion (spektrale Pixelnummern) und in horizontaler (räumliche Pixelnummer) ist der analysierte Ort des im Spalt befindlichen Lichtes.

Die Größe des Gitter-Spektrometers liegt bei ca. 400 x 200 x 100 mm.

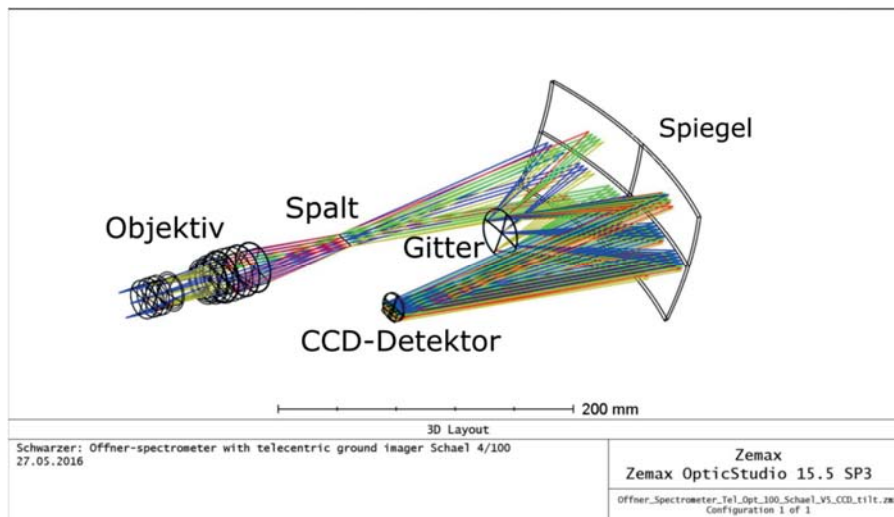


Abb. 2: Zemax Modell des Offner Spektrometers

3 Kalibrierung des Offner-Spektrometers

3.1 Effektivität des Spektrometers

In der Abb. 3 sind die bekannten, theoretischen Einflüsse auf die Gesamteffektivität des Offner-Spektrometers dargestellt. Das betrifft die Transmission des Objektivs, den Reflexionsgrad des Spiegels, die Effektivität des Gitters in der +1.Ordnung und die Elemente des Detektors. Neben der Quanteneffizienz des Sensors, mit Typenbezeichnung CMV2000, müssen noch zwei Schutzgläser aus den Materialien N-BK7 und D263 im Detektoraufbau berücksichtigt werden. Während die UV-Kante durch die Gläser und den Sensor zusammen gebildet wird, ist im infraroten Bereich die Quanteneffizienz des Sensors maßgeblich für ein Absinken der Gesamteffektivität verantwortlich.

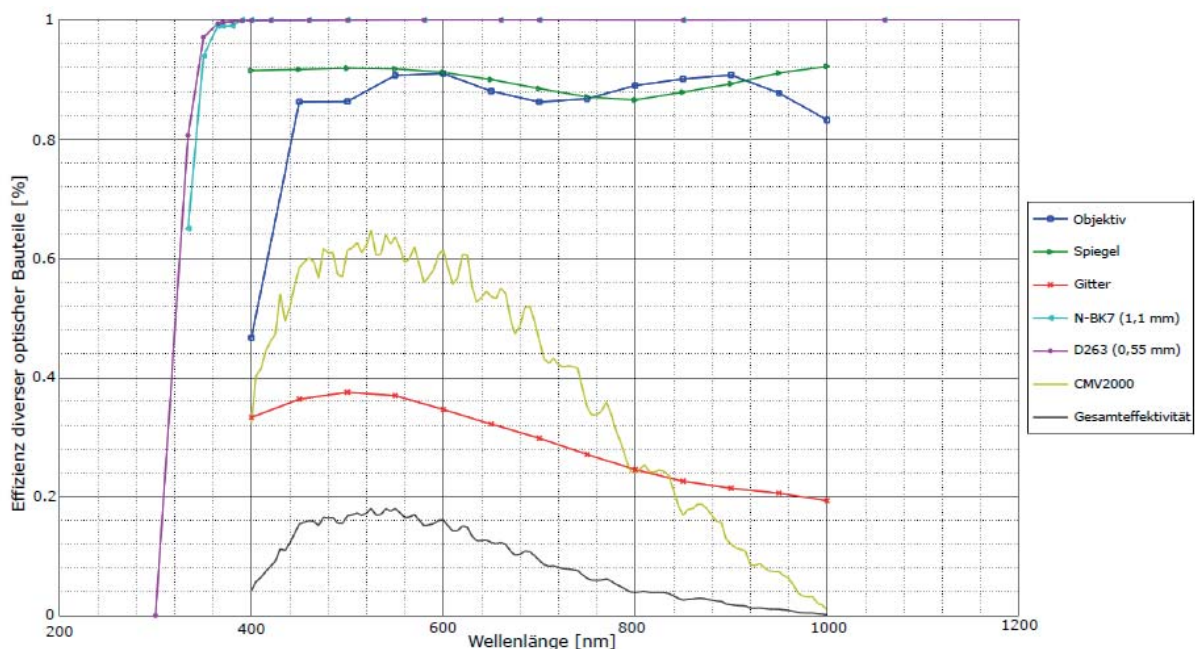


Abb. 3: Theoretische Gesamteffektivität des Offner-Spektrometers

Eine spektrale Kalibrierung kann nun insofern vorgenommen werden, dass Korrekturfaktoren für jede Wellenlänge eingeführt werden, die zu einer gleichbleibenden Quanteneffizienz führen. Es gilt zu beachten, dass auch das Rauschen verstärkt wird und kein verbessertes Signal-Rausch-Verhältnis die Folge ist.

3.2 Radiometrische Kalibrierung

Das erwartete Signal \hat{S} in DN (Digital Numbers) lässt sich folgendermaßen bestimmen

$$\hat{S}[\text{DN}] = \eta_{\text{SG}}^{\text{Det}} \left[\frac{\text{DN}}{e^-} \right] \cdot \eta_{\text{QE}}(\lambda) \left[\frac{e^-}{\text{photon}} \right] \cdot \tau_{\text{int}} [\text{s}] \cdot A_{\text{Det}} [\text{m}^2] \cdot \frac{\lambda}{hc} \left[\frac{\text{m}}{\text{Ws}^2 \text{ m/s}} \right] \cdot E \left[\frac{\text{W}}{\text{m}^2} \right] + \text{DS} \quad (1)$$

Tab. 2: Einflussgrößen zur Signalberechnung

$\eta_{QE}(\lambda)$	Quanteneffizienz	$[e^-/\text{photon}]$
η_{SG}^{Det}	Overall system gain	$[\text{DN}/e^-]$
A_{Det}	Detektorfläche	$[\text{m}^2]$
τ_{int}	Integrationszeit	$[\text{s}]$
E	Bestrahlungsstärke	$[\text{W}/\text{m}^2]$
DS	Temperaturabhängiges Dunkelsignal	$[\text{DN}]$
λ	Wellenlänge	$[m]$

Wir erwarten mindestens drei Rauschkomponenten: Photonenrauschen, Dunkelstrom und das Lese- oder Ausleserauschen (read noise). Während Dunkelstrom- und Photonenrauschen Poisson verteilt sind, wird der read-noise durch eine Normalverteilung beschrieben.

Das Signal wurde bezüglich der PRNU (Pixel Response Non-Uniformity) und DSNU (Dark Signal Non-Uniformity) korrigiert. Beispiele für Korrekturmatrizen sind in der Abb. 1. 4 dargestellt.

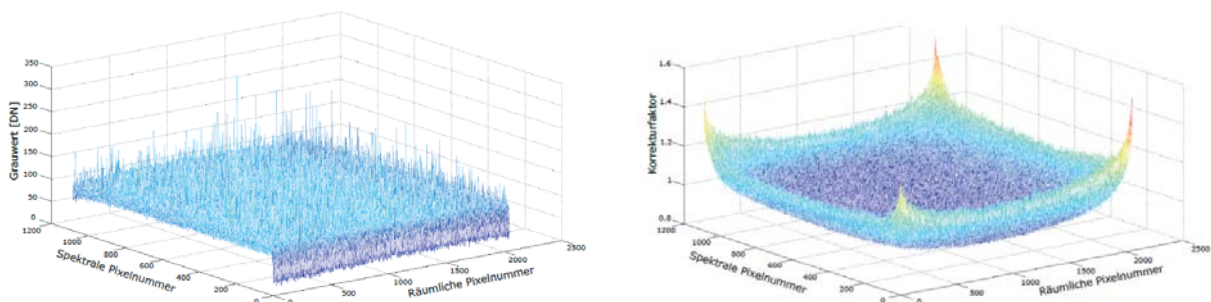


Abb. 4: DSNU-Korrekturwerte für eine Integrationszeit von 100 ms (links); PRNU-Korrekturmatrizen (rechts)

3.3 Keystone und Smile

Keystone und Smile entspricht der klassischen Verzeichnung in räumlicher und spektraler Richtung. Zur Bestimmung des Smiles wird eine stabförmige Penray-Spektrallampe (Argon) parallel zum Spalt des Spektrometers vor das Objektiv positioniert. Um die spektralen Verschiebungen der Spektrallinien im Subpixelbereich bestimmen zu können, wird an sie eine Normalverteilung angepasst:

$$G(\lambda) = C \cdot \exp \left[-\frac{(\lambda - \mu)^2}{2\sigma^2} \right] \quad (2)$$

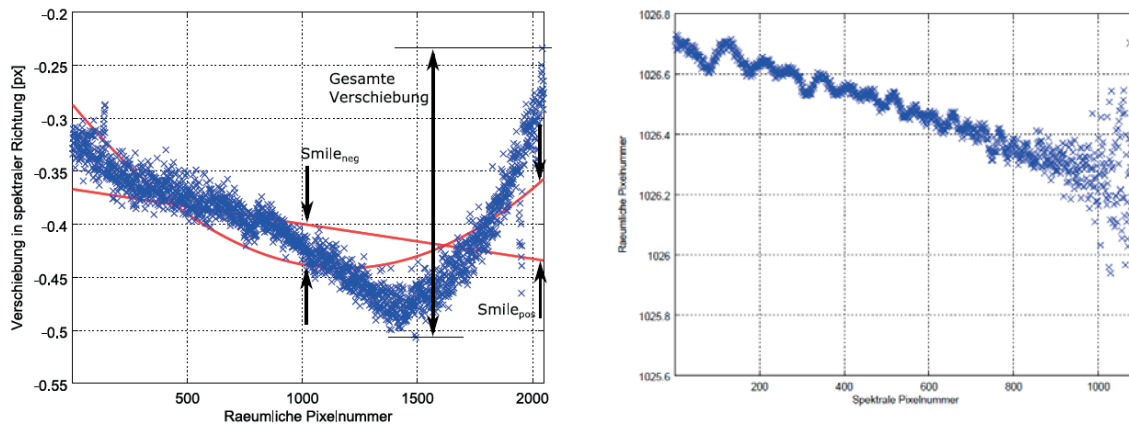


Abb. 5: Smile der 763,5106 nm-Linie mit quadratischen und linearen Fit (links); Keystone bei der mittleren räumlichen Pixelnummer 1026,5 (rechts)

In Abb. 5 (links) ist die Verschiebung des Maximums μ , der Gaußfunktion $G(\lambda)$, in spektraler Richtung dargestellt. Besonders auffällig erscheint der stärkere Einbruch um die räumliche Pixelnummer 1500. Der Grund liegt in der noch nicht ausreichend guten Justage des Detektors. Der Keystone wird an 5 räumlichen Pixeln mithilfe eines Targets ermittelt, das in der Fokalebene des Kollimators liegt. Das Target bildet ein schmaler Spalt, der sich senkrecht zum Spektrometer-Spalt befindet und von einer Halogenlampe beleuchtet wird. Zur Bestimmung der Verzerrung wird der Schwerpunkt entlang der räumlichen Richtung für jeden spektralen Pixel berechnet und in Abb. 5 (rechts) dargestellt. Ab der spektralen Pixelnummer 800 ist eine starke Streuung der Messwerte erkennbar. Dies liegt an einer zu geringen Emission von kurzwelligem Licht der Halogenlampe und einem damit verbundenen schlechten Signal-Rausch-Verhältnis.

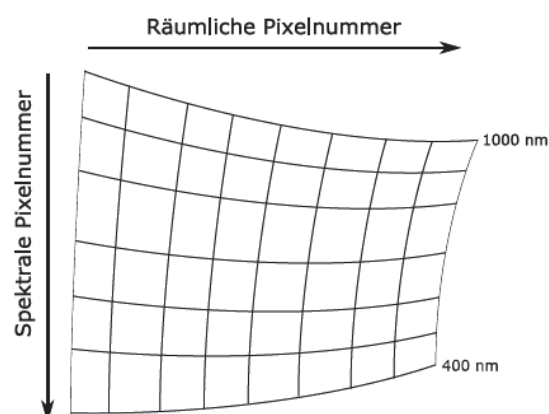


Abb. 6: Stark vereinfachte Zusammenfassung von Keystone und Smile

In der Abbildung 6 findet sich die qualitative Zusammenfassung der Keystone und Smile-Messungen. Durch die nicht ausreichend gute Justage des Detektors ist keine Symmetrieachse auffindbar.

3.4 Die lineare Dispersion und spektrale Bandbreite

Die theoretische lineare Dispersion berechnet sich aus:

$$D_L = \frac{dy}{d\lambda} = f \cdot \frac{1}{g \cdot \cos \left[\arcsin \left(\frac{\lambda}{g} \right) \right]} \quad (3)$$

Mit einer Brennweite des Offner-Spiegels von $f = 117,06 \text{ mm}$ und einer Gitterkonstante von $11,5 \text{ } \mu\text{m}$ bestimmt sich die lineare Dispersion $1/D(\lambda = 700 \text{ nm}) = 0,54 \text{ nm/px}$.

Praktisch wird die lineare Dispersion des Spektrometers mit einer Pen-Ray Spektrallampe (Argon) gemessen. Dazu wird aus der Intensitätsverteilung (siehe Abb. 7, links) der Ort des Maximums geeigneter Spektrallinien über die Wellenlänge aufgetragen (Abb. 7, rechts). Geeignet bedeutet hierbei, dass die Linien spektral nicht zu nah beieinanderliegen dürfen und ein ausreichend hohes Signal-Rausch-Verhältnis aufweisen. Nur die mit schwarzer Wellenlänge gekennzeichneten Linien gehen in die Berechnung der Lineardispersion ein. Der Anstieg der Gerade im Diagramm 7 (rechts) entspricht der linearen Dispersion. Er ist in diesem Beispiel $1/D = 0,52 \text{ nm/px}$ und damit etwas geringer als theoretisch erwartet. Mit der Lineardispersion berechnet sich die spektrale Breite des Lichtes, welches analysiert werden kann, von 419 nm bis 988 nm .

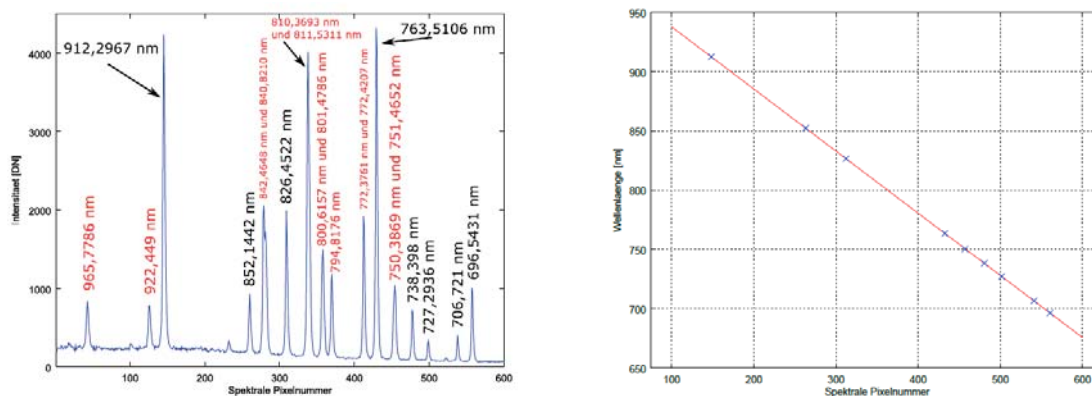


Abb. 7: Intensität entlang der spektralen Pixel beim räumlichen Pixel 1024 (links), Zuordnung der Wellenlänge zu jedem spektralen Pixel zur Berechnung der linearen Dispersion (rechts)

3.5 Spektrale Auflösung

Die spektrale Auflösung wird durch den Gauss-Fit an die Spektrallinie $763,5106 \text{ nm}$ für jedes räumliche Pixel bestimmt (Abb. 8). Auffällig sind die Sprünge an den räumlichen Pixelnummer 381, 810, 1044 und 1472, die symmetrisch um den Pixel 927 liegen. Vermutlich kommen die

Sprünge durch das spezielle Fertigungsverfahren des optischen Gitters. Die spektrale Auflösung $\Delta\lambda$ nimmt mit wachsender Wellenlänge ab, bei einem gleichbleibenden Verhältnis $\Delta\lambda/\lambda$. Wird die spektrale Auflösung als 75% der Halbwertsbreite definiert, liegt sie zwischen 0,41 nm ($\lambda = 419$ nm) und 0,99 nm ($\lambda = 988$ nm).

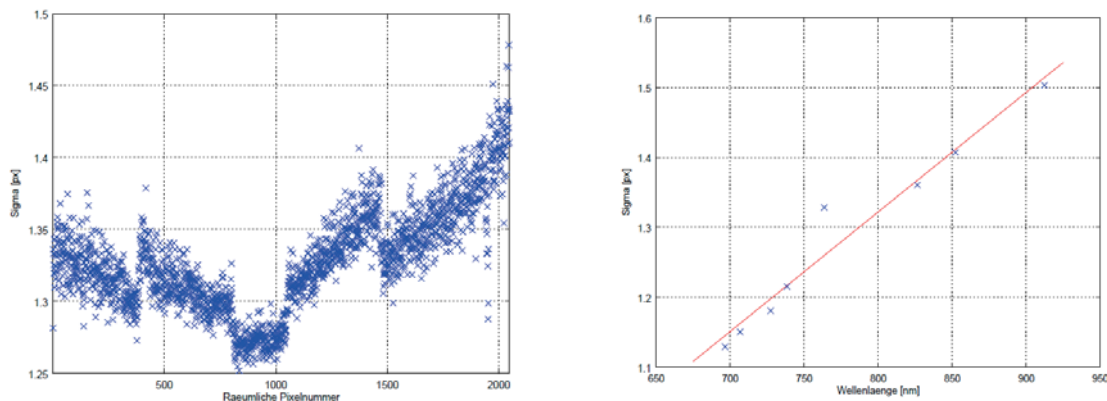


Abb. 8: Sigma der Normalverteilung aus Fit an der Spektrallinie 763,5106 nm als Funktion der räumlichen Pixelnummer (links); Sigma der Gaußverteilung für verschiedene Spektrallinien (rechts)

4 Filter-Spektrometer

4.1 Aufbau

Auf der Sensorfläche befinden sich nach Abb. 9 die Fabry-Pérot-Resonatoren über eine Breite von 8 Pixeln.

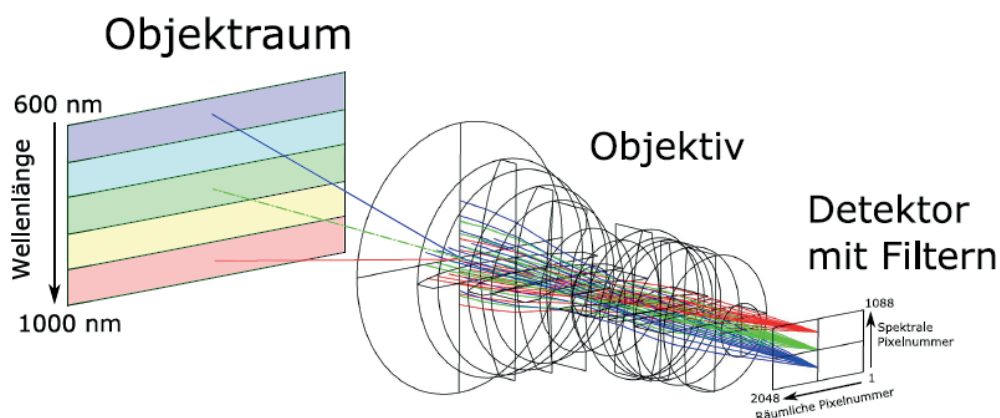


Abb. 9: Hyperspektrale Kamera mit Fabry-Pérot Filtern. Das Objekt wird in Streifen mit unterschiedlichen Wellenlängen (siehe Farben) abgebildet.

4.2 Transmissionsverhalten der Fabry-Pérot-Resonatoren

Die Phasendifferenz der interferierenden Strahlen im Fabry-Pérot-Resonator ändert sich in Abhängigkeit vom Einfallswinkel. Damit hängt die Transmission der Filter des Spektrometers von

der Blendenöffnung des Objektivs ab. Je weiter die Blende geöffnet ist, umso breiter wird die Transmissionskurve durch die Verschiebung der Wellenlänge der maximalen Transmission. Außerdem verschiebt sich das Reflexionsband der Bragg-Reflektoren mit verändertem Einfallswinkel.

Es wird das Transmissionsverhalten der Filter experimentell mit einem Monochromator bestimmt (Abb. 10). Die spektrale Bandbreite am Austrittspalt beträgt 1 nm. Die Wellenlänge wird in 0,5 nm-Schritten verändert. Die Abb. 10 (links) zeigt viele Einbrüche der Halbwertsbreite (FWHM) der Resonatoren. Außerdem ist die Verschiebung der FWHM zwischen senkrechtem Lichteinfall (ohne Objektiv) und schrägem Lichteinfall (Objektiv der Blende f4) erkennbar. Abb 10 (rechts) zeigt die Wellenlänge der maximalen Transmission bei größerer Blendenöffnung.

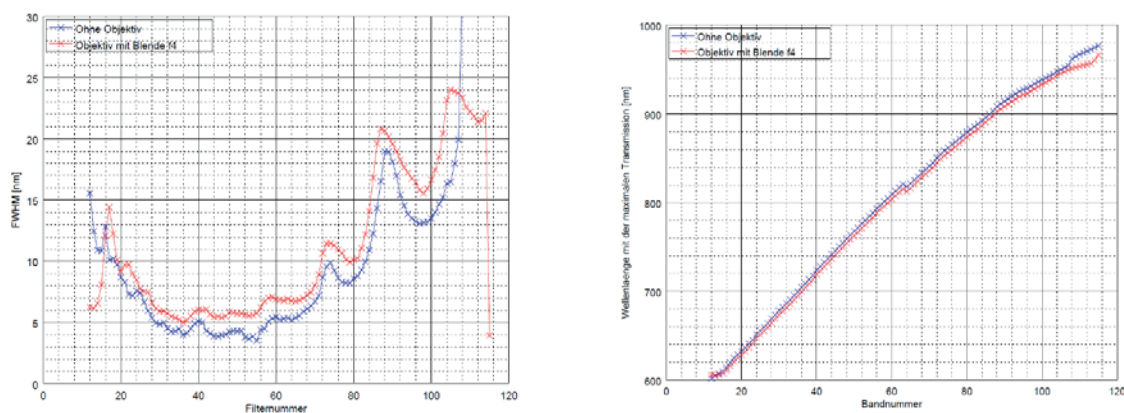


Abb. 10: Die gemessene FWHM für jeden Filter bei der Blendenzahl 4, verglichen mit dem Detektor ohne Objektiv (links); Wellenlänge der maximalen Transmission für jeden Filter mit einem Objektiv der Blendenzahl 4 und der Vergleich mit dem Detektor ohne Objektiv (rechts)

Das Transmissionsverhalten wird durch die Einführung des Verhältnisses v in der Abb. 11 zusammengefasst. Er bildet sich aus der Ratio des Wellenlängenabstandes (Wellenlänge der maximalen Transmission) zweier benachbarter Filter und dessen mittlere Halbwertsbreite:

$$v = \frac{|\lambda_1 - \lambda_2|}{0.5 \cdot (\text{FWHM}(\lambda_1) + \text{FWHM}(\lambda_2))} \quad (4)$$

Es zeigt sich, dass bei Wellenlängen von größer als 850 nm und kleiner als 650 nm die Transmissionskurven zweier benachbarter Filter mehr als 50 % überlappen ($v < 0,5$). In diesen Bereichen ist durchaus ein größerer Abstand der Transmissionskurven möglich, ohne Einbußen hinsichtlich der spektralen Auflösung. Sie wird im grünen, wie auch im Bereich $v < 0,5$, alleine durch die Halbwertsbreite der Filter festgelegt.

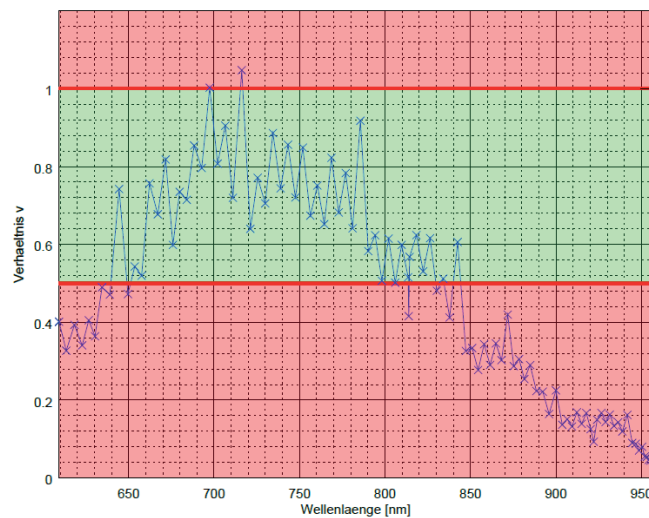


Abb. 11: Verhältnis v in Abhängigkeit von der Wellenlänge

4.3 Hellsignalungleichförmigkeits-Korrektur

Mit einer Ulbricht-Kugel wird auf das Filter-Spektrometer (mit Objektiv) geleuchtet. Die Blendenzahl liegt bei 1,65 und der Fokus wird ins Unendliche gesetzt. Es werden bei mehreren Integrationszeiten Bilder aufgenommen. Anschließend wird für jeden Filter diejenige Integrationszeit verwendet, welche im Mittel zu einer Grauwert von ca. 50% der Sättigung führt. Die Matrix der PRNU-Korrekturfaktoren ist in der Abb. 12 (links) wiederzufinden.

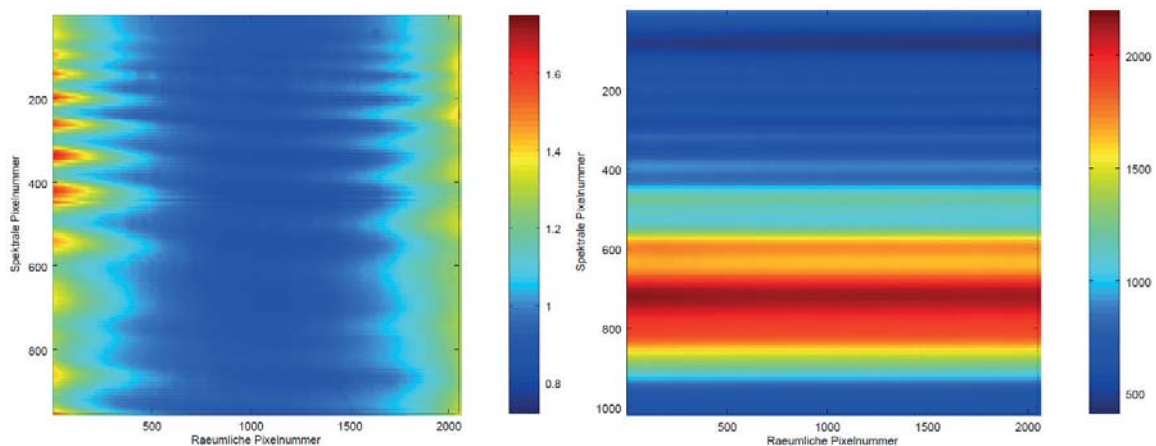


Abb. 12: PRNU-Korrekturfaktoren (links); Detektierte Intensität in Digital Numbers (DN) der Halogenlampe nach der PRNU-Korrektur (rechts)

Obwohl der Detektor PRNU-Korrigiert ist, variiert die Intensität nach Abbildung 11 (rechts) in spektraler Richtung.

In der Abb. 13 wird die detektierte Intensität nach Abb. 12 (rechts) mit der Quanteneffizienz des Detektors und der Halbwertsbreite der Filter verglichen. Die nach der PRNU-Korrektur detektierte Intensität ist zum einen auf den Detektor zurück zu führen, welcher unterschiedliche spektrale Empfindlichkeiten aufweist. Zum anderen wird bei einer breiten Transmissionskurve der Filter mehr Licht auf den Detektor fallen.

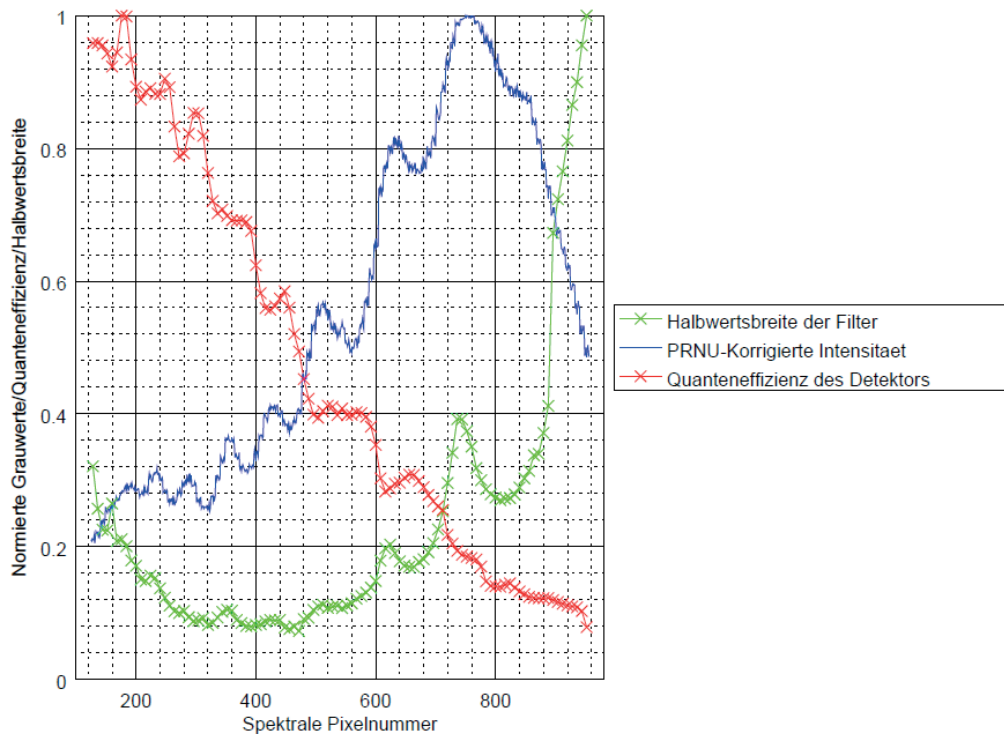


Abb. 13 Vergleich der PRNU-Korrigierten Intensität mit der Quanteneffizienz des Detektors, sowie der Halbwertsbreite der Filter

5 Fazit

Die Tab. 3 zeigt die Unterschiede in der spektralen Auflösung der beiden Spektrometer auf. Sie ist deutlich schlechter und weist außerdem Nicht-Linearitäten auf.

Tab. 3: Zusammenfassung der spektralen Auflösung

Spektrometerart	Spektrale Bandbreite	Spektrale Auflösung
Offner-Spektrometer	419-988 nm	0,41-0,99 nm (Linear)
Filter-Spektrometer	606-966 nm	5,1-23,94 nm (Nichtlinear)

Vor beide Spektrometer wird nun eine Argon-Spektrallampe positioniert und die Intensität detektiert. Die Abb. 14 zeigt den Intensitätsverlauf. Die Spektrallinien können von dem Filter-Spektrometer nicht oder nur schwer unterschieden werden.

In LIANG (2005) ist eine hyperspektrale Kamera so definiert, dass sie 100-200 Wellenlängenbänder besitzt mit jeweils einer Bandbreite von 5-10 nm. Eine multispektrale Kamera dagegen habe 5-10 Bänder mit einer Bandbreite von 70-400 nm. Das untersuchte Filter-Spektrometer passt nur zu einem gewissen Grad in die Definition einer hyperspektralen Kamera. Dagegen ist die spektrale Auflösung deutlich besser als die einer multispektralen Kamera. Das Filter-Spektrometer kann schlussendlich als Bindeglied zwischen hyper- und multispektral eingeschätzt werden. Anwendungen für diese sehr kompakte Kamera finden sich dort, wo keine große spektrale Auflösung gefragt ist und Platzmangel vorherrscht.

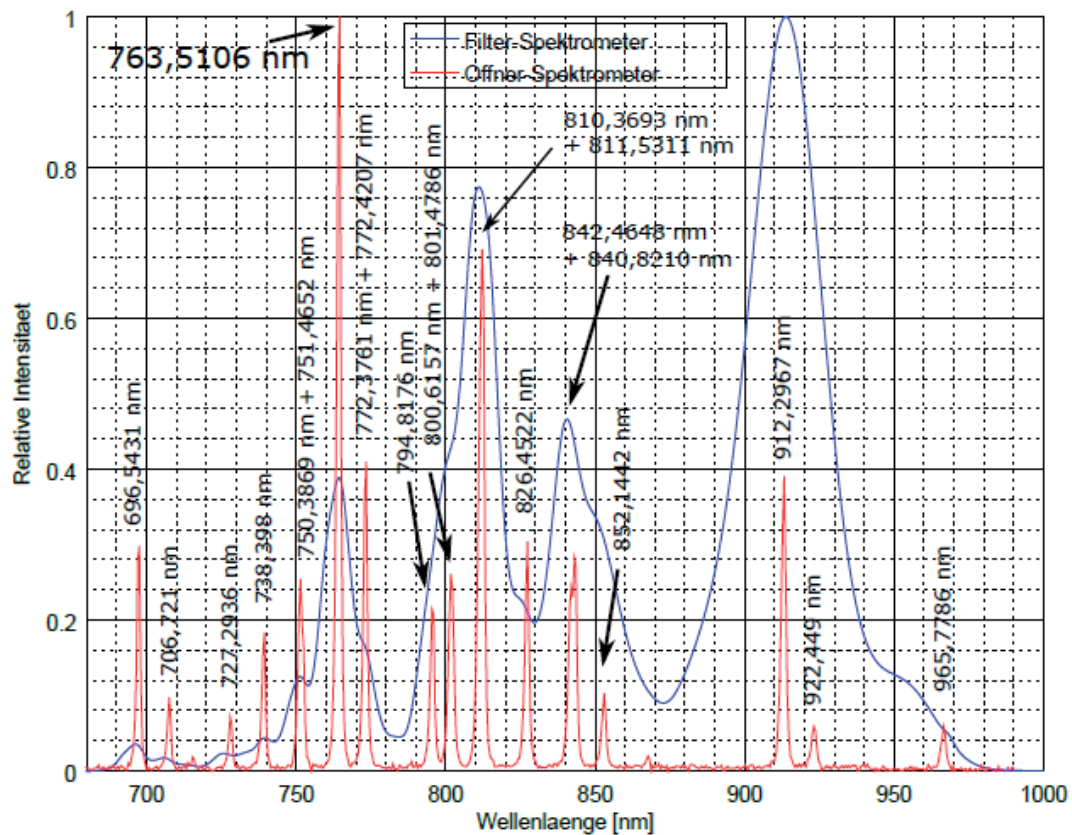


Abb. 14: Vergleich der detektierten Intensität in Abhängigkeit von der Wellenlänge vom Filter- und Offner-Spektrometer bei der Einstrahlung mit einer Argon-Spektrallampe

6 Literaturverzeichnis

- LIANG, S., 2005: Quantitative Remote Sensing of Land Surfaces. Wiley Series in Remote Sensing and Image Processing, John Wiley & Sons.
- TACK, N., LAMBRECHTS, A., SOUSSAN, P. & HASPELAGH, L., 2012: A Compact, High-speed, and Low-cost Hyperspectral Imager. Proc. SPIE 8266, Silicon Photonics VII, 82660Q.
- ECKARDT, A., HORACK, J., LEHMANN, F., KRUTZ, D., DRESCHER, J., WHORTON, M. & SOUTULLO, M. 2015: DESIS (DLR Earth Sensing Imaging Spectrometer for the ISS-MUSES platform). IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 1457-1459.
- IMEC, 2014: Hyper Spectral Imaging. https://www.ximea.com/support/attachments/4675/XIMEA_imec_HSI_technology-Part-V1.1.pdf, Version 2014, letzter Zugriff am 29.04.2016

Crowd basierte Verortung historischer Fotografien und Konzepte zur verbesserten Aufarbeitung historischer Archive mit Hilfe geoinformatischer Konzepte – Methodischer Ansatz des EU Projektes PhotoStruk

WOLFGANG DORNER¹, ANNE WEINFURTNER¹, LIBOR DOSTÁLEK², KLÁRA PAČLÍKOVÁ³, KATEŘINA HORNÍČKOVÁ³ & MICHAL PREUSZ³

Zusammenfassung: Historische Fotoarchive bieten für die Raumwissenschaften ein großes Potenzial räumlich relevante historische Informationen zu gewinnen. Die Aufarbeitung dieser Archive stellt gleichzeitig eine große Herausforderung dar und bietet Chancen für den Einsatz geoinformatischer Konzepte und Technologien. Im Projekt PhotoStruk sollen Crowdsourcing basierte Verfahren herangezogen werden, um durch Georeferenzierung und Orientierung von Fotografien große Archivbestände zu erschließen und systematisiert auf Grundlage von Geoinformation zugänglich zu machen. Hierzu werden Aufnahmen des Ateliers Seidel aus dem Bayerisch-Böhmischen Grenzraum herangezogen und eine web- und app-basierte Methodik erprobt.

1 Einleitung

Historische Fotografien sind wertvolle Dokumente unserer Vergangenheit und gleichzeitig selbst Forschungsgegenstand. Die Größe der Archivbestände, deren stetes Wachstum durch Nachlässe und Leihgaben sowie der niedrige Digitalisierungsgrad bei Fotografien wie auch Dokumentation der Bestände sind ein Hindernis für die Nutzung dieser Daten.

Im Projekt „PhotoStruk - Analyse historischer PHOTOgraphien für die virtuelle RekonSTRUKtion von Kulturgütern in der Bayerisch-Böhmischen Grenzregion“ (Analýza historických FOTOgrafií pro virtuální rekonSTRUKci kulturního dědictví v česko-bavorském příhraničí) setzen sich die Südböhmische Universität Budweis und die Technische Hochschule Deggendorf in Kooperation mit dem Fotoatelier Seidel mit den Möglichkeiten der Aufarbeitung dieser Archivbestände auseinander. Das Projekt orientiert sich dabei an drei zentralen Fragen, die auch den Kern eines möglichen Lösungsansatzes darstellen. Zum ersten wird der Frage nachgegangen, wie digitalisierte Bestände historischer Fotografien langzeitarchiviert werden können. Die zweite Fragestellung setzt sich intensiv mit geoinformatischen und fotogrammetrischen Aspekten auseinander: Kann durch die Kombination von Experten und Crowdsourcing basierten Verfahren die Aufarbeitung von Archiven beschleunigt werden und kann dabei insbesondere die räumliche Perspektive auf die Fotografien einen besonderen Beitrag zur Erschließung leisten? Kern dieser Fragestellung ist zum einen der räumlich-perspektivische

¹ Technische Hochschule Deggendorf, Technologie Campus Freyung, Grafenauerstr. 22, D-94078 Freyung, E-Mail: [wolfgang.dorner, anne.weinfurtner]@th-deg.de

² Jihočeské univerzity v Českých Budějovicích, Přírodovědecká fakulta, Branišovská 1760, CZ-370 05 České Budějovice, E-Mail: dostalek@prf.jcu.cz

³ Jihočeské univerzity v Českých Budějovicích, Filozofická fakulta, Branišovská 31a, CZ 370 05 České Budějovice, E-Mail: [klara.paclikova, katerina.hornickova, preusz.michal]@gmail.com

Inhalt der Fotografien, zum anderen aber auch der Ansatz, ob eine räumliche Herangehensweise sowohl der Dokumentation durch Georeferenzierung der Fotografien und beinhalteten Objekte, als auch die räumliche Suche nach Fotografien und darin enthaltenen Objekten ein Ansatz ist, der unterschiedlichen Nutzergruppen entgegenkommt. Drittens stellt sich dann die Frage, ob dieser Ansatz der Aufarbeitung der Archive einen Beitrag leisten kann, umfangreiches (bis dato unerschlossenes) Material für die weiterführende Forschung, z.B. der Geschichtswissenschaften und Archäologie, zur Verfügung zu stellen.

Ziel des Projektes PhotoStruk, als Vorhaben der angewandten Forschung und Entwicklung, ist die Entwicklung Web und App gestützter Verfahren, um Fotografien der breiten Öffentlichkeit zugänglich zu machen und (insbesondere) Informationen zu Ort und abgebildeten Objekten durch Crowdsourcing-Verfahren zu erschließen. In Folge werden auf den gewonnen Daten Historiker und Archäologen Rekonstruktionen verlorener Landschaften, Ortschaften und Gebäude aufbauen. Der Beitrag stellt die grundlegenden Überlegungen des Projektes sowie die geplante Methodik vor. Anhand erster Beispiele und Daten werden diese Ansätze veranschaulicht.

2 Problemstellung und Herausforderung

Historische Fotografien sind als Dokument und Kulturgut in zweierlei Hinsicht wichtig. Sie sind zum einen Kulturgut per se und Gegenstand wissenschaftlicher Untersuchungen der Kunst- und Bildwissenschaften. Zum anderen dienen sie vielen Disziplinen als wichtiges Dokument bei der Erschließung historischer Landschaftssituationen, Siedlungs- und Baugeschichte sowie als Untersuchungs- und Planungsgrundlage für Denkmalschutz, Landschaftsplanung und Geomorphologie.

Gerade die perspektivische und räumliche Dimension einer Fotografie hat dabei besondere Bedeutung. Die meisten Fragestellungen der oben genannten Disziplinen betrachten Fotoarchive aus einer räumlichen Perspektive, da die Fragestellungen meist von konkreten räumlichen Objekten geleitet werden, die im Vordergrund der Untersuchung stehen. Die Verfügbarkeit von Aufnahmen für die entsprechenden Zwecke ist damit unter anderem von einer Erschließung des Archivs nach meist Ortsnamen oder (baulichen) Objekten abhängig. Alternative ist eine Sichtung des Archivs und der Versuch der Zuordnung der Fotografien zum konkreten Untersuchungsthema oder -gegenstand. Ein zweiter Aspekt ist die (nicht nur photogrammetrische) Auswertung der Fotografien aus perspektivischer Sicht, die zum Beispiel im Fall der Rekonstruktion von Landschaften im Sinne der Geomorphologie, Denkmalschutz, Archäologie oder Baugeschichte unterschiedliche Ansichten des selben Objekts als auch Darstellungsebenen von der Übersicht und räumlichen Kontext bis hin zu Detailaufnahmen erfordert. Die Verfügbarkeit und räumliche Erschließung historischer Fotografien bzw. der Archive ist damit eine wesentliche Grundvoraussetzung für weiterführende Forschungen.

Die große Menge an Fotografien in historischen Archiven lässt sich in ihrer Dokumentation und Aufarbeitung mit herkömmlichen Methoden nicht bewerkstelligen. Der Digitalisierungsgrad in diesen Archiven ist gering und insbesondere die Metadokumentation schwierig. Hierzu müssen Archiv-, Findebücher und historische Aufzeichnungen digitalisiert bzw. als Texte eingegeben werden. Bei Beständen von mehreren zehntausenden bis hunderttausenden von Fotografien pro

Archiv bzw. Sammlung ist dies mit aktuellen Personalressourcen in privaten, kommunalen und staatlichen Archiven nicht leistbar. Das Atelier Seidel, Projektpartner im vorgestellten Vorhaben, verfügt über einen Gesamtbestand von etwa 160.000 Fotografien des späten 19. Jahrhunderts bis in die 1940er. Davon sind derzeit 20.000 Fotografien digitalisiert, aber mit sehr unterschiedlichem Qualitätsgrad dokumentiert. Die dabei teilweise vorhandene oder vorgenommene Dokumentation, wie in Abb. 1 dargestellt, ist nicht ausreichend, um ein Foto korrekt zuordnen zu können. Das Bild der Ilz bei Passau mit Blick auf Veste Unterhaus, Innstadt und Donau bei Passau ist im Archiv mit den Schlagwörtern „Landschaft“, „Häuser“, „Aussicht“, „Fluss“ verschlagwortet. Hinweise auf Ortsnamen fehlen vollständig, da die Situation vermutlich einem Archivar oder Wissenschaftler im nur 50km entfernten Cesky Krumlov nicht geläufig ist. Die Ortskenntnis der regionalen Bevölkerung stellt aber in diesem Sinne ein Potenzial dar.



Abb. 1: Fotografie von Passau aus dem Bestand des Atelier Seidel (Quelle: <http://fotobanka.seidel.cz/#fotobanka/detail/104090401010012640001>)

Gleichzeitig setzen aktuelle Digitalisierungskonzepte bei der Systematisierung nach räumlichen Gesichtspunkten ggf. nur auf die Verortung des Bildursprungs oder die Angabe von Orts- und Objektnamen, z.B. als Teil der Bildbeschreibung. Die Auffindbarkeit für Anwender wird dabei verbessert, steht aber nicht in Relation zu den Möglichkeiten, die eine Georeferenzierung der Bilder bzw. eine systematische räumliche und perspektivische Aufarbeitung an Potenzialen bieten würde.

Das Beispiel der Ortschaft Kapellen in Südböhmen zeigt dabei die Herausforderung vor denen die Geschichtswissenschaft und Archäologie stehen. Die Ortschaft wurde in der Nachkriegszeit verlassen bzw. die Bevölkerung abgesiedelt. 1959 wurde die Kirche durch eine gezielte Sprengung zerstört. Zur Rekonstruktion der Siedlungsgeschichte werden durch die Archäologie neben Ergebnissen von Ausgrabungen nun auch historische Karten und Luftbilder (Abb. 2) sowie Fotografien herangezogen. Eine Abfrage unter dem Schlagwort Kaplicky (historisch tschechischer Name „Kapličky u Vyššího Brodu“) liefert im digitalisierten Archivbestand des Atelier Seidel fünf Fotografien. Zusätzliche Probleme bereiten dabei die Verschlagwortung unter

dem deutschen Namen „Kapellen bei Hohenfurth“, der allerdings von der historisch ebenfalls geläufigen Flur- und Ortsbezeichnung Kappeln abweicht. Unter der räumlich übergeordneten Ortsbezeichnung „Vyšší Brod“ finden sich ebenfalls weitere Aufnahmen. Eine Georeferenzierung der Fotografien, die eine Suche deutlich erleichtern würde, wurde nicht vorgenommen. Ein räumlich gestützte Suche, z.B. als Abfrage über eine Karte und Nähebeziehungen ist nicht verfügbar. Auch im Bild dargestellte Objekte (Kirche) werden in den Schlagworten nicht näher aufgeführt.



Abb. 2: Ansicht der Ortschaft Kappeln nach 1900 (links oben, Quelle: Atelier Seidel), bei der Sprengung der Kirche 1959 (rechts oben, Quelle: Kubištová (2014)) und heute (links unten, Quelle: Kubištová (2014)) sowie heute im Luftbild (rechts unten, Quelle: <http://kontaminace.cenia.cz>)

Archive stehen damit sowohl für den Erhalt als auch die Konservierung der Bestände vor der Herausforderung große Mengen an Fotomaterial zu digitalisieren und diese digitalisierten Archivbestände auch soweit zu erschließen, dass sie für Zwecke der Forschung oder Bereitstellung für die Öffentlichkeit genutzt werden können. Zu einem großen Problem wird dabei die fehlende Metadokumentation, da diese entweder in den Originaldokumentationen der Archive nicht vorhanden ist oder die Digitalisierung und Zuordnung handschriftlicher Archiv- und Findebücher einen deutlich höheren Aufwand erfordert, als die Digitalisierung der Fotografien. Etablierte Verfahren der Aufbereitung geraten in ihre Grenzen, da diese zu zeitintensiv sind bzw. die Budgets zu gering und die Bestände zu groß.

Für wissenschaftliche Zwecke ist damit nur ein Bruchteil der eigentlich verfügbaren Fotografien zugänglich. Oder Wissenschaftler müssen den Weg einer aufwändigen Recherche in den analogen Archiven gehen. Eine vollumfängliche dreidimensionale Abbildung historischer Bauwerke auch unter Berücksichtigung der zeitlichen Entwicklung ist damit nur schwer oder nicht darstellbar.

3 Hintergrund

Die Bewältigung des Aufkommens und Bestandes historischer Fotografien stellt eine Herausforderung für Archive dar (DIJCK, 2010). Gerade die Größe der Bestände (GREIPL, 2013) in Verbindung mit dem geringen Dokumentationsgrad sind als Hauptprobleme anzusehen. Vorhaben wie das Crowdsourcing-Vorhaben des Smithsonian Institutes (KALFATOVIC, KAPSALIS, SPIESS, CAMP, & EDSON, 2009) unter Verwendung der Fotoplattform Flickr oder das Projekt „Virtuelles Kartenforum 2.0“ (BILL, KOLDRACK, & WALTER, 2015) zeigen, dass crowdsourcing-basierte Verfahren sowohl in den Archivwissenschaften als auch der Geoinformatik das Potenzial haben, bei der Aufarbeitung von Archiven einen Beitrag zu leisten. Andererseits gibt es zahlreiche Disziplinen, wie Umweltwissenschaften (z.B. Kadmon & Harari-Kremer, 1999; Turner, 1990), Naturgefahrenmanagement (z.B. Chandler & Brunsten, 1995), Denkmalschutz (z.B. Yilmaz, Yakar, Gulec, & Dulgerler, 2007), Geschichtswissenschaften und Kunstgeschichte (z.B. Arias, Ordóñez, Lorenzo, & Herraes, 2006) oder Archäologie (z.B. Estes, Jensen, & Tinney, 1977), die auf historische Fotografien als Datenquelle zugreifen. Neben Verschlagwortung, Beschreibungen und Datumsangaben ist dabei auch die Georeferenzierung eine etablierte Methode der Metadokumentation von Fotografien (LU, PARK, & HU, 2010).

Auch auf den Websites des Ateliers Seidel gibt es eine Kommentarfunktion. Die ca. 3.500 Kommentare über einen Zeitraum von knapp sechs Jahren geben einen Hinweis darauf, dass es ein Interesse an historischen Fotografien gibt. Dabei sind in vielen Fällen die Kommentare nach Kuratierung durch das Archivpersonal oder Wissenschaftler in die Verschlagwortung eingeflossen. Hierbei sind insbesondere viele Ortsnamen auffällig. Auch regionale Crowdsourcing Vorhaben, wie die Facebookgruppe “Das alte Passau” und “Du kommst aus Passau, wenn ...” haben in den vergangenen Jahren mehrere hundert historische Fotografien und Ansichtskarten gesammelt und kommentiert und unterstützen damit die Annahme, dass die Bereitstellung entsprechender Funktionen durch die breite Öffentlichkeit wahr- und angenommen wird.

4 Vorgehen

Das Projekt basiert auf der Annahme, dass durch Georeferenzierung und sogar eine erste Bildorientierung eine hinlängliche Dokumentation des Archivmaterials gelingen kann, die dabei hilft, Disziplinen, die stark von historischen Aufnahmen abhängig sind, einen schnelleren und zielgerichteteren Zugang zu Archivmaterial zu ermöglichen. Das Projekt PhotoStruk setzt bei der Aufarbeitung von Fotoarchiven auf Crowdsourcing basierte Verfahren. Das Vorgehen beruht auf der Hypothese, dass sich durch Crowdsourcing die Kenntnisse regionaler Experten, Zeitzeugen und fachlich interessierte Personenkreise erschließen lassen.

In der Herangehensweise (Abb. 3) setzt PhotoStruk auf eine web- und smartphone-basierte Lösung, die es Zeitzeugen und regionalen Experten erlauben soll, den Aufnahmen Ortsnamen zuzuweisen, dargestellte Objekte in Bildern zu georeferenzieren bzw. mit Augmented Reality die Aufnahmesituation und den Bildursprung historischer Fotografien zu rekonstruieren, indem die historische Fotografie in das Kamerabild des Smartphones eingepasst wird. Hierbei soll insbesondere untersucht werden in welchem Umfang ein Archivbestand aufbereitet und welche

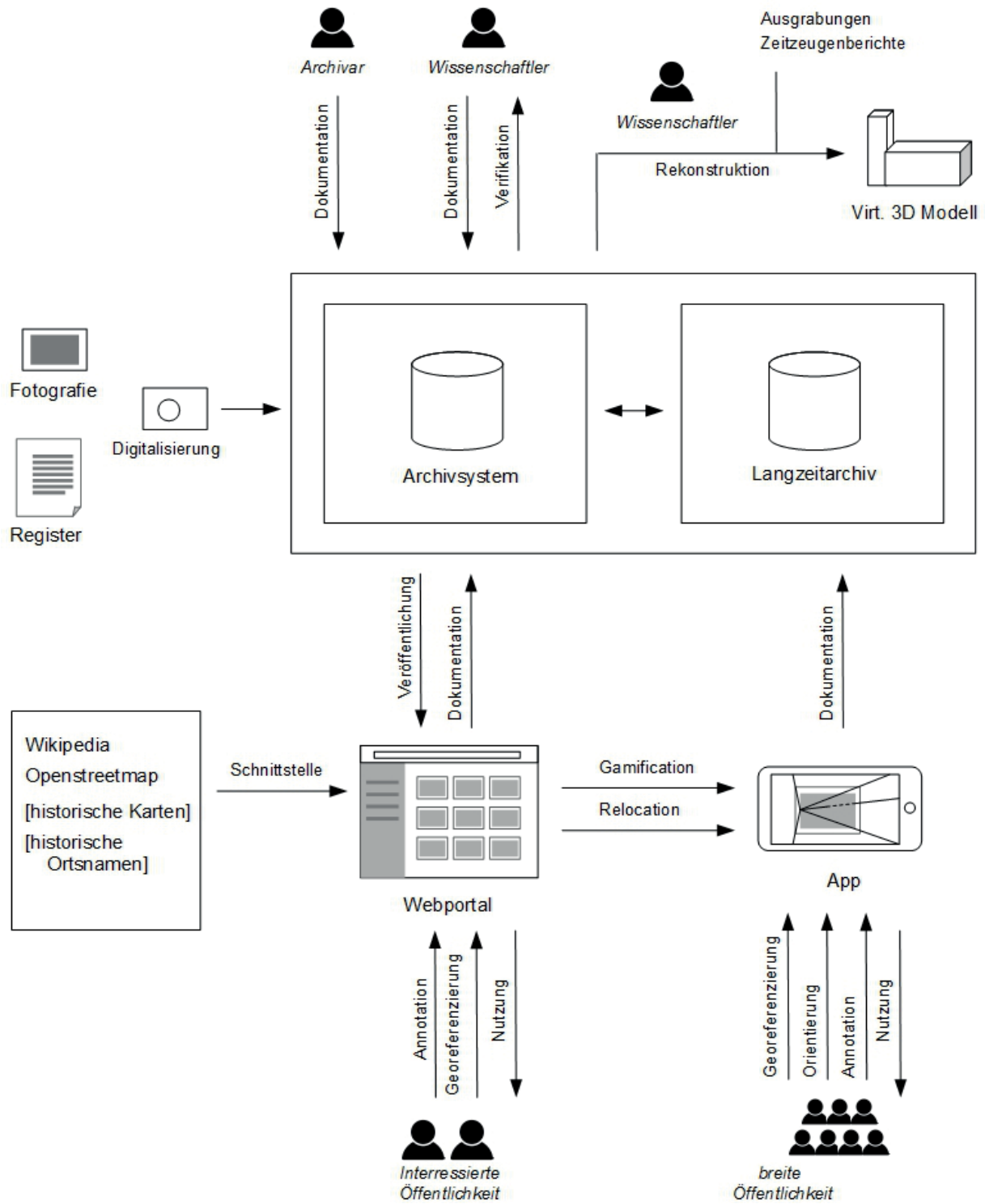


Abb. 3: Vorgehensmodell/Konzept

Genauigkeit bei Georeferenzierung und Bildorientierung erreicht werden kann. Dabei sollen gezielt unterschiedliche Möglichkeiten der Verortung eines Bildes parallel herangezogen werden, um unterschiedliche Herangehensweisen bei der Verortung zu vergleichen bzw. zu analysieren, in wie weit ein Crowdsourcing gestütztes Verfahren zu einer laufenden Verbesserung beitragen kann, in dem durch fortlaufende und mehrfache Erfassung durch unterschiedliche Personen Lagemaße und Genauigkeiten sukzessive verbessert werden.

Digitalisierte Aufnahmen werden dazu aus einem Archivsystem heraus über eine Webplattform zugänglich gemacht. Diese erlaubt der interessierten Öffentlichkeit den Zugriff auf die Daten und die Interaktion in Form von Kommentaren und Annotationen, Georeferenzierung sowie Verknüpfung mit anderen Datenquellen wie z.B. (historischen) Ortsnamen und lexikalischen Artikeln. Auch Smartphone basierte Ansätze sowie Gamification sollen als Möglichkeiten der Nutzerinteraktion überprüft werden. Hierbei wird insbesondere als Ansatz eine „Digitale Relocation“ erprobt, also den Anwendern eine Interaktion mit der Fotografie im Umfeld ihrer Entstehung (Aufnahmeort, abgebildete Bauobjekte/Landschaften) zu ermöglichen. Hierzu werden in einem Location Based Service den Anwendern bereits georeferenzierte Aufnahmen im originären Kontext präsentiert.

Für Bilddaten soll dabei ein Ansatz erprobt werden, der eine weiterführende Georeferenzierung und darauf basierend Prozessierung erlaubt. Dazu werden neben dem Standort des Fotografen auch die Standorte und Ortsnamen von Objekten erfasst und georeferenziert, die im Bild abgebildet sind (Abb. 4). Dies soll eine spätere objekt- und ortsbasierte Suche vereinfachen, die nach Annahme im Projekt vielen Suchanfragen sowohl der interessierten Öffentlichkeit als auch von Fachanwendern zu Grunde liegt.

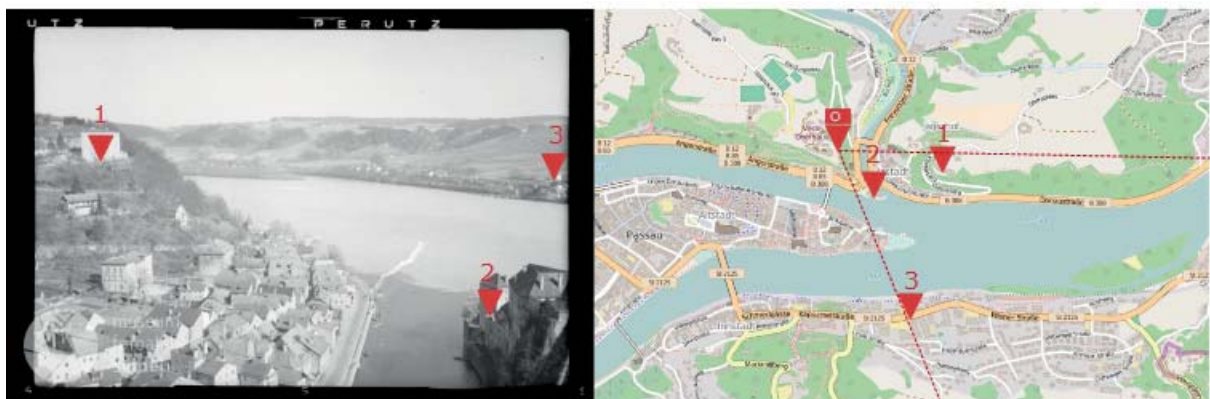


Abb. 4: Ansatz zur Georeferenzierung und groben Bildorientierung durch die Erfassung des Standorts des Fotografen als auch abgebildeter Objekte

Mit Hilfe von Workshops werden in Zusammenarbeit mit Anwendern der Archäologie, Denkmalpflege und Landschaftsplanung Methoden der räumlichen Suche in Archivbeständen im Sinne neuer Herangehensweisen für die Archivwissenschaften sowie Ansätze zur Präsentation der Fotografien auf Basis von Virtual und Augmented Reality sowie kartografischer Ansätze untersucht, um eine bessere Zugänglichkeit für die Nutzergruppen der Archive zu erreichen. Wissenschaftler werden in diesem Kontext sowohl als Nutzer der Ergebnisse gesehen, können aber auch einen Beitrag durch Verifikation der Ergebnisse leisten.

Das Projekt setzt sich aus acht inhaltlichen Arbeitspaketen zusammen (Abb. 5). Insbesondere soll dabei das Zusammenspiel von Ansätzen der “Georeferenzierung und Orientierung” sowie “Visualisierung und Nutzerinteraktionen” auf Webplattformen und in Smartphone Anwendungen in Verbindung mit Öffentlichkeitsarbeit iterative getestet werden. Hier steht im Vordergrund ob und wie die Aufmerksamkeit der Öffentlichkeit gewonnen werden und dieses Interesse zur Aufarbeitung der Archive zielführend eingesetzt werden kann.

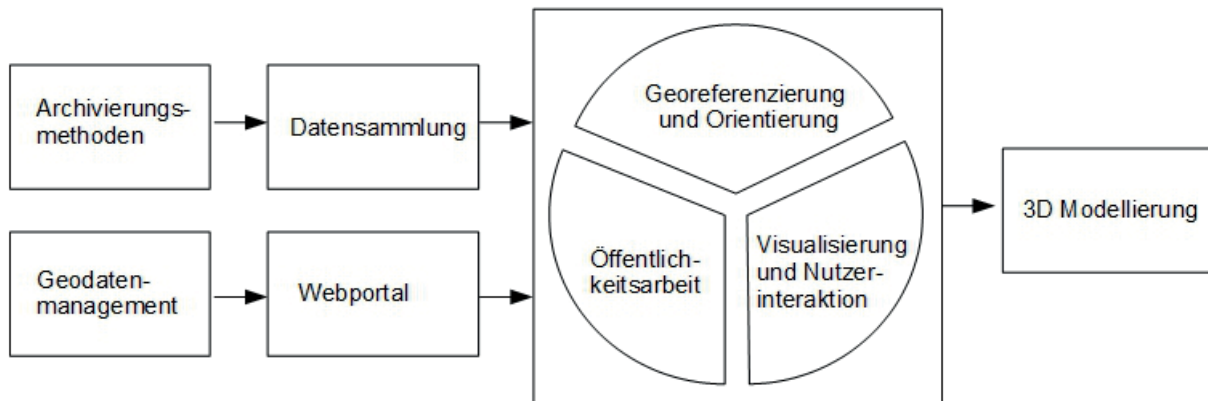


Abb. 5: Projekttablauf und Struktur

5 Geplante Fallstudien

Für die bayerisch-böhmische Grenzsituation werden hierzu undokumentierte, aber digitalisierte Archivbestände herangezogen. Hierzu werden Teile des Fotoarchivs des Ateliers Seidel der Öffentlichkeit zugänglich gemacht. Das Atelier Seidel war ein Fotoatelier im Südböhmischen Krumau und ist heute als Museum und Archiv Träger der Sammlung. Das Archiv beherbergt eine Sammlung von ca. 160.000 Fotografien, die von Seidel Vater und Sohn zwischen ca. 1890 und 1949 mit Schwerpunkt in der bayerisch-böhmischen Grenzregion gemacht wurden. Von 160.000 Fotografien der Sammlung wurden in den vergangenen Jahren 20.000 digitalisiert und verschlagwortet. Die Qualität der Verschlagwortung ist dabei sehr unterschiedlich und führt, wie oben bereits beschrieben, oft nur auf Umwegen zu einem Ergebnis.

Im Projekt sollen dabei im Schwerpunkt Landschafts- und Ortsbilder untersucht werden, die infolge Kollegen aus Archäologie und Geschichtswissenschaften dazu dienen sollen untergegangene Siedlungen und Bauwerke im Sudetenland zu identifizieren und zu rekonstruieren.

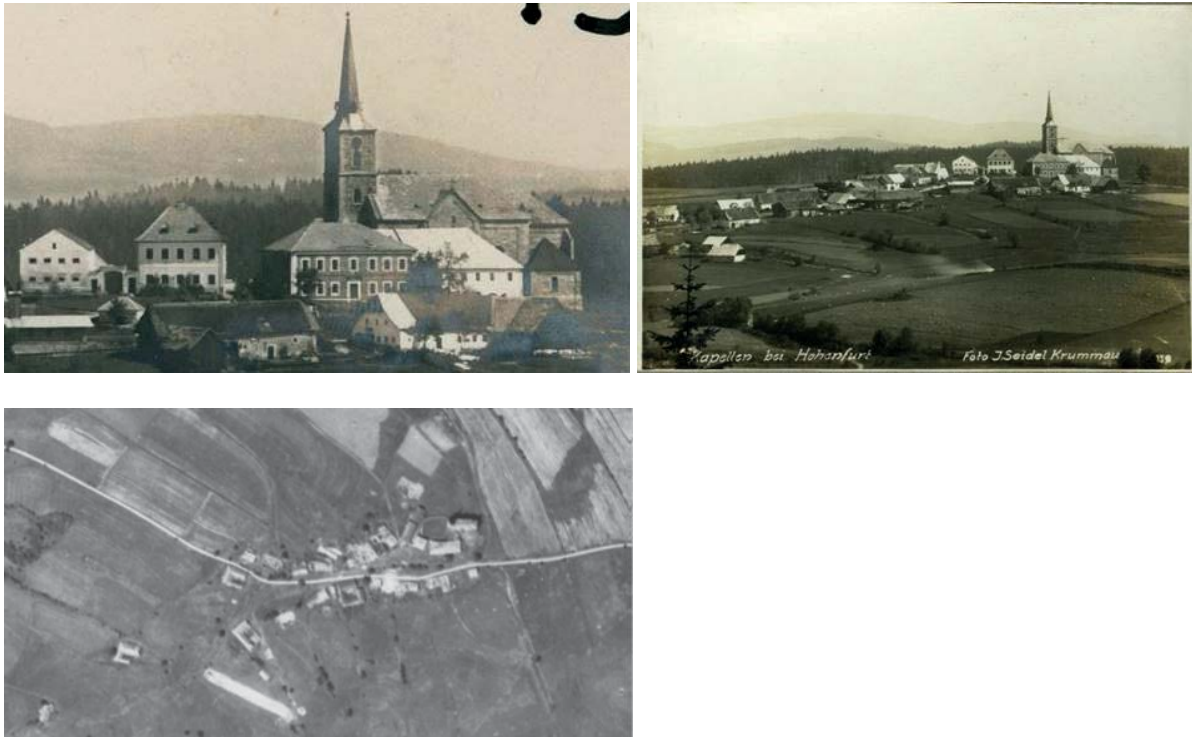


Abb. 6: Hauptansichten der Ortschaft Kapellen bei Hohenfurt (o.l. und o.r. Quelle: Atelier Seidel) in Verbindung mit einer historischen Luftbildaufnahme (Unten Quelle: <http://kontaminace.cenia.cz/>)

Die Ortschaft Kaplice/Kapellen, für die bereits erste Daten vorliegen (Abb. 6), zeigt deutlich das Dilemma von heute. Es sind nach aktuellem Stand zwar hinlänglich Aufnahmen verfügbar, die einem geschulten Betrachter dabei unterstützen können Rekonstruktionen der damaligen Siedlung zu erstellen. Für eine dreidimensionale digitale Rekonstruktion sind diese aber nicht ausreichend, da sie meist Hauptblickrichtungen auf die gesamte Ortschaft bieten.



Abb. 7: Historische digitalisierte Gebäudegrundrisse der Ortschaft Kaplice auf einem aktuellen Luftbild (Links Quelle: [on ags.cuzk.cz](http://ags.cuzk.cz)), auf einem aktuellen Geländemodell (Mitte Quelle: Tschech. Vermessungsverwaltung) und einer Überlagerung von beidem.

Es besteht die begründete Vermutung, dass auch von innerhalb des Ortes Detailaufnahmen, z.B. auf Gruppenaufnahmen und Architektur fotografien vorliegen, diese konnten aber bis dato nicht identifiziert werden.

Die Möglichkeiten der digitalen Analyse sind damit bis dato auf den Verschnitt historischer Daten mit aktuellen Geodaten beschränkt (Abb. 7), die bereits für Archäologen und Historiker in Verbindung mit Fotografien Anhaltspunkte für Grabungen und weiterführende Auswertungen bieten bzw. als Basis für eine erste räumliche Orientierung dienen (Abb. 8).

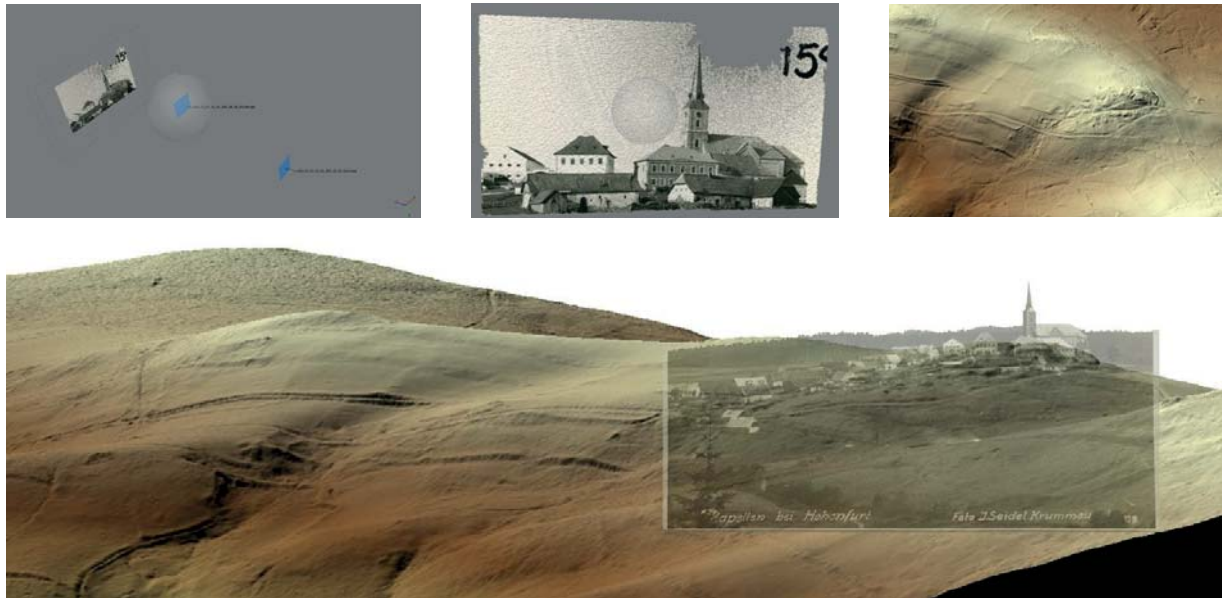


Abb. 8: Erste Ansätze zur Rekonstruktion von Landschaft und Bauwerken durch Georeferenzierung von historischen Fotografien (o.l. und o.m.) in Verbindung mit einem digitalen Geländemodell (o.r.) und Überlagerung in einer perspektivischen Rekonstruktion

6 Fazit und Ausblick

Eine erste qualitative Auswertung des Archivmaterials des Ateliers Seidel hat gezeigt, dass darin ein wissenschaftliches Potenzial liegt und Dokumente verfügbar wären, um wissenschaftliche Fragestellungen zur Fotografie als auch weitergehende wissenschaftliche Arbeiten in der bayerisch-böhmischen Grenzregion zu unterstützen. Obwohl digital verfügbar, ist bis dato die Auswertung des digitalisierten Archivmaterials mit einem erheblichen Rechercheaufwand verbunden. Gleichzeitig ist das Material nicht ausreichend, um zum Beispiel eine dreidimensionale Rekonstruktion zerstörter oder teilweise zerstörter Ortschaften und Gebäude zu erlauben. Auch eine zeitliche Analyse der Entwicklung von Landschaften und Siedlungen ist alleine auf dieser Datengrundlage nicht darstellbar. Die verfügbaren Kommentare, meist anonymer Nutzer, auf der Website des Archivs zeigen, dass bereits in der Vergangenheit (anscheinend regionale Experten und Zeitzeugen) einen Beitrag zur Aufarbeitung und Dokumentation des Archivmaterials geleistet haben, da viele Kommentare nach einer Verifikation durch die Archivmitarbeiter direkt Eingang in die Verschlagwortung der zugehörigen Aufnahmen gefunden haben. Umfang und Anzahl der Kommentare könnten aber

gleichzeitig Indikator dafür sein, dass die bis dato zur Verfügung gestellten Funktionen nicht ausreichend sind, um eine breitere Gruppe an Personen einzubinden.

In einem nächsten Schritt sollen nun die vorhandenen Erkenntnisse, wie Crowdsourcing für die Annotation und Georeferenzierung historischer Fotografien herangezogen werden und können in einer ersten Testanwendung implementiert und in Verbindung mit weiterführenden Ansätzen erprobt werden.

Hinweis: Das Projekt PhotoStruk ist ein gemeinsames Projekt der Südböhmischen Universität Budweis und der Technischen Hochschule Deggendorf in Zusammenarbeit mit dem Atelier Seidel aus Cesky Krumlov. Das Projekt wird gefördert durch die Europäische Union Ziel ETZ Freistaat Bayern – Tschechische Republik 2014 – 2020 (Interreg V) durch den Europäischen Fonds für regionale Entwicklung.

7 Literaturverzeichnis

- ARIAS, P., ORDÓÑEZ, C., LORENZO, H., & HERRAEZ, J., 2006: Methods for documenting historical agro-industrial buildings: a comparative study and a simple photogrammetric method. *Journal of Cultural Heritage* **7**(4), 350-354.
- BILL, R., KOLDRACK, N., & WALTER, K., 2015: Georeferenzierung alter topographischer Karten – Crowdsourcing versus Bildverarbeitung. *AGIT – Journal für Angewandte Geoinformatik*, (1), 540-550.
- CHANDLER, J. H., & BRUNSDEN, D., 1995: Steady state behaviour of the black ven mudslide: The application of archival analytical photogrammetry to studies of landform change. *Earth Surface Processes and Landforms* **20**(3), 255-275.
- DIJCK, J. VAN., 2010: Flickr and the culture of connectivity: Sharing views, experiences, memories. *Memory Studies* **4**(4), 401-415.
- ESTES, J. E., JENSEN, J. R., & TINNEY, L. R., 1977: The Use of Historical Photography For Mapping Archaeological Sites. *Journal of Field Archaeology* **4**(4), 441-447.
- GREIPL, E. J., 2013: Bilder-Welten: Das Fotoarchiv des Bayerischen Landesamtes für Denkmalpflege. *AVISO - Zeitschrift für Wissenschaft u und Kunst in Bayern*, (1), 38-43.
- KADMON, R., & HARARI-KREMER, R., 1999: Studying Long-Term Vegetation Dynamics Using Digital Processing of Historical Aerial Photographs. *Remote Sensing of Environment*, **68**(2), 164-176.
- KALFATOVIC, M. R., KAPSALIS, E., SPIESS, K. P., CAMP, A. V., & EDSON, M., 2009: Smithsonian Team Flickr: a library, archives, and museums collaboration in web 2.0 space. *Archival Science* **8**(4), 267-277.
- LU, C., PARK, J., & HU, X., 2010: User tags versus expert-assigned subject terms: A comparison of LibraryThing tags and Library of Congress Subject Headings. *Journal of Information Science* **36**(6), 763-779.
- TURNER, M. G., 1990: Landscape changes in nine rural counties in Georgia. *Photogrammetric Engineering & Remote Sensing* **56**(3), 379-386.
- YILMAZ, H. M., YAKAR, M., GULEC, S. A. & DULGERLER, O. N., 2007: Importance of digital close-range photogrammetry in documentation of cultural heritage. *Journal of Cultural Heritage* **8**(4), 428-433.

Segmentierung großer Punktwolken mittels Region Growing

MARKUS PÖCHTRAGER¹

Zusammenfassung: Die stetig wachsenden Datensätze mit 3D-Punktwolken gewonnen durch terrestrisches Laserscanning und Airborne Laserscanning bieten einerseits vielfältige Möglichkeiten in der Datenanalyse und Datenprozessierung, erfordern andererseits jedoch eine hohe Effizienz der prozessierenden Algorithmen. Die Segmentierung von Punktwolken liefert eine Gruppierung von gleichartigen Punkten anhand eines Homogenitätskriteriums und bildet damit eine geeignete Grundlage für viele Prozessierungsketten. Dabei stehen alle Punkte über die Homogenität in Beziehung zueinander und müssen gemeinsam prozessiert werden. Um auch Datensätze behandeln zu können, die für eine Verarbeitung im Ganzen zu groß sind, wurde im Rahmen dieser Arbeit ein Konzept für eine Segmentierung von solchen großen Punktwolken vorgestellt, bei dem diese Kette zwischen den Punkten aufgebrochen wird. Für eine effiziente Prozessierung wird die Punktwolke zuerst in rechteckige Teilpunktwolken (Kacheln) aufgeteilt. Diese Kacheln werden unabhängig voneinander segmentiert und abschließend werden gleichartige Segmente aus benachbarten Kacheln zusammengeführt. Dabei wird auch die Punktwolke wieder vereint. Ziel dieses kachelbasierten Konzeptes ist auch, eine Möglichkeit der Parallelisierung zu schaffen. Wie diese Arbeit zeigt, bleiben die Ergebnisse von der Wahl der Kachelgröße unbeeinflusst und sind nur vom Homogenitätskriterium abhängig. Durch beliebige Wahl der Kachelgröße und parallele Prozessierung können Speicherbedarf und Laufzeit optimiert und die Anforderungen an den verarbeitenden Rechner reduziert werden.

1 Einleitung

Während Laserscanning zur Aufnahme von dreidimensionalen Punktwolken in verschiedensten Fachgebieten zunehmend Verbreitung findet, steigen die Anforderungen an Prozessierungs- und Auswertemethoden für die gewonnenen Punktwolken. Zum einen weil immer komplexere Prozessierungsketten entwickelt werden, zum anderen durch die immer größer werdenden Datenmengen. Für die automatisierte Prozessierung von großen Datenmengen werden Methoden, welche die Speicherung und Verwaltung von Punktwolken vereinfachen, benötigt. Dabei ist man unter anderem an Verfahren zur Datenkompression interessiert, um den Speicherbedarf zu reduzieren, aber auch an der Entwicklung von Speicherstrukturen, welche einen schnellen Zugriff auf die Punktinformationen ermöglichen. Es stellt sich hier die Frage, wie Strukturinformation in Punktwolken gefunden werden kann, um einen effizienteren Zugriff auf gleichartige Punkte zu ermöglichen. Wie VOSSELMAN et al. 2004 argumentiert, liefert die Segmentierung einen erheblichen Beitrag für diese Strukturfindung in Punktwolken. Bei der Segmentierung wird ein Datensatz auf Basis eines Homogenitätskriteriums in Segmente von Punkten mit gleicher Eigenschaft unterteilt. Das dieser Arbeit zugrunde liegende Segmentierungsverfahren *Seeded Region Growing* (ADAMS & BISCHOF 1994) ist Standard, aber durch den Arbeitsspeicherplatz beschränkt anwendbar. Für die Lösung dieses Problems wurde ein Konzept entwickelt, bei dem

¹ TU Wien, Department für Geodäsie und Geoinformation; Gußhausstraße 27-29, A-1040 Wien
E-Mail: markus.poechtrager@tuwien.ac.at

die Punktwolke in Teilpunktwolken (Kacheln) mit beliebiger Größe unterteilt werden kann. Die Segmentierung erfolgt dabei auf den Kacheln und es muss dadurch nie die gesamte Punktwolke im Arbeitsspeicher gehalten werden, sondern immer nur die gerade verarbeiteten Kacheln.

2 Konzept

Das Konzept dieser Arbeit basiert auf einem *Split-and-Merge* Ansatz. Dafür wird die Punktwolke im ersten Schritt in Kacheln unterteilt (*Split*). Alle Kacheln zusammen bilden eine lückenlose und überlappungsfreie Abdeckung der gesamten Punktwolke. Im nächsten Schritt wird auf jeder dieser Kacheln eine von anderen Kacheln unabhängige Segmentierung durchgeführt. Abschließend werden die Punkte aus den segmentierten Kacheln wieder zu einer Punktwolke vereint. Gleichartige Segmente – also Segmente, bei denen zumindest zwischen zwei Punkten das Homogenitätskriterium erfüllt ist – aus benachbarten Kacheln werden dabei zusammengeführt (*Merge*). Ein Ziel dieses Konzeptes muss sein, dass die Ergebnisse dieser *Split-and-Merge* Segmentierung den Ergebnissen einer Segmentierung auf der gesamten Punktwolke entsprechen (CHEN & PAVLIDIS 1990). Das Segmentierungsergebnis soll also von der gewählten Kachelgröße unabhängig sein.

2.1 Datenorganisation

Das Konzept wurde auf Basis der Software OPALS (PFEIFER et al. 2014) entwickelt, kann aber grundsätzlich auch auf anderen Datenstrukturen etabliert werden. Der verwendete *OPALS Data Manager* übernimmt die Aufteilung und persistente Speicherung der Punktwolke in Kacheln und verwaltet außerdem für jeden Punkt eine Liste mit Attributinformationen. Ist eine Kachel der Punktwolke im Arbeitsspeicher geladen, können diese Punktattribute – wie z.B. Intensität, Farb-Kanäle oder die Segment-ID – gelesen oder geschrieben werden. Beim Laden einer Kachel wird mit den darin liegenden Punkten *on-the-fly* ein *kd*-Baum im Arbeitsspeicher aufgebaut. Die Datenstruktur des *kd*-Baumes liefert einen schnellen Index für Punkte im *k*-dimensionalen Raum und ermöglicht eine effiziente Nachbarschaftssuche in der Punktwolke (BENTLEY 1975). Gerade bei der Segmentierung mit *Seeded Region Growing* spielt die Suche von benachbarten Punkten für einen Ausgangspunkt (Saatpunkt) eine entscheidende Rolle.

2.2 Segmentierung mit Seeded Region Growing

Die Segmentierung mit *Seeded Region Growing* teilt die Punktwolke von einem Saatpunkt ausgehend in disjunkte Segmente ($S_i \cap S_j = \emptyset$, für $i \neq j$). Die Vereinigung aller Segmente S_i ist dabei wieder die ursprüngliche Punktmenge. Für die Segmentierung auf den Kacheln werden zunächst alle Punkte einer Kachel geladen und mit ihnen der *kd*-Baum aufgebaut. Anschließend werden alle Punkte p_i mit $i = 1, \dots, m$ in der Baumstruktur durchlaufen. Wurde ein Punkt p_i noch keinem Segment zugewiesen, wird ein neues Segment S_i mit p_i als Saatpunkt angelegt. Für jeden Saatpunkt p_i eines Segmentes wird eine Nachbarschaftssuche ausgeführt. Die Nachbarschaftssuche liefert eine Liste von Punkten p_n mit $n = 1, \dots, m$ die ein symmetrisches Nachbarschaftskriterium *NH* erfüllen (z.B.: Punkte innerhalb einer Kugel mit Suchradius r). Für die Nachbarschaft *NH* zweier Punkte muss folgende Bedingung erfüllt sein:

$$NH(p_i, p_n) == NH(p_n, p_i)$$

Ist ein Punkt p_n ein Nachbarpunkt von p_i , so muss auch p_i ein Nachbarpunkt von p_n sein.

Erfüllt ein benachbarter Punkt p_n zusätzlich auch ein gewähltes Homogenitätskriterium P , wird dieser Punkt zum Segment S_i hinzugefügt und reiht sich außerdem in eine Liste zukünftiger Saatpunkte des Segmentes SQ_i ein. Das Homogenitätskriterium muss ebenfalls symmetrisch sein, damit dieses *Split-and-Merge* Konzept unabhängig von der Wahl der Kacheln immer die gleichen Ergebnisse liefert. Auch hier gilt also folgendes:

$$P(p_i, p_n) == P(p_n, p_i)$$

Erfüllt ein Punkt p_n das Homogenitätskriterium zu p_i , so muss auch p_i das gleiche Kriterium zu p_n erfüllen.

Nachdem alle Nachbarpunkte von p_i auf Homogenität überprüft und gegebenenfalls zum Segment hinzugefügt wurden, wird aus der Liste SQ_i der nächste Saatpunkt des Segmentes S_i gewählt. Das Segment wächst dadurch in alle Richtungen, bis keine neuen Punkte das Homogenitätskriterium erfüllen. Danach wird das Segment abgeschlossen und mit dem nächsten Punkt p_i aus dem *kd*-Baum, welcher noch nicht prozessiert wurde, ein neues Segment initialisiert. Die Segmentierung erfolgt auf jeder Kachel also sequentiell. Ein neues Segment wird erst dann angelegt, wenn das vorherige abgeschlossen ist.

Während bei *Seeded Region Growing* eine Parallelisierung auch über die Verwendung von n Saatpunkten, von denen ausgehend die Segmente parallel wachsen, denkbar wäre, wird sie in diesem Konzept über die unabhängige Segmentierung auf den n Kacheln erreicht.

2.3 Zusammenführung der Kacheln

Um das erwartete Segmentierungsergebnis auf der gesamten Punktwolke zu bekommen, müssen abschließend die gleichartigen Segmente aus den Kacheln vereint werden. Für eine effiziente Lösung dieser Aufgabenstellung wird eine Nachbarschaftsbeziehung zwischen den Kacheln aufgebaut. Dafür wird jede Kachel in neun Teilbereiche unterteilt (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Neben dem inneren Bereich (0), in der jene Punkte liegen, die für die Zusammenführung keine Relevanz haben, gibt es acht Randbereiche (1)-(8). Die Größe der Randregionen ergibt sich aus dem Suchradius r der Nachbarschaftssuche in der Segmentierung. Alle Punkte, die bei der Segmentierung in zumindest einer Nachbarkachel eine Rolle spielen können, liegen dadurch in Randregionen.

Für jede Kachel (K0) werden außerdem die angrenzenden Kacheln ermittelt (K1-K8) und mit benachbarten Kachelregionen sogenannte *Merge*-Nachbarschaften gebildet. Im Wesentlichen werden 4er- (**Fehler! Verweisquelle konnte nicht gefunden werden.** – *rot*) und 6er-Nachbarschaften (**Fehler! Verweisquelle konnte nicht gefunden werden.** – *grün* und *blau*) unterschieden. Die beteiligten Randbereiche in einer *Merge*-Nachbarschaft können über die geschaffene Index-Struktur, bestehend aus Kachel-ID und Randbereichs-ID, schnell ermittelt werden:

z.B.: $[K0, \text{Region } 1] \rightarrow [[K8, \text{Region } 3], [K1, \text{Region } 5], [K2, \text{Region } 7]]$

(1)	(2)	(3)
(8)	(0)	(4)
(7)	(6)	(5)

Abb. 1: Kachel-Regionen

K1	K2	K3
(5) (7)	(6)	(5) (7)
(3) (1)	(2)	(3) (1)
(4) (8)	K0	(4) (8)
(5) (7)	(6)	(5) (7)
(3) (1)	(2)	(3) (1)
K7	K6	K5

Abb. 2: Kachel-Nachbarschaft

Im tatsächlichen Zusammenführungsschritt werden dann alle Kacheln der Punktwolke durchlaufen. In jeder Kachel werden die Randbereiche (Kachel-Region 1-8) unabhängig voneinander - parallel oder sequentiell - geladen. Für jede Kachel-Region werden außerdem die relevanten Nachbarregionen geladen und mit allen Punkten dieser *Merge*-Nachbarschaft ein lokaler *kd*-Baum aufgebaut. Mit den Punkten im *kd*-Baum wird eine zweite, lokale Segmentierung durchgeführt. Dabei werden zwei Segmente zusammengeführt wenn ein Punkt p_i und sein benachbarter Punkt p_n innerhalb des Suchradius folgende Bedingungen erfüllen:

- i. *Der Punkt p_n ist noch nicht prozessiert*
- ii. *Die Punkte p_i und p_n liegen nicht im gleichen Segment*
- iii. *Die Segmente der Punkte p_i und p_n besitzen noch keine Zuordnung (Mapping) zueinander*
- iv. *Das Homogenitätskriterium $P(p_i, p_n)$ wird erfüllt*

Alle Segmente, die durch die Zusammenführung zu einem Segment vereint wurden, werden in einem Mapping-Eintrag zusammengefasst. Jeder Eintrag in der Mapping-Liste beinhaltet also zwei oder mehr Segment-IDs. Über diese Mapping-Liste können die endgültigen Segment-IDs für die Punkte bestimmt werden.

3 Ergebnisse

Die Tauglichkeit des Konzeptes wird im Rahmen der Arbeit mit unterschiedlichen Testdatensätzen geprüft. Wir erwarten dabei, dass die Ergebnisse von der Wahl der Kachelgröße unbeeinflusst bleiben. Zusätzlich wird der Einfluss der Kachelgröße auf die Laufzeit analysiert.

3.1 Segmentierungsergebnis

Der für den anschaulichen Vergleich der Segmentierungsergebnisse bei unterschiedlichen Kachelgrößen verwendete, synthetische Testdatensatz *Ring (100x100 m)* besteht im Wesentlichen aus drei Teilbereichen. Einem inneren kreisförmigen Bereich, einem in z-Achse um 1 m angehobenen Ring, sowie dem äußeren Randbereich. Eine Segmentierung mit dem einfachen Homogenitätskriterium

$$P(p_i, p_n) = |p_{i.z} - p_{n.z}| < 0.01$$

wird für verschiedene Kachelgrößen (10 m, 50 m, 100 m) durchgeführt und das Ergebnis zahlenmäßig (Tab. 1) und visuell (**Fehler! Verweisquelle konnte nicht gefunden werden.** und **Fehler! Verweisquelle konnte nicht gefunden werden.**) analysiert.

Die unterschiedlichen Farbwerte in der Punktwolke repräsentieren die den Punkten zugeordneten Segment-IDs. Alle drei Segmentierungsergebnisse zeigen die erwarteten, gleichen Segmente. Zur Veranschaulichung der unterschiedlichen Berechnungsvorgänge sind in **Fehler! Verweisquelle konnte nicht gefunden werden.** die Segmente vor der Zusammenführung zu sehen.

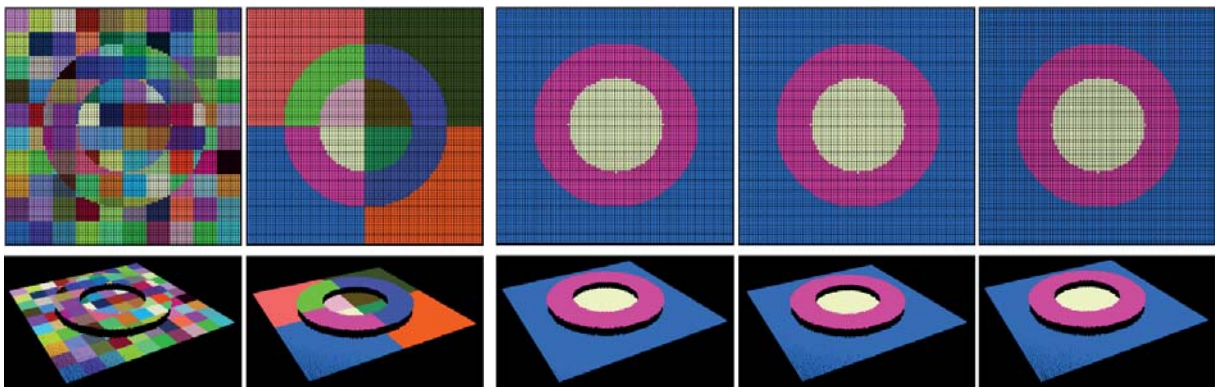


Abb. 3: Segmentierungsergebnisse vor der Zusammenführung; Kachelgrößen 10m, 50m; Grundriss und 3D

Abb. 4: Segmentierungsergebnisse nach der Zusammenführung; Kachelgrößen 10m, 50m und Segmentierung ohne Kachelung; Grundriss und 3D

Tab. 1: Einfluss der Kachelgröße auf Segmente – Datensatz Ring

Kachelgröße [m]	Anz. d. Kacheln	Segmente vor Merge	Segmente Final
10	100	140	3
50	4	12	3
100	1	-	3

3.2 Laufzeitanalyse

Unter der Annahme, dass keine signifikanten Laufzeitunterschiede zwischen der Segmentierung des gesamten Datensatzes und der verteilten Segmentierung auftreten, wurden auf einem herkömmlichen Business-Notebook (Intel® Core™ i5 M560 @ 2.67 GHz, 4 GB RAM) Laufzeittests durchgeführt. Obwohl der Prozessor eine Parallelisierung auf 4 Threads ermöglicht,

werden dabei die Berechnungen in einem einzelnen Thread ausgeführt. Eine Parallelisierung würde sowohl die Segmentierung, als auch die Zusammenführung erheblich beschleunigen. An dieser Stelle sind wir jedoch am direkten Laufzeitvergleich interessiert. Für die Laufzeitanalyse wird eine ALS Punktwolke mit etwa 2.5 Millionen Punkten, aufgenommen bei einer Befliegung im Naturschutzgebiet *Zonser Grind* am Niederrhein, verwendet.

Tab. 2: Einfluss der Kachelgröße auf die Laufzeit – Datensatz *Niederrhein*

Kachelgröße [m]	Anz. d. Kacheln	Segmente vor Merge	Segmente Final	Laufzeit Segmentierung auf Kacheln [s]	Laufzeit Zusammenführung [s]	Laufzeit Gesamt [s]
75	90	26409	1019	1309	1075	2384
100	56	21012	1019	1234	775	2009
125	36	19256	1019	1310	506	1816
150	30	15456	1019	1317	559	1876
200	20	12960	1019	1259	402	1661
250	12	9575	1019	1316	212	1528

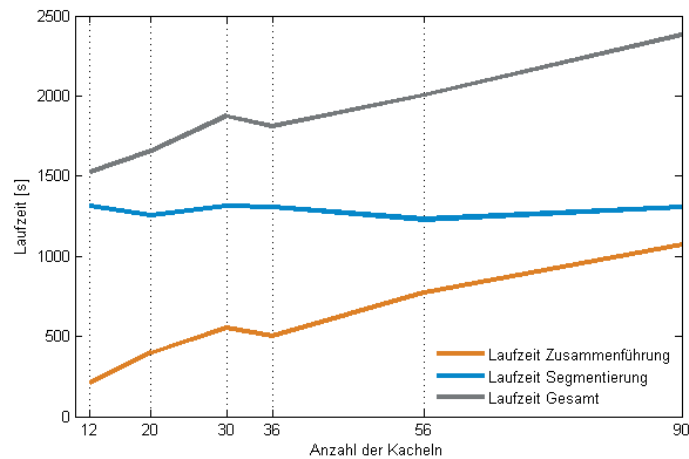


Abb. 5: Laufzeitvergleich, Segmentierung bei unterschiedlichen Kachelgrößen - Datensatz *Niederrhein*

Die Ergebnisse der Laufzeitanalyse zeigen, dass die Kachelgröße – wie erwartet – keinen Einfluss auf die Rechenzeit der Segmentierung hat. Die Laufzeit für die Zusammenführung nimmt hingegen mit steigender Kachelanzahl linear zu. Der Grund dafür liegt in der für die Zusammenführung relevanten Randbereichsfläche, welche mit zunehmender Kachelanzahl ebenfalls linear wächst (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Mit den wachsenden Randbereichen steigt auch die Anzahl der Punkte, die in der Zusammenführung erneut auf Homogenität überprüft werden.

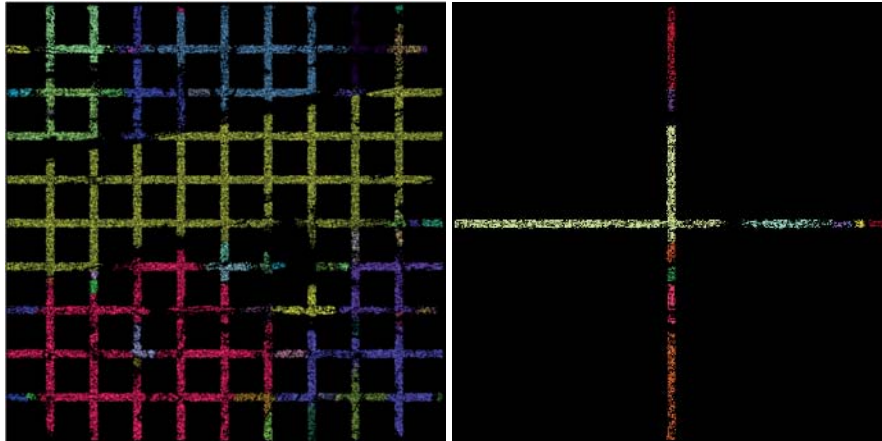


Abb. 6: Vergleich der Randbereiche aller Kacheln; Farbwerte \triangleq Segment-IDs; 1 m Suchradius; 10 m Kacheln (li) und 50 m Kacheln (re)

4 Fazit & Ausblick

Wie die Ergebnisse bestätigen, liefert das entwickelte Konzept nach Segmentierung und Zusammenführung unabhängig von der gewählten Kachelgröße die gleichen Segmente. Die Ergebnisse hängen somit nur vom gewählten Homogenitätskriterium und dem gewählten Suchradius ab. Durch eine geeignete Wahl der Kachelgröße kann die Laufzeit für die Berechnung optimiert werden. Zusätzlich bietet das vorgestellte Konzept großes Potenzial für Parallelisierung. Sowohl die Segmentierung als auch die Zusammenführung der Kacheln ermöglichen eine verteilte Berechnung auf mehreren Threads. Bei guter Skalierung kann dabei, bei einer Verteilung auf n Threads, die Laufzeit auf ein n -tel gesenkt werden.

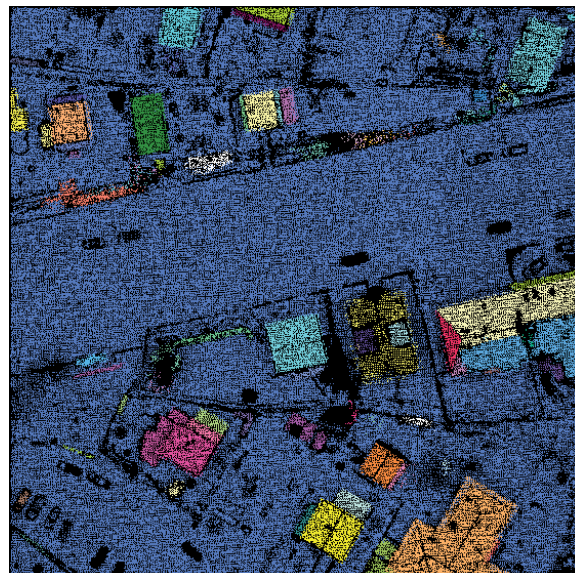


Abb. 7: Segmentierungsergebnis Testdatensatz Wien; Homogenitätskriterium $|p_i \cdot \text{NormalZ} - p_j \cdot \text{NormalZ}| < 0.01$; Suchradius $r = 1$ m; Unterschiedliche Farbwerte repräsentieren Segment-IDs

5 Literaturverzeichnis

- ADAMS, R. & BISCHOF, L., 1994: Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**(6), 641-647.
- BENTLEY, J. L., 1975: Multidimensional Binary Search Trees Used for Associative Searching. *Communications of the ACM* **18**(9), 509-517.
- CHEN, M.-H. & PAVLIDIS T., 1990: Image seaming for segmentation on parallel architecture. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(6), 588-594.
- PFEIFER, N., MANDLBURGER, G., OTEPKA, J. & KAREL, W., 2014: OPALS – A framework for Airborne Laser Scanning data analysis. *Computers, Environment and Urban Systems* **45**, 125-136
- VOSELMAN, G., GORTE, B.G., SITHOLE, G. & RABBANI, T., 2004: Recognising structure in laser scanner point clouds. *International archives of photogrammetry, remote sensing and spatial information sciences* **46**(8), 33-38

Konzept zur 3D-Erfassung, Analyse und Visualisierung eines dynamischen Bienennests

QUIRIN SCHREYER¹, EBERHARD GÜLCH² & JÜRGEN TAUTZ³

Zusammenfassung: Seit einigen Jahren arbeitet das Team HOBOS (Honeybee Online Studies) an der Erforschung der Bienenwelt. Im Mittelpunkt der Untersuchungen stehen Bienenvölker, die rund um die Uhr überwacht und beobachtet werden. In Zusammenarbeit mit der Audi Stiftung für Umwelt entsteht das neue und einmalige Forschungsprojekt Smart HOBOS, das Bienen beim freien Bau eines Bienennests mit modernsten Beobachtungsinstrumenten untersucht. Die 3D-Erfassung des dynamischen Bienennests stellt einen Teil des neuen Projekts dar. Aufgabe der Masterthesis war die Entwicklung eines geeigneten Prozesses zur Vermessung, Analyse und Visualisierung dieses Bienennests.

1 Einleitung und Motivation

Das Projekt Smart HOBOS wurde im Herbst 2014 in Zusammenarbeit der Universität Würzburg, dem Team HOBOS und der Audi Stiftung für Umwelt entwickelt. (HOBOS 2016) Auf dem Audi-Werksgelände in Münchsmünster nahe Ingolstadt entsteht die Smart HOBOS Station (Abb. 1).



Abb. 1: Smart HOBOS Station Münchsmünster (AUDI STIFTUNG FÜR UMWELT, 2016)

¹ Landesamt für Digitalisierung, Breitband und Vermessung, Alexandrastr. 4, 80538 München, E-Mail: Quirin.Schreyer@ldbv.bayern.de

² Hochschule für Technik Stuttgart, Schellingstr. 24, 70174 Stuttgart, E-Mail: Eberhard.Guelch@hft-stuttgart.de

³ Julius-Maximilians-Universität Würzburg, Josef-Martin-Weg 52, 97074 Würzburg, E-Mail: Tautz@hobos.de

Diese verwendet eine einzigartige, neue Methodik zur Beobachtung und Erforschung des Bienenvolks. In der Smart HOBOS Station sollen die Honigbienen in einer möglichst naturnahen Umgebung ihr Nest auf natürliche Art und Weise bauen und nicht durch menschliche Einflüsse gestört werden. Eine Infrarotkamera, eine Wärmebildkamera und ein optisches 3D-Erfassungssystem werden auf einem Roboterarm befestigt. Dieser ermöglicht, das Bienennest jederzeit aus verschiedensten Perspektiven zu beobachten. Die Entwicklung eines Konzepts zur Anwendung des hier beschriebenen 3D-Erfassungssystems bildet eine der Aufgaben der Masterthesis. Die 3D-Daten sollen darüber hinaus zur Visualisierung und für tiefer gehende Analysen genutzt werden. Mit Smart HOBOS betritt HOBOS Neuland, da es ein „derartiges Forschungsprojekt bisher noch nicht gegeben hat“ (HOBOS 2016). Demnach ist noch nicht bekannt, welche Ergebnisse zu erwarten sind. Es besteht jedoch eine hohe Wahrscheinlichkeit, dass die bisherigen Erkenntnisse über die Honigbiene in ihrem natürlichen Verhalten vertieft werden. Der Mensch ist folglich in der Lage, seinen Umgang mit Honigbienen optimal an deren Bedürfnisse anzupassen, um sie in Zukunft besser zu schützen. Das in der Masterthesis zu entwickelnde Konzept dient der Erfassung, Analyse und Visualisierung eines Teils der hierfür notwendigen Daten.

2 Realisierung der 3D-Erfassung und Visualisierung mit der Kinect

Der erste praktische Teil der Arbeit umfasste die Entwicklung eines automatisierten Arbeitsablaufs, der es ermöglichte, ein dynamisches Bienennest mit einer Microsoft Kinect V2 (Kinect) zu erfassen und das hieraus erstellte 3D-Modell online zu visualisieren. Hierbei wurden verschiedene 3D-Messprogramme in einer Simulationsumgebung an einem Bienennestmodell (Abb. 2) getestet und gegenübergestellt.



Abb. 2: Vorlage (HOBOS, 2016) und Nachbau eines frei gebauten Bienennests

2.1 Anwendung und Vergleich der Softwarepakete

Die Anwendung und der anschließende Vergleich der Softwarepakete zeigte eine Übersicht der Programme, die für die Smart HOBOS Station geeignet sind. Der Kinect Fusion Explorer der Firma Microsoft (MICROSOFT 2016) dient eher als Grundlage zur Programmierung, die die Möglichkeiten der Kinect zum 3D-Scan aufzeigt. Für eine geeignete Anwendung wären hier umfangreiche Programmierarbeiten notwendig. Arctec Studio der Firma Arctec 3D (ARCTEC

EUROPE 2016) hingegen könnte in einer Weiterentwicklung der 3D-Erfassung in der Smart HOBOS Station interessant werden. Die Software umfasst den gesamten Arbeitsablauf von der 3D-Erfassung über die Nachbearbeitung bis hin zur Visualisierung. Des Weiteren werden auch professionelle Laserscanner unterstützt, die eine höhere Auflösung als die Kinect bieten. Für die aktuelle Konfiguration der Smart HOBOS Station mit der Kinect als 3D-Erfassungssystem ist KScan3D von LMI Technologies (LMI TECHNOLOGIES 2016) am besten geeignet. Zum einen sind die Ergebnisse optisch ansprechender und zum anderen auch besser hinsichtlich der Genauigkeit im Vergleich zur Software Kinect Fusion Explorer. Des Weiteren ist KScan3D kostenlos verfügbar. Die Skriptfähigkeit rundet die Entscheidung für KScan3D ab, da es ohne weiteres möglich ist, den Arbeitsablauf von der 3D-Erfassung bis zum Export des 3D-Modells zu automatisieren.

2.2 Visualisierung auf der HOBOS-Webseite

Die Online-Visualisierung erfolgte auf einer zu diesem Zweck programmierten Webseite (siehe Abb. 3). Die entwickelte Visualisierung der 3D-Modelle bietet eine ansprechende Möglichkeit, die erfassten Daten darzustellen, zu untersuchen und folglich neue Erkenntnisse über den Wabenbau der Honigbiene zu erlangen.

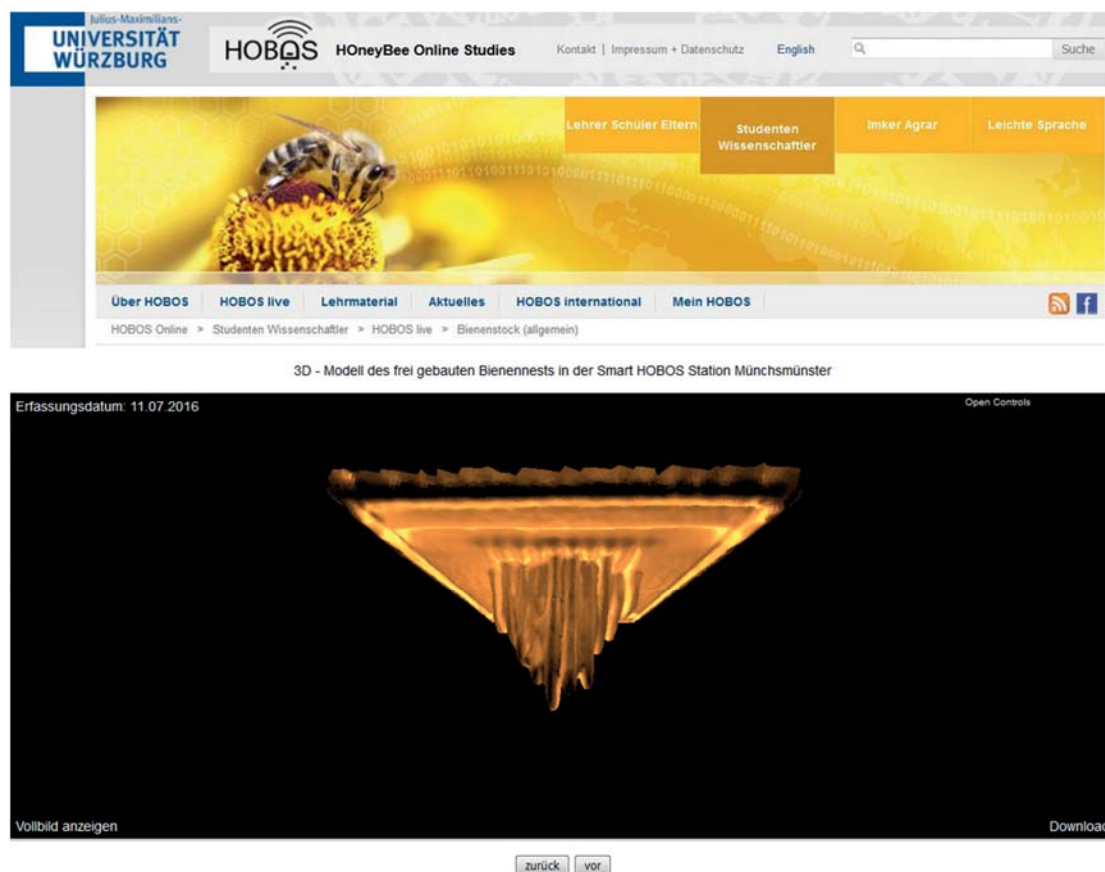


Abb. 3: Online-Visualisierung der 3D-Bienennestmodelle

Mit der integrierten Download-Funktion ist es möglich, die Daten lokal für tiefer gehende Analysen bereitzustellen. In einem ausführlichen Anwenderhandbuch wurden alle Schritte beschrieben, die zur Ausführung des entwickelten Prozesses von der 3D-Erfassung bis hin zur Online-Visualisierung notwendig sind.

3 Messungen mit dem Faro Freestyle 3D und Vergleich

Zur Beurteilung der Genauigkeit der mit der low-cost Kamera Kinect erfassten 3D-Modelle erfolgten Vergleichsmessungen (Abb. 4) mit dem professionellen Handlaser-Scanner Faro Freestyle 3D der Firma Faro (FARO 2016).

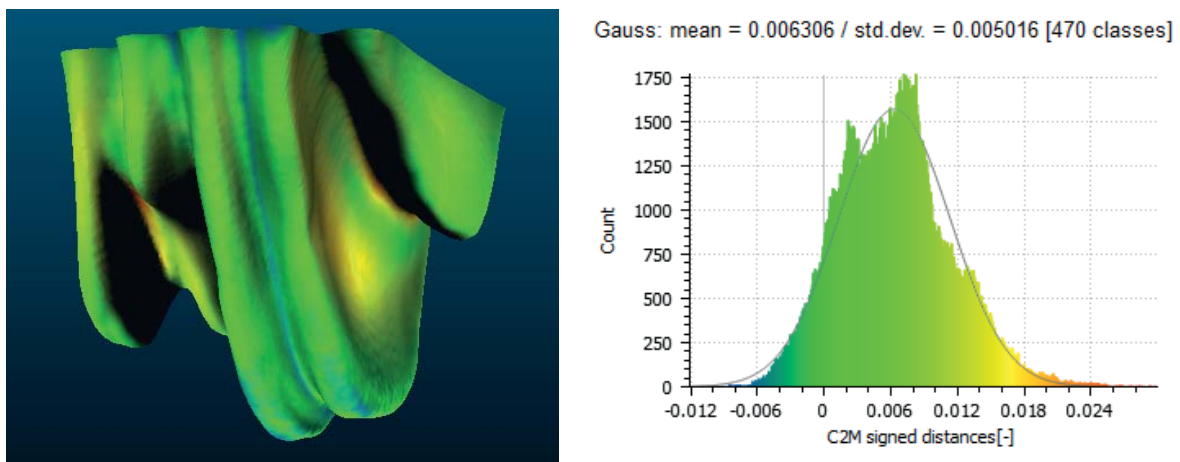


Abb. 4: Vergleich eines 3D-Modells der Kinect mit einem 3D-Modell des Faro Freestyle 3D

Die Ergebnisse zeigten durchschnittliche Abweichungen der 3D-Modelle von wenigen Millimetern, die bei einem solch hohen Preisunterschied von mehreren tausend Euro durchaus bemerkenswert sind. Es bleibt zu bedenken, dass die Kinect ursprünglich zur interaktiven Steuerung von Videospielen konzipiert wurde. Dennoch konnten für diesen speziellen Einsatzzweck gute Ergebnisse erzielt werden.

4 Weitere Analyse und Verarbeitung der digitalen Nestmodelle

Den Abschluss der Arbeit bildete die weitere Analyse und Verarbeitung der erfassten 3D-Daten. Zunächst wurde eine Deformationsanalyse durchgeführt, die sowohl quantitativ als auch visuell die Entwicklung des Bienennestmodells abbildete. In einem weiteren Schritt wurden die Bienen simuliert, die sich später in der Smart HOBOS Station auf der Nestoberfläche befinden. Mit Hilfe einer digitalen Retusche wird es ermöglicht, die Bienen von der Oberfläche der 3D-Modelle zu entfernen.

4.1 Deformationsanalyse

Die Deformationsanalyse (Abb. 5) der 3D-Daten ermöglichte diese über die Visualisierung hinaus noch besser zu verwerten. Mit Hilfe der dargestellten Analysemethoden ist es möglich,

die Entwicklung des frei gebauten Bienennestes innerhalb der Baustadien übersichtlich nachzuverfolgen.

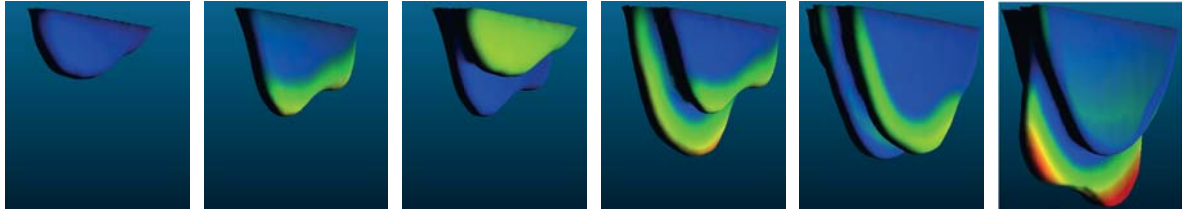


Abb. 5: Deformationsanalyse am Beispiel einer Modellentwicklung

Dennoch zeigten die Testdaten auf, dass Probleme auftreten können, wie zum Beispiel die zu schmale Erfassung einzelner Waben. Sofern diese Einschränkungen berücksichtigt werden, sind die abgebildeten Methoden zur Analyse durchaus wertvoll und bieten einen Mehrwert für die Verwertung der Daten.

4.2 Digitale Retusche der Bienen von der Nestoberfläche

Im Rahmen der Masterthesis wurden bisher die Bienen vernachlässigt, die sich in der Smart HOBOS Station auf dem Bienennest befinden werden. Die Bienen sind eine variable Komponente, deren Verhalten sich nur erahnen lässt. Zu Beginn des Nestbaus werden die Waben wohl vollständig von Bienen bedeckt sein. Demzufolge werden bei der 3D-Erfassung tatsächlich die Bienen und nicht die Bienennestoberfläche aufgenommen. Je größer die Waben werden, umso häufiger sind freie Flächen auf der Bienennestoberfläche zu erkennen. Aus diesem Grund wurde ein Prozess entwickelt, der es ermöglicht, die Bienen digital von der 3D-Modelloberfläche zu retuschieren (Abb. 6). Der entwickelte Prozess zeigte, dass eine vollständige Reproduktion der Nestoberfläche möglich ist.

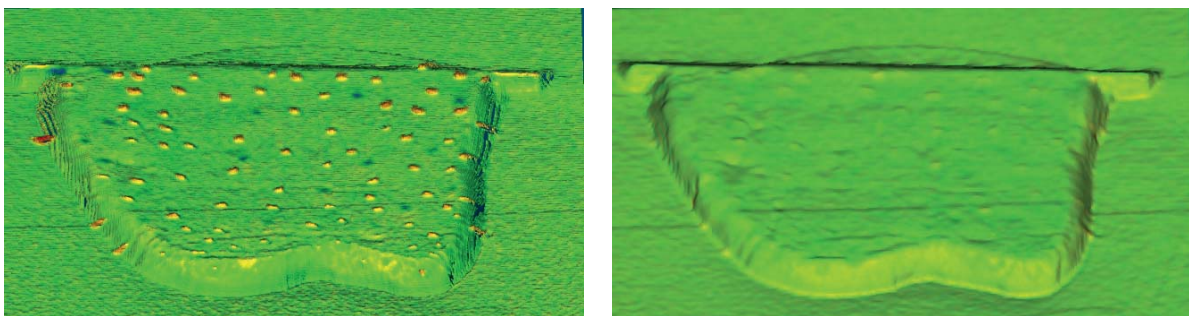


Abb. 6: Digitale Retusche der Bienen von der Oberfläche des 3D-Modells (links vorher/rechts nachher)

Aufgrund der Dauer der Aufnahme und der Auflösung der Kinect konnte der Prozess nur auf einzelne Bereiche der Nestoberfläche angewendet werden. Durch die Verwendung mehrerer 3D-Scanner, die gleichzeitig das Bienennest aus verschiedenen Positionen erfassen, wäre mit dem abgebildeten Prozess eine vollständige Retusche der Bienen von der gesamten Nestoberfläche theoretisch möglich. Die Retusche ist jedoch erst möglich, wenn die Bienen nach und nach alle Bereiche der Oberfläche mindestens einmal freigegeben haben. Die geringen Abweichungen

zwischen retuschierter und tatsächlich aufgenommener Nestoberfläche im Submillimeterbereich waren erstaunlich und bestätigten, dass der entwickelte Prozess gute Ergebnisse liefert.

5 Fazit & Ausblick

Die zuvor beschriebenen Teilaspekte finden sich im entwickelten Konzept wieder, das nun in der Smart HOBOS Station an einem realen, frei gebauten Bienennest angewendet werden kann. Auf dieser Arbeit aufbauend ist es denkbar, in Zukunft mit einem professionellen Laserscanner die Auflösung der 3D-Modelle zu erhöhen. Es entstehen dadurch neue Aufgaben, die gelöst werden müssen. Zum einen steigt die Dateigröße enorm an. Dies hat zur Folge, dass eine flüssige Online-Visualisierung höhere Hardwareansprüche erfordert. Zum anderen werden die Bienen auf der Nestoberfläche die Aufnahmen deutlich beeinflussen. Die entwickelten Analyse- und Visualisierungsmöglichkeiten sollten hingegen problemlos auch auf höher auflösende 3D-Modelle anwendbar sein. Eine weitere denkbare Verwertung der 3D-Modelle könnte ein 3D-Zeitrafferfilm sein. Mit einer Virtual-Reality-Brille könnte die Entwicklung des Bienennests in einem solchen Film noch eindrucksvoller nachvollzogen werden. Für Schülerinnen und Schüler könnten hieraus auch interaktive Lernspiele entwickelt werden. Der Entwicklungskreis der Kinect von der interaktiven Steuerung von Videospiele über die 3D-Erfassung des Bienennests zurück zum Einsatz als Steuerung in Lernspielen wäre schließlich geschlossen.

Zusammenfassend ist eine automatisierte und dennoch flexible Möglichkeit entstanden, die Kinect innerhalb der Smart HOBOS Station sinnvoll einzusetzen. Die Ergebnisse der Analyse- und Visualisierungsfunktionen werden hoffentlich umfangreiche und bisher unbekannte Erkenntnisse über den Superorganismus Bienen liefern und folglich dessen Schutz verbessern.

6 Literaturverzeichnis

- ARCTEC EUROPE, 2016: Artec 3D - Scanner - Komplette 3D - Scanlösungen. <https://www.artec3d.com/de/>, letzter Zugriff am 03.05.2016.
- AUDI STIFTUNG FÜR UMWELT, 2016: Audi Stiftung für Umwelt fördert Bienenforscher. <http://www.audi-umweltstiftung.de/auws/brand/de/projekte/hobos/impressionen.html>, letzter Zugriff am 18.07.2016.
- FARO, 2016: 3D-Messtechnik von FARO. www.faro.com/de-de/, letzter Zugriff am 03.06.2016.
- HOBOS, 2016: HOBOS - Honey Bee Online Studies. <http://hobos.eu/>, letzter Zugriff am 18.07.2016.
- LMI TECHNOLOGIES, 2016: 3D Mesh Generation with KScan3D Software and Kinect. <http://www.kscan3d.com/>, letzter Zugriff am 16.06.2016.
- MICROSOFT, 2016: Kinect Fusion. Online verfügbar unter <https://msdn.microsoft.com/en-us/library/dn188670.aspx>, letzter Zugriff am 18.07.2016.
- TAUTZ, J., 2015: Die Erforschung der Bienenwelt. Neue Daten - neues Wissen. 2. Aufl. Stuttgart: Klett MINT GmbH.

Integration und Bewertung photogrammetrischer Methoden und GIS zur Dokumentation von archäologischen Projekten am Beispiel von Gheriat el-Garbia (Libyen)

AGNES DINKEL¹, LUDWIG HOEGNER¹ & MANFRED STEPHANI¹

Zusammenfassung: In der vorliegenden Arbeit wird der Einsatz von photogrammetrischen Methoden und die Nutzung von Geoinformationssystemen für die Dokumentation des Denkmalbestands am Beispiel des römischen Ensembles von Gheriat el-Garbia (Libyen) diskutiert. Neben der Erzeugung verschiedener photogrammetrischer Produkte wird geprüft, mit welchen Auswirkungen für dieses Aufgabengebiet zu rechnen ist, wenn der Aufwand für die Aufnahme variiert wird. Dabei werden insbesondere die Variation der Anzahl der verwendeten Bilder und die Minimierung der Passpunkte unter Berücksichtigung der Passpunktconstellation und der resultierenden Genauigkeit beleuchtet. Exemplarisch wird gezeigt, dass ein GIS nicht nur eine integrative und effektive Möglichkeit bildet, heterogene Daten zu verwalten, sondern parallel zu den Arbeiten auch für die Planung und das Qualitätsmanagement genutzt werden kann. Die Erfassung von Denkmälern ist damit ein Themengebiet, das auch von der Geodäsie und der Geoinformation im Blick behalten werden sollte.

1 Was weg ist, ist weg

Obwohl die sog. Welterbekonvention in den 40 Jahren ihres Bestehens von mehr als 190 Staaten ratifiziert wurde (UNESCO 1963), befindet sich das kulturelle Erbe der Menschheit in einer prekären Situation: Jährlich werden zahlreiche Objekte nicht nur durch medienwirksame Zerstörungsaktionen radikaler Gruppierungen in Krisengebieten, sondern auch durch Ignoranz, Habgier und Armut auf der ganzen Welt zerstört. Eine Trendwende dieser Entwicklung ist derzeit nicht abzusehen, vielmehr nimmt die Zerstörung immer mehr an Geschwindigkeit zu.

Daher gewinnt die schnelle Erfassung und Dokumentation des aktuellen Zustands der historischen Monumente eine immer zentralere Bedeutung. Neben den Fachdisziplinen der Denkmalpflege, der Bauforschung und den verschiedenen archäologischen Disziplinen, sind hier besonders die Geodäsie und die Geoinformation gefordert, da diese die dazu benötigten Instrumentarien und Methoden bereitstellen können.

Die hier vorgestellte Arbeit diskutiert am Beispiel des Denkmalbestands des römischen Ensembles von Gheriat el-Garbia (Libyen) (Mackensen 2010, 2011)², in welcher Form photogrammetrische Methoden und GIS zur Dokumentation in der Denkmalpflege genutzt werden können, und welche

¹ Technische Universität München, Fachbereich für Photogrammetrie und Fernerkundung, Arcisstr. 21, D-80333 München, E-Mail: Agnes.dinkel@tum.de, ludwig.hoegner@tum.de, manfred.stephani@bv.tum.de

² Die verwendeten Daten wurden im Rahmen des vom LMUexcellent-Programm geförderten Projekts „Das severische Kastell MYD(--)/ Gheriat el-Garbia (Libyen)“ im Rahmen von archäologischen Feldkampagnen 2009 und 2010 unter der Leitung von Prof. Michael Mackensen vom Institut für Provinzialrömische Archäologie der LMU erhoben. Für die Zurverfügungstellung der Daten sei den Verantwortlichen an dieser Stelle herzlich gedankt.

Vorteile sich hierdurch ergeben. Detailliert wird insbesondere untersucht, wie sich die Variation der bei einer photogrammetrischen Auswertung verwendeten Parameter auf die Qualität und Genauigkeit der Ergebnisse auswirken.

2 Photogrammetrie

Als Testbeispiel für die photogrammetrischen Untersuchungen dienen die Reste eines römischen Wachturms, der in einer Entfernung von ca. 1 km nordwestlich vom römischen Militärkastell von Gheriat el-Garbia liegt (Abb. 1). Der runde Turm wurde während der archäologischen Kampagne außerplanmäßig besucht und dokumentiert (Mackensen 2010), da er schon teilweise zerstört und vom Einsturz bedroht ist.

Dafür wurde zum einen eine Bilderstrecke von 47 sich überlappenden Aufnahmen (> 60%, Nikon D300) um den Turm herum angefertigt. Zum anderen wurden 14 Passpunkte rundum an Ober- und Unterkante des Turms angebracht und tachymetrisch mit Katastergenauigkeit vermessen (Abb. 2). Auf dieser Basis kann eine verdichtete 3D-Punktwoke des Turms berechnet werden, die die Basis für alle weiteren Produkte bildet (Abb. 3), wobei hier die Software von Pix4D, Matlab und Cloud Compare genutzt werden: Damit lassen sich Oberflächen- und Volumenmodelle des Turms ableiten, die für Animationen oder zum 3D-Druck von realen Modellen verwendet werden können. Für die Wissenschaftler entscheidender sind jedoch die aus Schnitten und Profilen der Punktwoke ableitbaren Kenngrößen des Turms, wie Durchmesser, Mauerdicke, erhaltene Höhe des Bauwerks etc. und der Grundriss. Auch Abrollungen der äußern Mauerschale des Turms bilden für die Bauforscher und Archäologen eine wichtige Grundlage für Kartierungen von Bauphasen, Bearbeitungspuren oder Schäden am Bauwerk.



Abb. 1: Römischer Wachturm in Gheriat el-Garbia (Libyen): Ansicht von Südwesten

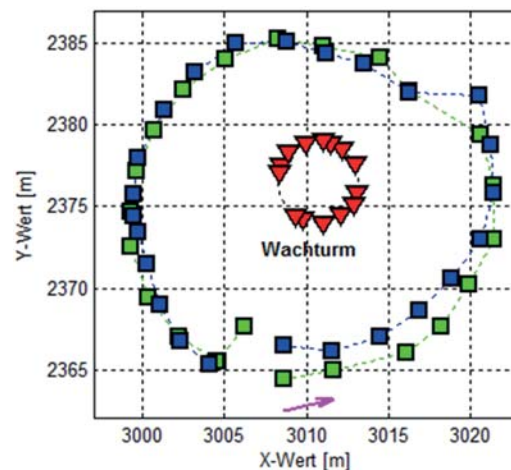


Abb. 2: Römischer Wachturm in Gheriat el-Garbia (Libyen): Konstellation der Passpunkte (Dreiecke) und Kamerapositionen (Quadrate) für die photogrammetrische Erfassung des Turms; Pfeil: Laufrichtung

Allerdings stellt sich die Frage, ob der hier vorgenommene v.a. zeitliche und technische Aufwand für die Dokumentation auch reduziert werden kann, und dennoch die avisierte Genauigkeitsvorgabe (Katastergenauigkeit) erreicht bzw. beibehalten werden kann.

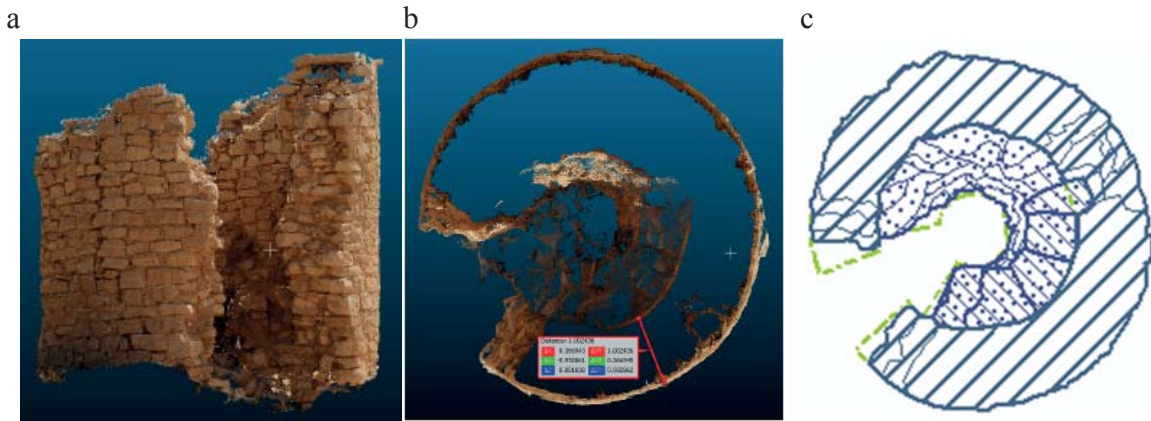


Abb. 3: Römischer Wachturm: a) 3D-Punktwolke, b) Abgreifen von Kenngrößen: Mauerstärke aus einem Querschnitt der Punktwolke c) Aus Punktwolke abgeleitetes Produkt: Grundriss des Bauwerks

Dazu werden verschiedene Parameter reduziert, deren Konstellation variiert und das mit diesen berechnete Ergebnis mit der beschriebenen Maximalkonfiguration verglichen. Einer der veränderten Parameter stellt die Anzahl der für die Berechnung der Punktwolke verwendeten Bilder dar. Mit der Reduzierung der Bilderzahl treten verschiedene Effekte auf, von denen besonders die Lageänderung der berechneten Kamerapositionen sehr plastisch veranschaulicht, mit welchen Effekten hier zu rechnen ist: Abbildung 4 zeigt die Residuen zwischen der Lage der Kamera bei der Maximalkonfiguration und der bei einer Bilderzahl von 25, 17 und 12. Diese drei Konfigurationen werden hier exemplarisch für eine Entwicklung gezeigt, die sich mit abnehmender Bilderzahl abzeichnet. Dabei kann mit 12 Bildern gerade noch eine geschlossene Lösung für das Bauwerk berechnet werden. Die Konfiguration mit 25 Bildern zeigt zwar kleinere Abweichungen von der Maximalkonfiguration, jedoch weisen sie ein einheitliches Muster auf. Im Gegensatz dazu geht diese einheitliche „Ordnung“ bei der Konfiguration mit 17 Bildern verloren, die Residuen besitzen unterschiedliche Beträge und beginnen zu divergieren. Bei der kleinstmöglichen Konfiguration mit nur 12 Bildern ist, obwohl noch immer eine Lösung für das gesamte Bauwerk erstellt werden kann, kein einheitlicher Trend mehr beobachtbar. Vielmehr zerfallen die berechneten Punkte in zwei Gruppen, die sich teils mit großen Beträgen gegeneinander verschieben. Diese Beobachtungen zeigen, dass sich schon lange vor dem Versagen einer Lösung deutliche Fehler ansammeln, obwohl eine Punktwolke und die Kamerapositionen berechnet werden können. Ähnliche Phänomene zeichnen sich auch für die Kamerakalibrierung, Höhenwerte und die Genauigkeit der Punkte ab. Daher sollte bei einer photogrammetrischen Aufnahme eines Bauwerks nicht auf eine vergleichsweise ausführliche Dokumentation mittels Bildern verzichtet werden, was angesichts des geringen Zeitaufwands, den weitere Aufnahmen bedeuten, kein großes Problem und nur geringen Zeitverlust darstellen sollte.

Eine weitere Möglichkeit, Zeit zu sparen, bildet die Reduzierung der Passpunkte auf ein Minimum von drei. Allerdings stellt sich die Frage, welche Konstellation der Passpunkte für ein solches

Objekt am besten geeignet ist. Zu diesem Zweck wurden die Auswirkungen von Konstellationen mit drei Passpunkten berechnet.

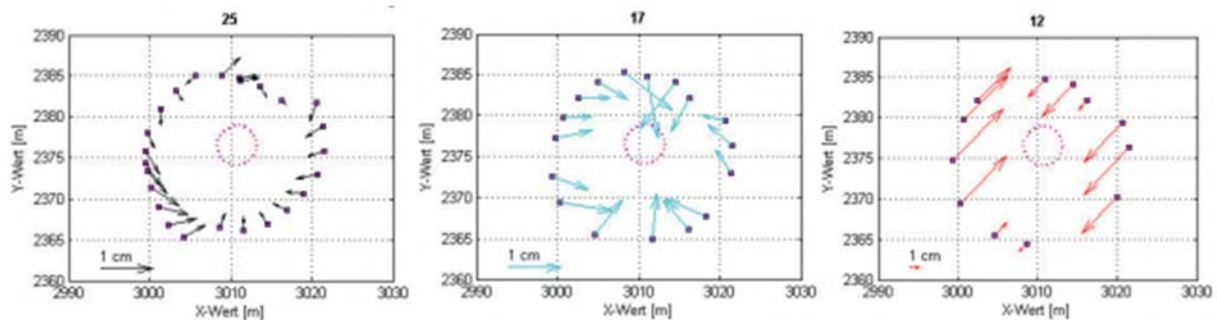


Abb. 4: Römischer Wachturm: Einfluss der Bilderzahl (25, 17, 12) auf die Lage der Kamerapositionen im Vergleich zur Maximalkonfiguration

Diese Konstellationen lassen sich wie folgt beschreiben: alle Passpunkte befinden sich auf einer Seite des Objekts, an seiner Unterkante oder an seiner Oberkante. Wiederum verdeutlichen die Änderung der Lagekoordinaten der Kamerapositionen im Vergleich zur Maximalkonfiguration anschaulich die Effekte (Abb. 5): Auch hier lassen sich unterschiedliche Muster bei den Residuen beobachten. Bei einer einseitigen Anbringung der Passpunkte besitzen die Residuen den kleinsten Betrag und zeigen eine Rotation ungefähr um die Mitte des Turms in Richtung der Bewegung bei Aufnahme der Bilder. Die Skalierung wird dabei annähernd beibehalten. Im Gegensatz dazu weisen die anderen Konstellation zwar keine Rotation auf, dafür aber deutlich größere Beträge und eine deutliche Skalierung der Kamerapositionen: Bei der Konstellation mit Passpunkten an der Oberkante wird der Maßstab deutlich verringert, wogegen er sich bei der Konstellation an der Unterkante deutlich vergrößert. Vergleicht man also die verschiedenen Konstellationen, so schneidet die einseitige Anbringung der Passpunkte vor allem aufgrund des geringen Verschiebungsbetrags am besten ab. Ein ähnliches Bild zeigt sich auch beim Vergleich der Höhenkomponente der Kamerapositionen, den Parametern der parallel durchgeführten Kamerakalibrierungen und dem Vergleich der verdichteten Punktwolken. Letzterer besitzt für die einseitige Passpunktanbringung einen Mittelwert von 4,7 mm bei einer Standardabweichung von 2,8 mm für die Differenz zur Maximalkonfiguration. Auch die Berechnung eines Worst-Case-Szenarios auf Basis der vorliegenden Daten und eine Abschätzung der Fehlerfortpflanzung zeigen, dass maximal mit einer Abweichung von ca. 8 mm für die einseitige Konstellation von der Maximalkonfiguration zu rechnen ist. Der Wert liegt jedoch weit unter der Genauigkeit der Passpunkte selbst, die lediglich Katastergenauigkeit besitzen. Dies zeigt, dass sich zwar Effekte aufgrund der Passpunktreduzierung deutlich nachweisen lassen, diese aber keine relevante Größe besitzen. Für die Messung der Passpunkte vor Ort kann der Zeitaufwand reduziert werden: Es kann darauf verzichtet werden, den Tachymeter für die Einmessung der Passpunkte auf allen Seiten des Objekts aufzustellen. Bei einer einseitigen Anbringung der Passpunkte kann dieser Vorgang so auf ein Minimum reduziert werden. Allerdings sollte man, um mögliche Messfehler aufzudecken zu können, die Anzahl der Passpunkte mindestens auf vier erhöhen und so die nötige Redundanz

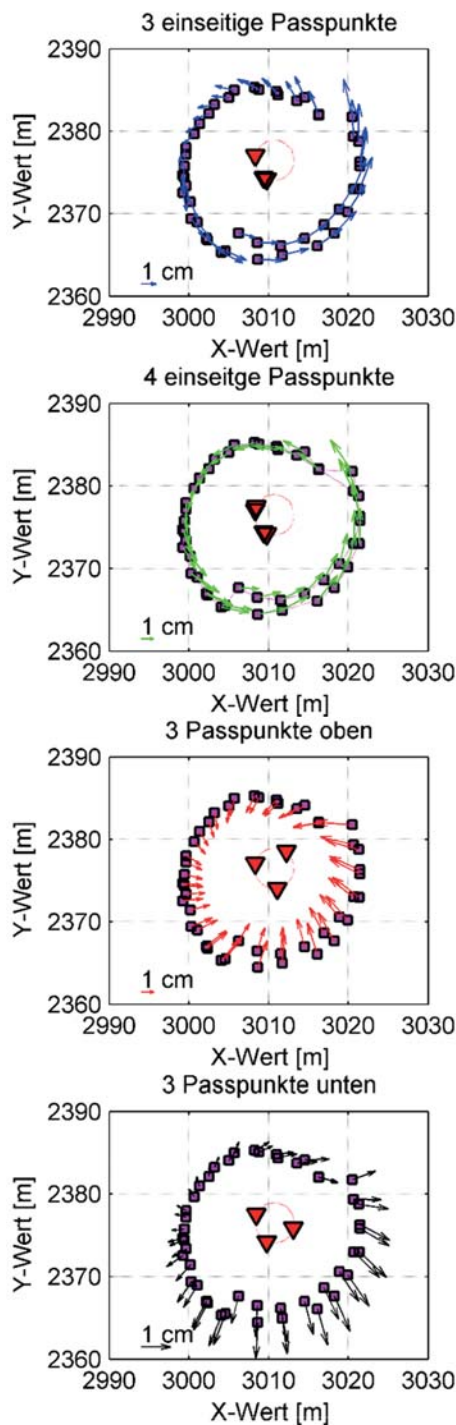


Abb. 5: Einfluss Passpunktanstellung auf die Lage-Koordinaten der Kamerapositionen. Residuen im Vergleich zur Maximalanstellung: einseitige Anbringung von 3 bzw. 4 Passpunkten, Anbringung von 3 Passpunkten an Ober- bzw. Unterkante des Wachturms.

schaffen. Zudem gelten diese Beobachtungen nur für einen geschlossenen Bildverband. Eine Untersuchung für offene, nicht in sich geschlossene Bildverbände muss getrennt hiervon betrachtet werden.

Bildmessung stellt somit eine alternative Methode zur Erfassung von Denkmälern dar, zumal sich historische Objekte aufgrund ihrer variationsreichen Oberfläche gut für photogrammetrische Methoden eignen. Durch die sehr starke Verknüpfung der einzelnen Bilder eines geschlossenen Verbandes entsteht ein stabiles 3D-Modell, welches nur noch durch Passpunkte skaliert, gedreht und gelagert werden muss.

3 GIS

Eine weitere, zeitgemäße Möglichkeit, die Erfassung von Kulturdenkmälern zu unterstützen, bildet die Verwaltung und Sicherung der aufgenommenen Daten in Form eines

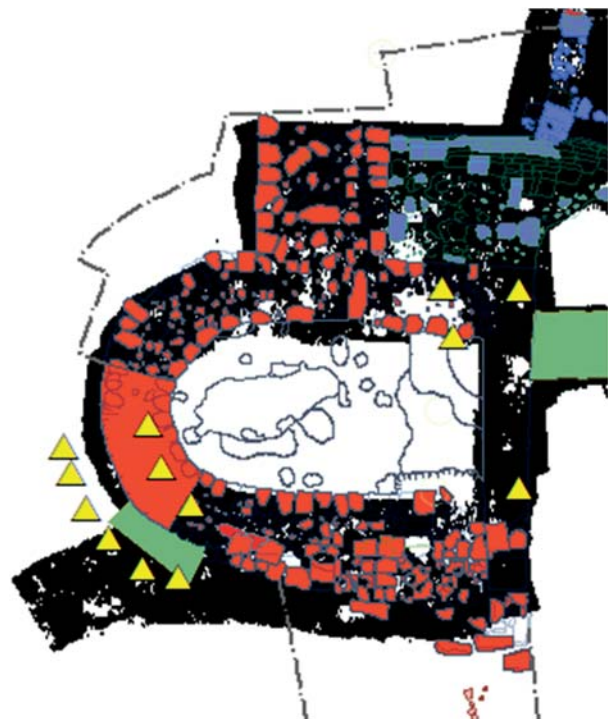


Abb. 6: Gheriat el-Garbia: Porta Decumana des Militärkastells: Vergleich 3D-Scan (schwarz), tachymetrische Messung (gelbe Dreiecke) und analoge Messung (Planumszeichnung mit Passpunkten; sonstige Farben) in ArcMap (Esri)

Geoinformationssysteme (GIS). So können nicht nur heterogene Daten verschiedenster Fachdisziplinen wie Bauforschung, Archäologie, Geophysik und verschiedene biologische Disziplinen etc. übersichtlich und effektiv miteinander verwaltet und organisiert werden, sondern auch verschiedene Dokumentationsmedien wie Kartenmaterial, Satellitenbilder, lokale Messungen mit Tachymeter oder 3D-Scanner, photogrammetrische Messungen oder analoge Messungen, abgeleitete Pläne und Beschreibungen. Ein parallel zu den Arbeiten geführtes GIS ermöglicht darüber hinaus eine tagesaktuelle Übersicht über den Fortschritt der Arbeiten und eine Planung und Koordination der beteiligten Teams. So verwendet, können GIS-Systeme einen Beitrag zum Qualitätsmanagement leisten.

Durch den unmittelbaren Vergleich können beispielsweise Inkonsistenzen und mögliche Fehler zwischen Daten, die mit unterschiedlichen Messmethoden aufgenommen worden sind, aufgedeckt und korrigiert werden. Abbildung 6 zeigt den Vergleich zwischen 3D-Scandaten, tachymetrischen Messungen und einer analogen Planungszeichnung eines Tors des römischen Militärlagers von Gheriat el-Garbia. In diesem Fall rühren die Widersprüche daher, dass die Daten in zwei unterschiedlichen Jahren des Projekts aufgenommen wurden, wobei erst im zweiten Jahr ein ausgeglichenes Festpunktfeld zur Verfügung stand. Eine Korrektur der Daten der ersten Kampagne ist daher notwendig. An diesem Beispiel kann man nachvollziehen, dass ein GIS nicht nur eine Datenbank ist, in der Ergebnisse zusammengeführt werden, sondern vielmehr auch als aktives Hilfsmittel für Qualitätsmanagement und Planung während der Arbeiten genutzt werden kann.

4 Fazit

In der Arbeit wird gezeigt, dass die Erhaltung des kulturellen Erbes der Menschheit nicht nur ein Thema für die entsprechenden Fachdisziplinen ist. Vielmehr können auch Ingenieurdisziplinen wie die Geodäsie und die Geoinformation einen substanziellen Beitrag zur Erfassung des Denkmalbestands liefern, indem Methoden wie die Bildmessung von Objekten und integrative Systeme wie GIS auch in diesem Themengebiet zum Einsatz kommen. Es gilt, auch in Zukunft ein Auge darauf zu haben, inwieweit neue Methoden und Techniken auch in diesem Aufgabengebiet Anwendung finden können. Durch das Zusammenwirken von etablierten und neuen technischen Verfahren kann ein Mehrwert sowohl vor Ort als auch bei der Auswertung der erhobenen Daten erzielt werden.

5 Literaturverzeichnis

MACKENSEN, M., 2010: Das severische Vexillationskastell Myd(---)/Gheriat el-Garbia am limes Tripolitanus (Libyen): Bericht über die Kampagne 2009. *Mitteilungen des Deutschen Archäologischen Instituts, Römische Abteilung* **116**, 363-458.

MACKENSEN, M., 2011: Das severische Vexillationskastell Myd(---) und die spätantike Besiedlung in Gheriat el-Garbia (Libyen): Bericht über die Kampagne im Frühjahr 2010. *Mitteilungen des Deutschen Archäologischen Instituts, Römische Abteilung* **117**, 247-375.

UNESCO, 1963: Übereinkommen zum Schutz des Kultur- und Naturerbes der Welt. <http://unesco.de/infothek/dokumente/uebereinkommen/welterbe-konvention.html>; Stand 30.10.2016.

Automatically Extraction and Reconstruction of Cupola Geometries of Orthodox Churches from Precision Point Clouds

MARIA CHIZHOVA¹, ANDREY GURIANOV², DMITRII KOROVIN³,
ANSGAR BRUNN⁴ & UWE STILLA⁵

Abstract: Complex geometry extraction from point clouds is an actual problem in reverse engineering. Simple geometrical models (like parallelepipeds, prisms, pyramids, cones, spheres) were already applied in construction and machine-building modeling, but are not sufficient for high quality BIM now. This work, which is carried out in the context of virtual reconstruction of destroyed orthodox churches, presents a robust and efficient method of cupola (domes) and tambour geometry extraction from precise point clouds. The rich diversity of architectural forms, which are defined by many parameters, does not allow to consider this problem as a trivial duty, because usual geometry extraction methods fail for these object types. The new developed algorithm is presented and realized.

1 Introduction

1.1 Motivation

Ancient cultural and science developments, accumulations of practical human experience, led to fantastic achievements in industrial and civil constructions. These achievements are reflected in various architectural forms, construction materials, and technologies.

Simple geometrical models (like parallelepipeds, prisms, pyramids, cones, spheres) were already applied in construction and machine-building modeling. However, such geometrical forms are insufficient for modern industry requirements, which arises the need for more difficult mathematical models, and, therefore, the identification relations between the parameters of the geometrical models and the projected constructions.

This work is carried out in the context of actual research in the virtual reconstruction of destroyed orthodox churches, in which the cupola (dome) geometry extraction plays an important role. The rich diversity of architectural forms does not allow us to consider this problem as a trivial duty.

Analytical surfaces have a broad application in the various branches of technique and construction, being relevant for the geometry description of complex church components. A significantly high

¹ Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Röntgenring 8, D-97070 Würzburg/
Technische Universität München, Arcisstr. 21, D-80333 München, E-Mail: mariatschishowa@yahoo.de

² Staatliche Universität Iwanowo, Fakultät für Höhere Mathematik und Informationstechnologien,
Ermak Str. 39, 153025 Iwanowo, Russland, E-Mail: a.v.gur.2008@mail.ru

³ Staatliche Universität der Energietechnologien in Iwanowo, Rabfakowskaja Str. 34, 153003 Iwanowo,
Russland, E-Mail: dmitriyikorovin@list.ru

⁴ Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Röntgenring 8, D-97070 Würzburg,
E-Mail: ansgar.brunn@fhws.de

⁵ Technische Universität München, Arcisstr. 21, D-80333 München, E-Mail: stilla@tum.de

number of parameters, which complicate a geometry detection task from a point cloud, defines such kind of surfaces.

The aim of this work is the development of a robust method of automatically dome geometry extraction from precision point clouds.

1.2 Related works

Complex geometry extraction from point clouds is a current problem in reverse engineering, that is considered in many articles. Due to the huge number of parameters in the context of big data problems, which is relevant for point clouds, the usual geometry extraction methods (e.g. extraction of geometrical primitives with Hough Transform or RANSAC) fail in this case. Moreover, our research object cannot be approximated with geometrical primitives in most cases (unlike to e.g. ALBY & GRUSSENMEYER 2012).

The mathematical construction of complex surfaces is one of the main objectives in engineering geometry. Different mathematical models and aspects of curve and complex surface approximation from point clouds are represented in BUREICK et al. (2016).

A spline interpolation is very common in point cloud processing (e.g. thin plate spline in non-rigid texture mapping procedure (FAN et al. 2012)) and curve fitting (WANG et al. 2004). BARAZZETTI et al. (2016) presents a modelling of bridges by NURBS surfaces. B-spline curves are extracted from an input point cloud, using its point subsets as control points for a curve approximation. B-spline basis functions are defined on the knot vector, which is a non-decreasing sequence of real numbers with external elements, the knots. BELYAEVA (2014) deals in her PhD-Thesis with a development of mathematical surface models (vector/matrix models) and surface transformation algorithms for practical tasks with an application of computer geometry. In particular, the construction of domes in orthogonal and cylinder coordinate systems and its approximation with cubic spline is considered.

However, an important modeling objective in the context of geometry detection from big point clouds is the determination of the minimal number of parameters, which describe necessary variations of dome geometries. Therefore, the extraction of significant information (e.g. features points) is a crucial task.

RUSU et al. (2008) developed a point descriptor, identifying points on planar, round, linear surfaces. In BUENO et al. (2016) the key points are detected using entropy values by planarity and curvature change. In GEVAERT et al. 2016 a maximal height difference, its standard deviation and number of points per bin are estimated, solving for key point extraction of an input point cloud. The bins have been defined from the geographical grid determined by a orthomosaic. Some algorithm, which are suitable for feature point detection in 2D images (Harris detector (HARRIS & STEPHENS 1988), SIFT (LOWE 1999; LOWE 2004)) are applicable for 3D point clouds, too.

There are some program applications allowing a semi-automatic modeling of architectural forms, in which the model parameters of the structure elements are estimated from user defined keypoints (KIVILCIM & DURAN 2016). Such approaches basis on an a priori created library of architectural elements (DORE et al. 2015). DORE et al. (2015) estimated and reconstructed destroyed architectural elements (domes and tambours) from horizontal cuts, according to a model from elements library.

CANCIANI et al. (2015) used point cloud cuts to extract key points and estimate a mathematical model of architectural elements for its further “path-wise” reconstruction.

According to our objectives, we want to provide a complex form reconstruction algorithm using a minimal parameter set, that will be suitable for incomplete data, too.

2 Method

The proposed method can be divided in two steps – dome block segmentation and dome/tambour geometry extraction. The second step is in the focus of this article and will be described according to the research object.

2.1 Pre-segmentation of dome blocks

The algorithm, considered in CHIZHOVA ET AL. 2016, bases on the idea, that there is a certain number of domes of orthodox churches (e.g. 1, 3, 5, 7, 9, 33). The quantity of domes is determined by orthodox construction canons. Our duty is to segment the point set $P_{x,y,z}$ into subsets P_i , which may belong to every type but also to a single dome. The number of such subsets corresponds to the possible number of the domes.

We consider horizontal cuts of the treated church point cloud, which are covered with an orthogonal grid. In each layer, at different heights, containing the point cloud cut of the dome and the grid, distances between grid nodes and cut points are estimated. If some points have almost an equal distance to the grid nodes, these points are assumed as a dome/tambour with a circular horizontal projection. A user defined deviation of the distances gives an “index of detection quality”, which maximal value allows to state, that the bottom of the dome block has been detected.

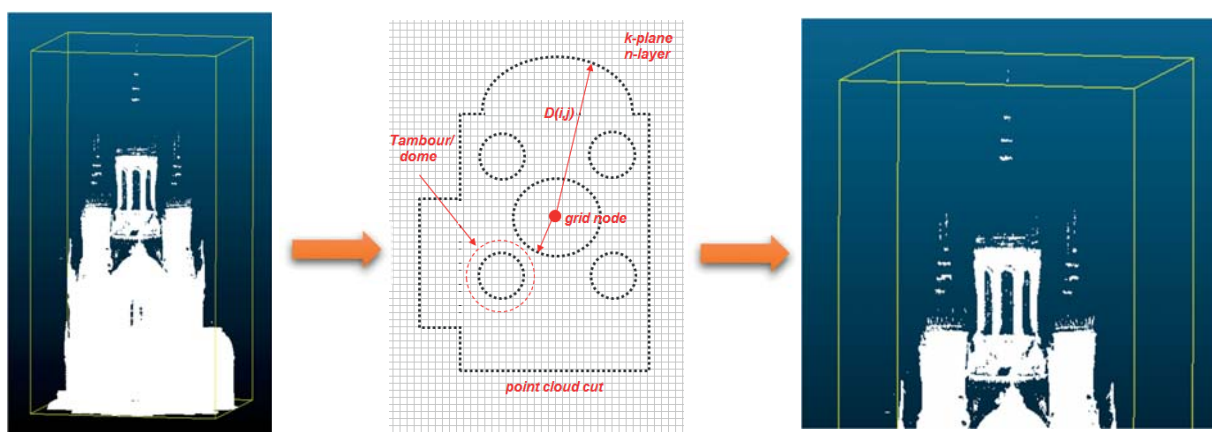


Fig. 1: Segmentation process of dome blocks (dome with tambour)

2.2 Mathematical modelling and point cloud simulation of domes

In this research, the mathematical model of the dome should consider the specific properties of point cloud analysis as well as the classification and storage procedure of recognized objects. Therefore, it is necessary to reduce a space dimension, in which an analysis is carried out, and to minimize a set of model parameters, keeping its generality. As a basic form of a dome, we will choose the most common onion form, which can be considered more universal in relation to other domes forms (e.g. oval domes). Besides, it is necessary to consider such types of domes (and its tambours) like ripped or “umbrella” domes, which are divided at the base into curved segments, which follow the curve of the elevation.

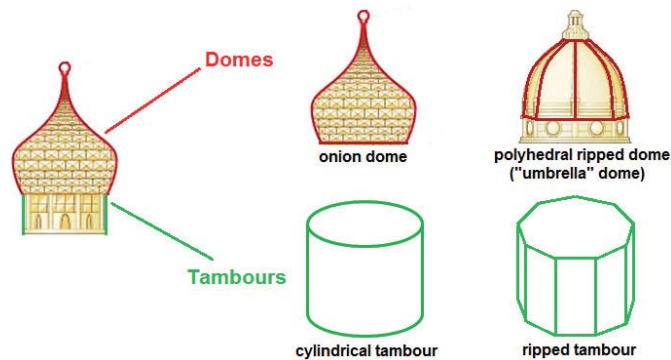


Fig. 2: Different types of domes and tambours

Obviously, the onion domes have an axial symmetry, being a surface of rotation, which are defined completely by a planar profile curve (meridian).

Therefore, the modeling of a dome form can be considered as a selection of a planar parametric curve, setting the meridian. Let's direct the abscissa axis along the dome axis from its basis to its top, then a radius of the dome cuts on different heights will be stored on the ordinate axis. Moving along the dome axis, the radius of its cut increases at first, reaching a maximum, then decreases, concerning an axis on its top. Obviously, there is a crease point between a maximum and top.

This kind of geometrical forms is described usually with parametric third order Bezier curves or cubic splines. However, in our case, this representation form cannot be used, because a parameter set of Bezier curve includes not only the points on this curve but external point too, setting its tangent and providing only directional information (a curve does not pass through these points).

Obviously, the reconstruction procedure of a tangent to the dome surface in the point cloud will have an error, which is bigger as an error by the definition of point coordinates on the dome surface. But a specific feature of the Bezier curves is, that the allocation of parametric tangents has more influence on a curve form as the coordinates of a points on this curve.

Moreover, a standard set of tangent segments with coordinates of its end points cannot be used in our case, because it will complicate a procedure of object classification and storage. A tangent setting with normalized vectors can solve as makeshift, but it will not improve a situation with error determination.

An application of a cubic spline is problematic in this case, too. It is known, that cubic spline parameters significantly depend on a piecewise-defined interval or its knots, building this spline. An effective classification will be interfered through a problem to offer a spline composition, uniform for all reconstructed dome forms.

Obviously, all listed above problems are insoluble in case of a geometry reconstruction from incomplete data, if a dome is partially destroyed or scanned with a big error.

Assuming enough simple prerequisites, it is possible to approximate a dome profile curve with a third degree polynomial, which coefficients are determined by coordinates of four reference points. In this case, the coefficient values or polynomial roots will be invariant concerning the choice of these reference points that allows using of polynomial coefficients or its roots as form parameters by object classification and storage.

Following points are chosen as reference (features) points:

- Dome basis: a point on the junction from the tambour to the dome ($P_0(x_0, y_0)$);
- Equator: a point on the maximal radius of the profile curve ($P_1(x_1, y_1)$);
- The crease point by the function change of the profile curve ($P_2(x_2, y_2)$);
- The point on the dome top ($P_3(x_3, y_3 = 0)$)

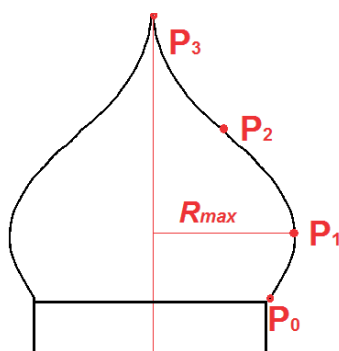


Fig. 3: Reference points on dome profile to be detected

Obviously, a point P_3 corresponds to one of the polynomial roots; therefore, a dome top can be chosen as one of the polynomial parameters and coefficients a , b , c of the second order polynomial as residual parameters. In this case, a profile curve model can be defined as

$$f(x) = (ax^2 + bx + c)(x - x_3)$$

This model was investigated during extensive computing experiments, in which course it has become clear that the cubic polynomial is not suitable due to wrong approximation of P_3 point vicinity. In particular, a condition of the profile contact with dome axis is violated, but it cannot be permitted by the construction rules of domes. Thus, a polynomial order has been raised to the fourth without increasing the parameter number for faults elimination. It is possible, if the difference $(x - x_3)$ will be squared.

Thus, the model can be defined by multiplication of second order polynomial to squared difference:

$$f(x)=(ax^2+bx+c)(x-x_3)^2$$

Due to x_3 -root multiplicity, the total number of parameters will remain equal to four; it is necessary to determine three residual unknown model parameters a , b , c (x_3 is already known). All calculation formulas have been received using a computer algebra system and exported into the C++ programming language. It allows to construct a generator for the simulation of domes point clouds. The program realization of the test data generator and geometry analyzer is made in a cross-platform integrated environment of software development (IDE) Qt5 using the PCL library version 1.8.

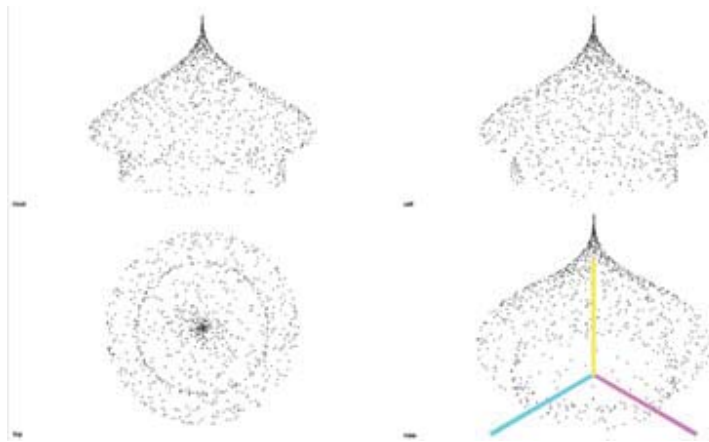


Fig. 4: Generated point clouds of domes

2.3 Extraction and reconstruction of the dome geometry

A detection procedure of dome the geometry consists of the following steps:

- Point cloud regularization;
- Definition and refining of dome axis coordinates;
- Profile curve extraction;
- Definition and refining of model parameters.

2.3.1 Point cloud regularization

Point cloud regularization allows reducing a point cloud capacity to values according to given spatial resolution and – therefore – modeling accuracy that relieves excessive calculations and increases the program efficiency. Moreover, it allows to eliminate the potential irregularity of the input point cloud. This procedure splits the point cloud space with cubic voxel grids. Further, all points in each voxel have been replaced with their centroids, which coordinates are defined by the average of such point coordinates.

However, in case of incomplete input data, it can be insufficiently for the construction of the profile curve. In particular, this procedure does not guarantee that the centroid of the considered layer will coincide surely with the center of the circle, being a dome cut in this layer.

2.3.2 Dome axis extraction

On the next step, a layered bottom-up “scanning” of the grid will be carried out for the dome axis definition and creation of its array. All horizontal grid layer, which include at least three points, will be analyzed with the following algorithm:

- 1) At first, a projection of all layer points to the plane $z=0$ will be calculated;
- 2) Then the initial proximity of the center and the circumcircle of these points will be defined. Three points will be chosen for this purpose according to the following rule: at first we choose two points in the point set, lying on the maximal distance from each other, then we find a third point, which is maximal distant from each of the first two points. Based on this data, the center coordinates and radius of the circle, set by the chosen points, are calculated. Such procedure allows to increase the definition accuracy of the initial parameter proximity in case of incomplete input data.
- 3) Further, circle parameters are specified by the conjugate gradient method for the minimization of the sum of squared distances from the circle to all points in this layer. The calculated coordinates of the circle center are accepted as dome axis coordinates in this layer and saved for a subsequent averaging on all analyzed layers.
- 4) Then we search for maximum and minimum distances from this center to the layer points. It is necessary for detecting of possible polygonal forms in the dome cut. A maximal distance is accepted as radius of regular polygon’s circumcircle and saved in the array element of profile curve according to an analyzed layer. A minimal distance is considered as the incircle radius of the same polygon using for the assessment of polygon sides number (n) with the following formula:

$$n=180^{\circ}/\arccos(R_{\min}/R_{\max})$$

Besides, an angle, which set a direction on the first found maximum, will be estimated in polar coordinate system.

Further, these parameters are specified by conjugate gradient method for deviation’s minimization of all layer points from constructed polygon. The calculated values will be stored for the subsequent averaging on the analyzed layers, which is carried out separately for dome and tambour.

- 5) After the “scanning” of the layers, which are important for the analysis, we build up an array containing information about layer height, maximal cut radius in the layer and polygon parameters. Average coordinate values of cut centers are accepted as dome axis allocation.

Then, an array analysis will be carried out for the parameter assessment of profile curve.

2.3.3 Profile extraction and model definition

A profile curve and its first derivative, defined by numerical differentiation, are needed for steady searching of reference (feature) points. A second derivative of profile curve could be useful for point P_2 search, but preliminary experiment shows, that an error of its values, defined by a numerical method, is very big due to noise with accidental outliers.

Therefore, only the function and its first derivative are analyzed. Array recognition is carried out according to increasing layer height.

The following feature points are detected from function and derivative analysis:

- 1) the point P_0 is detected by the maximal value of first derivative;
- 2) the point P_1 is detected by the maximal value of the function;
- 3) the point P_2 is detected by the minimal value of first derivative and
- 4) the point P_3 is detected by the maximal height value of dome point cloud.

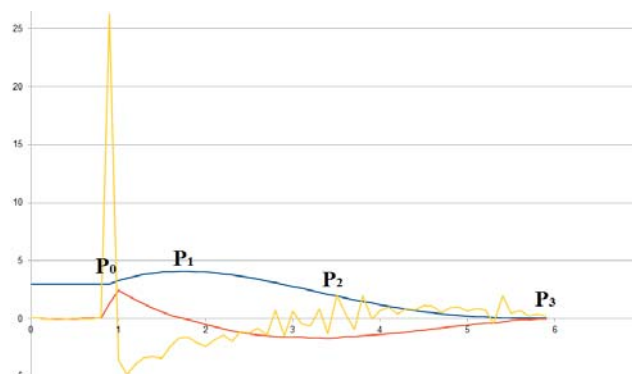


Fig. 5: Derivative analysis of extracted profile curve

The model parameters a , b , c , x_3 are calculated from the coordinates of found points and specified by a least squares method. Thus, a profile curve will be constructed based on this data.

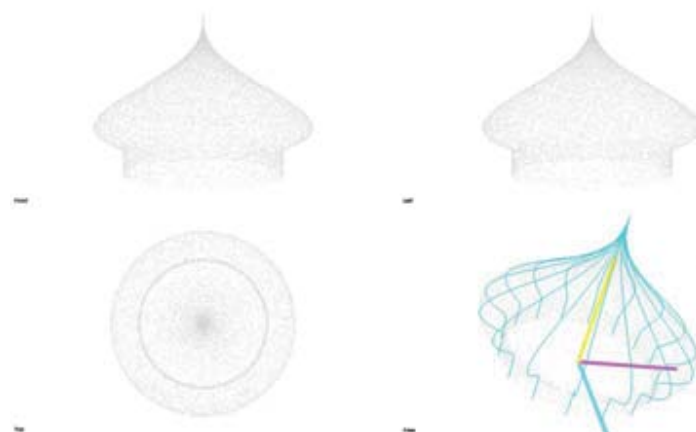


Fig. 6: Reconstructed profile curve from point cloud

In the final step, the averaging of polygon parameters is carried out for the dome and the tambour. A cut is polygonal, if the number of sides $2 < n < 14$. In all other cases it will be considered as a circle.

For the simplification of the classification problem, it could be better to store profile curve parameters instead of the coordinates of reference (feature) points. For this purpose, it is necessary to normalize coordinates of these points on the dome height, which is defined by the optimization, to repeatedly calculate and store profile parameters based on this normalization.

3 Conclusion

This article presents a robust and efficient method of the automatic extraction and reconstruction of complex geometrical cupola forms from point clouds. An example has been shown.

Practically, a calculation of the second derivation is not necessary because of oscillations of discrete point cloud. The algorithm shows a good geometry reconstruction in practical application.

4 References

- ALBY, E., & GRUSSENMEYER, P., 2012: From point cloud to 3d model, modelling methods based on architectural knowledge applied to fortress of chatel-sur-moselle (france). – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **39**(B5), 75-80.
- BARAZETTI, L., BANFI, F., BRUMANA, R., PREVITALI, M. & RONCORONI, F., 2016: Bim from laser scans... not just for buildings: Nurbs-based parametric modeling of a medieval bridge. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **III**(5), 51-56.
- BELYAEVA, Z., 2014: Geometric modelirovanie prostranstvennyh konstrukziy. PhD thesis. Uralische Föderale Universität benannt nach dem ersten Präsidenten Russlands B.N. Jelzin, Jekaterinburg, Russland.
- BUENO, M., MARTINEZ-SANCHEZ, J., GONZALEZ-JORGE, H. & LORENZO, H., 2016: Detection of geometric keypoints and its application to point cloud coarse registration. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **41**(B3), 187-194.
- BUREICK, J., NEUNER, H., HARMENING, C. & NEUMANN, I., 2016: Curve and surface approximation of 3d point clouds. Allgemeine vermessungs-nachrichten (avn) **123**(11-12), 315-327.
- CANCIANI, M., FALCOLINI, C., SACCONI, M. & SPADAFORA, G., 2013: From point clouds to architectural models: algorithms for shape reconstruction. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **40**(5/W1), 27-34.
- CHIZHOVA, M., KOROVIN, D., BRUNN, A. & STILLA, U., 2016: Automatical segmentation of cupola blocks from unorganized point clouds. Proceedings of GeoSpace Congress in Kiev.
- DORE, C., MURPHY, M., MCCARTHY, S., BRECHIN, F., CASIDY, C. & DIRIX, E., 2015: Historic Building Information Model (HBIM). The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **40**(5/W4), 351-357.

- FAN, Z., XIANFENG, H., ZHICHAO, Z., WEI, F. & DEREN, L., 2012: High precision texture reconstruction for 3d sculpture model. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39**(B5), 139-143.
- GEVAERT, C., PERSELLO, C., SLIUZAS, R. & VOSSELMANN, G., 2016: Classification of informal settlements through the integration of 2d and 3d features extracted from UAV data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **III**(3), 317-324.
- HARRIS, C. & STEPHENES, M., 1988: A combined corner and edge detector. *Proceedings of Fourth Alvey Vision Conference*, 147-151.
- KIVILCIM, C. & DURAN, Z., 2016: A semi-automated point cloud processing methodology for 3d cultural heritage documentation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41**(B5), 293-296.
- LOWE, D., 1999: Object recognition from local scale-invariant features. *Proceedings of the Seventh IEEE International Conference on Computer Vision* **2**, 1150-1157.
- LOWE, D., 2004: Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision* **60**(2), 91-110.
- QUATTRINI, R., MALINVERNI, E. S., CLINI, P., NESPECA, R. & ORLIETTI, E., 2015: From TLS to BIM. High Quality semantically-aware 3D Modelling of Complex Architecture. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40**(5/W4), 367-374.
- RUSU, R., BLODOW, N., MARTON, Z. & BEETZ, M., 2008: Aligning point cloud views using persistent feature histograms. *Proceedings of IEEE/RSJ International conference in Intelligent Robots and Systems*, 3384-3391.
- WANG, W., POTTMANN, H. & LIU, Y., 2004: Fitting b-spline curves to point clouds by squared distance minimization. Technical report. HKU CS Tech Report TR-2004- 11, <http://www.cs.hku.hk/research/techreps/document/TR-2004-11.pdf>.

Untersuchung der Eignung photogrammetrischer Methoden zur Erzeugung von 3D-Punktwolken aus historischen Bilddatenbeständen

Theresa Vietze¹, Danilo Schneider¹ & Ferdinand Maiwald¹

Zusammenfassung: In diesem Beitrag wird untersucht, inwiefern bildbasierte Methoden, insb. Structure-from-Motion geeignet sind, um aus historischen Fotografien 3D-Punktwolken zu erzeugen. Dabei werden Probleme aufgezeigt und Lösungen vorgestellt. Neben zu geringer Auflösung der Bilder und unbekannter Kamerageometrie, wurde als wesentliche Schwierigkeit die für die 3D-Rekonstruktion ungünstige Aufnahmekonfiguration (nur wenige Aufnahmerichtungen mit großer Basis) identifiziert, die mit Hilfe aktueller Aufnahmen desselben Objektes teilweise kompensiert werden konnte. Diese Möglichkeit besteht jedoch nur bei Objekten, die sich nicht signifikant verändert haben und aktuell noch existieren, was bei dem für diese Untersuchungen verwendeten Objekt, dem Kronentor des Dresdner Zwingers, der Fall war. Die aus historischen Fotografien erstellten 3D-Punktwolken bzw. daraus entwickelten 3D-Modellen können später genutzt werden, um historische Zustände eines Objektes „erlebbar“ zu machen, um bauliche Veränderungen nachzuweisen oder um wertvolle historische Objekte für die Zukunft zumindest virtuell zu „konservieren“.

1 Einleitung

Als Informationsquelle für die geschichtliche Forschung existieren neben Texten und Kunstwerken seit etwa 200 Jahren auch photographische Abbildungen. Historische Aufnahmen von bereits zerstörten oder veränderten Architekturobjekten besitzen einen hohen Informationsgehalt und stellen für die Geschichtswissenschaft und -didaktik einen großen Wert da. In der Photogrammetrie können aus Fotos, die aus unterschiedlichen Richtungen aufgenommen wurden mittels Methoden wie Structure-from-Motion (SfM) oder Stereoauswertung dreidimensionale Modelle erstellt werden. Ziel der Untersuchungen ist die Generierung eines dreidimensionalen Modells eines Objektes aus historischen Aufnahmen. Hierbei entstehen unterschiedliche Probleme für diese photogrammetrischen Methoden. Oft wurde ein Objekt nur aus demselben Blickwinkel oder von der „schönsten“ Seite fotografiert. Meist ist es aus diesem Grund nicht möglich ein vollständiges 3D-Modell zu rekonstruieren bzw. müssten die entsprechenden Objektbereiche mittels Annahmen des Bearbeiters interaktiv modelliert werden. Deshalb wurde für diesen Beitrag ein Objekt gewählt, für das eine hohe Anzahl von Aufnahmen vorhanden ist. Anhand des Bilddatenbestands der Fotothek der Sächsischen Landesbibliothek – Staats- und Universitätsbibliothek (SLUB) wird das Kronentor des Dresdner Zwingers rekonstruiert. Den Untersuchungen ging eine umfassende Literaturrecherche voraus, die aufzeigen soll, welche verschiedenen Ansätze zur photogrammetrischen Verarbeitung von historischen Aufnahmen bereits getestet wurden.

¹ Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstraße 10, D-01062 Dresden, E-Mail: theresa.vietze@mailbox.tu-dresden.de, danilo.schneider@tu-dresden.de, ferdinand.maiwald@tu-dresden.de

Vor der Verarbeitung der historischen Aufnahmen mittels SfM wurde das Datenmaterial nach Verwendbarkeit sortiert. Dabei wurde der Datenbestand subjektiv nach verschiedenen Kriterien reduziert. Aus den verbliebenen Aufnahmen wurden anschließend verschiedene Modelle erzeugt. In diesem Bearbeitungsschritt entstanden verschiedene Probleme, welche der Beitrag näher beschreibt. Dies beinhaltet größtenteils das Problem der vielen Aufnahmen von nahezu identischen Standorten. Es wird darauf eingegangen, wie dies behoben werden kann, z.B. durch Nutzung aktueller Fotos zur Orientierung weiterer historischer Aufnahmen. Abschließend wird die Qualität der erzeugten Modelle analysiert und mit anderen Modellen (Punktwolken erstellt aus aktuellen Bildern und mittels terrestrischem Laserscanning) verglichen.

1.1 Stand der Forschung

Objekte, die bisher häufig untersucht werden, sind archäologische Funde bzw. Fundorte sowie im Laufe der Zeit zerstörte oder veränderte Architektur. Oft liegen in diesen Fällen nur wenige und nicht speziell auf photogrammetrische Verfahren ausgerichtete Aufnahmen vor. Für archäologische Untersuchungen sollten beispielsweise im Jahre 1940 gefundene Dinosaurierspuren aus 17 Fotografien rekonstruiert werden (FALKINGHAM et al. 2014). Die einzelnen Spuren wurden im Laufe der Zeit ausgegraben und an verschiedene Museen verteilt oder sie sind verschollen. Da verschiedene Zeichnungen und die Fotografien des Gebietes zum Zeitpunkt der Entdeckung vorlagen, konnte die Erstellung eines 3D-Modelles mittels Structure-from-Motion (SfM) dazu beitragen, den ursprünglichen korrekten Verlauf der Spuren wiederherzustellen. Ein weiteres Beispiel ist die architektonische Rekonstruktion von serbischen Kirchen (STOJAKOVIC & TEPAV 2009). Aufgrund geringer Erhaltungsmaßnahmen sind dort oft Dächer oder Mauern eingefallen. Die Teile, die noch bestehen, wurden in dieser Untersuchung mit einem Laserscanner aufgenommen. Gebäudeteile, die zerstört waren, konnten durch historische Aufnahmen manuell wiederhergestellt werden. Aus der Verknüpfung beider Verfahren konnten verschiedene Kirchen neu modelliert werden (STOJAKOVIC & TEPAV 2009). Das Internet ist gefüllt mit den verschiedensten Fotos und Schnappschüssen von Sehenswürdigkeiten und Kulturdenkmälern. Mittels einer konkreten Suche nach diesen Touristenfotos konnten große Teile Roms in nur 24 Stunden mittels SfM rekonstruiert werden (AGARWAL et al. 2009). Des Weiteren konnte der Baaltempel in Palmyra durch touristische Fotografien und professionelle Panoramaaufnahmen dreidimensional rekonstruiert und virtuell für die Nachwelt erhalten werden (WAHBEH & NEBIKER 2016). Schließlich soll ein letztes Beispiel zeigen, wie die erzeugten Ergebnisse in ein Stadtmodell integriert werden können. Bei „4D Cities“ Projekten wurden historische Stadtmodelle erstellt, welche verschiedene zeitliche Epochen der städtischen Entwicklung zeigen (SCHINDLER & DELLAERT 2012; KERSTEN et al. 2012). Dadurch konnten Veränderungen an ausgewählten Gebäuden sehr schnell erkannt und zeitlich festgehalten werden.

1.2 Informationen zum Untersuchungsobjekt „Kronentor“

Der Zwinger ist eines der meistbesuchten Sehenswürdigkeiten der Stadt Dresden. Das Kronentor bildet dabei den westlichen Eingang in die Schlossanlage. Zu beiden Seiten schließen sich zwei Langgalerien an vor denen sich ein Wassergraben entlang zieht. Über diesen führt eine schmale Brücke zum Kronentor hin. Das Kronentor wurde 1714 als eine Art Tor mit Turm in barocker Architektur errichtet. Zwei Säulenpaare bilden die Hauptbestandteile des Durchganges. Das Tor

ist zweigeschossig angelegt, wobei das Obergeschoss nur über die Dachgänge der Seitenflügel erreicht werden kann. Eine Attika schließt die nach allen vier Seiten geöffnete Halle des oberen Stockwerks. Das Kronentor wird durch eine zwiebelartige Kuppel mit einer goldenen Krone abgeschlossen.

1945 wurden große Teile des Zwingers, unter anderem auch das Kronentor, stark beschädigt. Nach Kriegsende folgte dann der Wiederaufbau. Später wurde die Kuppel mit Goldbeschlägen versehen. Zwischen 2012 und 2016 wurde das Kronentor komplett saniert. Mittels historischer Bilder können ursprüngliche Bauzustände photogrammetrisch rekonstruiert werden. Dabei besteht die Möglichkeit, bauliche Veränderungen durch Vergleiche mit aktuellen Modellen aufzuzeigen.

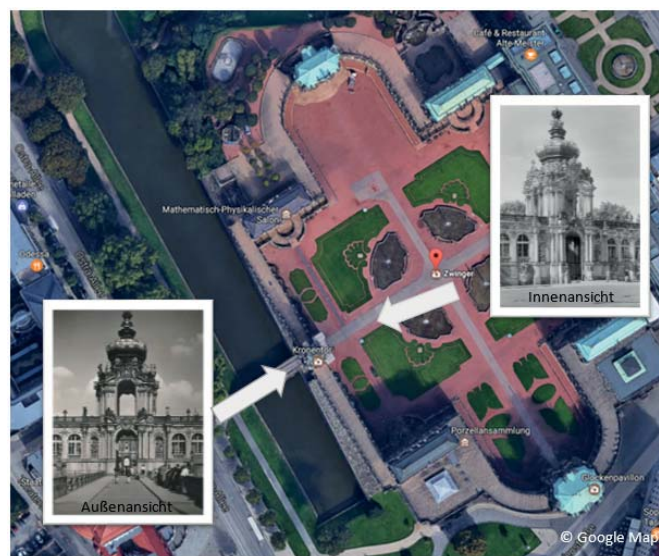


Abb. 1: Geländeübersicht des Zwingers mit Kronentor (aus Google Maps)

2 Datenquellen

Für den Bilddatenbestand des Kronentores des Dresdner Zwingers wurden aus der Mediathek der Sächsischen Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB) zunächst mehr als 800 historische Fotos aus unterschiedlichen Zeitepochen mittels Schlagwort-Suche gefiltert. Dabei konnte der Datensatz schon durch detailliertere Suchkriterien auf 499 Bilder reduziert werden. Dies stellt auf den ersten Blick zwar einen sehr großen Datenbestand dar, der sich aber anhand von weiteren Auswahlkriterien auf 44 Aufnahmen beschränken ließ. Diese Auswahlkriterien lassen sich in zwei Gruppen einteilen. Zunächst sind alle Aufnahmen entfernt worden, die augenscheinlich nicht das Kronentor zeigen bzw. nicht für eine spätere 3D-Rekonstruktion mittels Structure-from-Motion (SfM) geeignet sind. Dazu zählen unter anderem Zeichnungen, Postkarten und Skizzen. Auch Nachtaufnahmen und Aufnahmen, auf denen das Kronentor eingerüstet ist, verfälschen das erzielte Ergebnis. Aufnahmen mit großer Entfernung zum Objekt bieten nicht die nötige Detailgenauigkeit für eine Rekonstruktion. Es wurde dabei beachtet, dass der Bildausschnitt das Kronentor in voller Größe darstellt. Insgesamt konnte der Datensatz von 499 auf 89 Bildern reduziert werden. Daran anschließend wurde durch Beurteilung

der Qualität der Fotos eine weitere Selektion durchgeführt. Bilder mit einer zu geringen Scanauflösung eignen sich oft nicht für SfM, da bestimmte Details nicht wiedergegeben bzw. auf dem Foto dargestellt sind. Folglich wurden meist zu wenige Merkmalspunkte gefunden, wodurch diese Aufnahmen entweder gar nicht oder nur ungenau orientiert werden konnten. Auch beim anschließenden „Dense Matching“ wurden viele falsche Punktzuordnungen festgestellt, welche zu fehlerhaften Punkten in der Punktwolke führten. Überstrahlung oder Unschärfe verfälschen ebenso das Ergebnis. Einige wenige Bilder mussten aufgrund ihrer Qualität aussortiert werden, da sie sich stark verzogen hatten oder Knicke aufwiesen. Abschließend konnte ein Datensatz mit 44 Bildern erstellt werden (Beispielbilder in Abb. 2), von dem auszugehen war, dass dieser für die Erstellung von Punktwolken mittels SfM prinzipiell geeignet war.

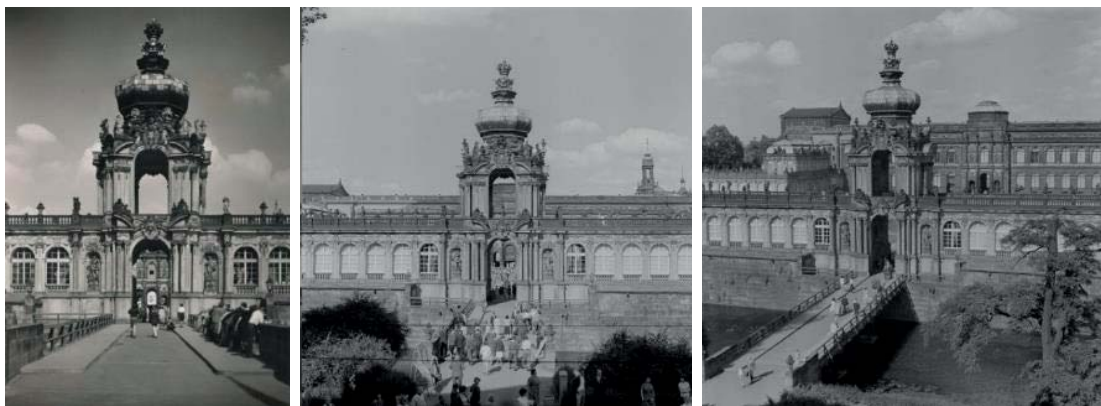


Abb. 2: Zur Punktwolken-Generierung verwendete historische Bilder des Dresdner Zwingers (Auswahl)

Bei Betrachtung des neuen Datensatzes war jedoch festzustellen, dass sich Aufnahmen von einigen Richtungen häuften. Von einigen Standpunkten gab es deutlich mehr Fotografien als von anderen Ansichten auf das Gebäude. Die historischen Aufnahmen stammen vor allem von Privatpersonen oder Zeitungsfotografen, welche das Kronentor nicht nach photogrammetrischen Aspekten fotografierten. Dies stellt für SfM ein Problem dar, da die Überlappungsbereiche der Fotos von zwei verschiedenen Standpunkten somit nicht hoch genug sind, um diese miteinander verknüpfen zu können. Außerdem ist die Basis zwischen den Bildern zu groß, um zuverlässig homologe Bildpunkte bestimmen zu können. Deshalb werden Fotos von dazwischenliegenden Standpunkten benötigt, welche aber entweder nicht existieren oder eine zu geringe Auflösung aufweisen. Die örtlichen Gegebenheiten schränken den Fotografen zusätzlich in der Wahl des Aufnahmestandpunktes ein. Somit können die Aufnahmen in zwei Gruppen unterteilt werden: Aufnahmen, von außerhalb des Zwingers und Aufnahmen, deren Standpunkt innerhalb lag (s. Abb. 1). Für diese zwei Gruppen wurde nachfolgend versucht zwei verschiedene Punktwolken zu erstellen. Anschließend sind diese noch in ein gemeinsames Koordinatensystem zu überführen.

3 Auswertung und Ergebnisse

Mittels Structure-from-Motion (LOWE 2004) wurde in dem Programm PhotoScan von Agisoft getestet, ob aus den historischen Aufnahmen Punktwolken erstellt werden können. Durch

vorangehende Recherchen und Erfahrungen versprach dieses Programm die besten Ergebnisse (REMONDINO et al. 2012; SANTAGATI et al. 2013; AGISOFT 2016). Alternativ wurden auch andere Programme, wie z.B. VisualSfM getestet. Der Vergleich der Programme soll aber hier nicht im Vordergrund stehen.

In diesem Kapitel werden Punktwolken gezeigt, welche aus den zuvor ausgewählten historischen Aufnahmen erstellt wurden. Dabei stellt eine Punktwolke die Seite des Kronentors dar, welche außerhalb des Zwingers liegt. Die andere Punktwolke repräsentiert den vom Schlosshof aus sichtbaren Objektbereich (Abb. 1). Die Kalibrierung der Kameras wird mittels Selbstkalibrierung von PhotoScan bestimmt. Die erstellten Punktwolken sind bisher noch maßstabslos und nicht absolut orientiert.

3.1 Generierung von Punktwolken ausschließlich aus historischen Aufnahmen

Zur Erstellung der Punktwolke, welche das Kronentor von außen zeigen soll, wurden zunächst 25 für qualitativ hochwertig empfundene Bilder in PhotoScan eingeladen. Bei dem Arbeitsablauf war jedoch festzustellen, dass einige Fotos nicht genügend überlappende Bereiche zu anderen aufweisen. In PhotoScan konnten deshalb automatisch lediglich 11 Bilder relativ zueinander orientiert und anschließend zum Generieren einer Punktwolke (Abb. 3) verwendet werden.



Abb. 3: Aus historische Bildern erstellte Punktwolke (Ansicht von außerhalb des Zwingers)

Zunächst einmal kann das Kronentor als solches erkannt werden. Kleine Details, wie z.B. die Sandsteinfiguren auf den Seitenflügeln sind nur schemenhaft zu erahnen, wie in Abb. 4 zu erkennen ist. Dies liegt daran, dass die gegebene Auflösung der Aufnahmen die feinen Details nicht wiedergeben kann. Der Zwinger mit Kronentor wurde 1945 teilweise zerstört. Durch den Wiederaufbau wurden kleine bauliche Details verändert bzw. Figuren ersetzt. Die genauen Aufnahmedaten der Bilder sind nicht bekannt. Es kann deshalb vermutet werden, dass Aufnahmen aus unterschiedlichen Bauepochen in den Bildverband eingegliedert wurden. Die baulichen Veränderungen führen zu Rekonstruktionsfehlern.



Abb. 4: Vergleich eines Ausschnitts aus historischer Aufnahme (links) und Punktwolke (rechts)

Es ist festzustellen, dass die ausgewählten Fotografien meist aus zentraler Position aufgenommen wurden (Abb. 5). Deshalb ist einerseits das Modell an einigen Stellen aufgrund von Verdeckungen lückenhaft. Andererseits führt die sehr kurze Basis zwischen den Kamerastandpunkten zu schleifenden Schnitten und verschlechtert damit die Genauigkeit der Bestimmung der Objektpunkte. Um ein besseres Ergebnis erzielen zu können, müssen die Randaufnahmen, welche zunächst nicht mit orientiert werden konnten, für eine bessere Strahlenschnittkonfiguration einbezogen werden.

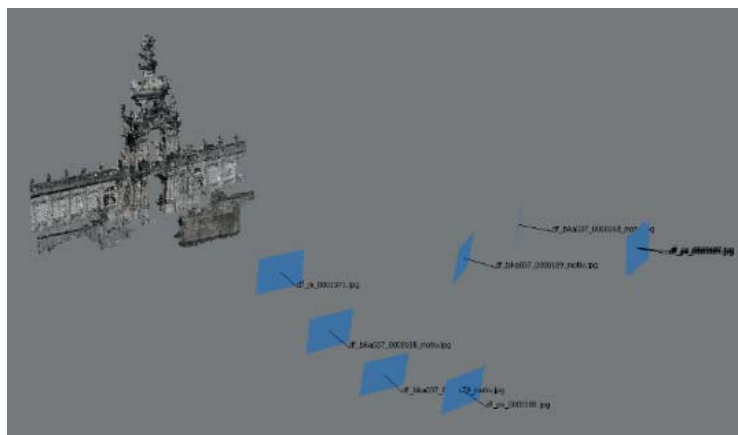


Abb. 5: Historische Punktwolke mit Aufnahmestandpunkten der verwendeten Aufnahmen

Die Grundlage für eine zweite Punktwolke bildeten Fotografien, welche von innerhalb des Zwingers aufgenommen wurden. Es ist jedoch festzustellen, dass ausschließlich Fotografien, welche von der linken Blickrichtung aufgenommen wurden, den Bildverband bilden. Aufnahmen von anderen Standpunkten, die die Strahlenschnittgeometrie stärken könnten, wurden wegen zu geringer Auflösung nicht mit einbezogen. Einige der verwendeten Bilder wurden außerdem von einem weiter entfernten Standpunkt aufgenommen (Abb. 6). Die gegebene Auflösung am Objekt reicht bei diesen nicht, um genaue Details am Kronentor zu erkennen. Um ein zufriedenstellendes Ergebnis erzielen zu können, ist eine Digitalisierung der analogen Fotografien mit deutlich höherer Auflösung notwendig. Dies setzt voraus, dass die analoge Vorlage ein größeres Auflösungspotential bietet als die aktuelle Digitalisierung darstellen kann. Es wird jedoch davon ausgegangen, dass ein hochauflösender Scan bei vielen Fotografien möglich ist, da häufig nur das Vorschau bild der Fotothek der SLUB abgespeichert wurde.

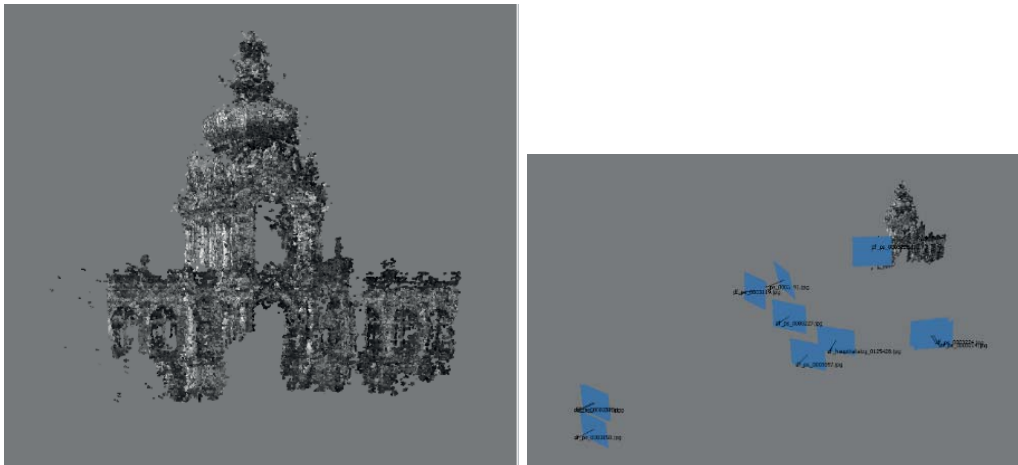


Abb. 6: Historische Punktwolke (Ansicht vom Zwingerhof) mit Aufnahmestandpunkten (rechtes Bild)

3.2 Generierung einer Punktwolke aus historischen und aktuellen Aufnahmen

Die folgenden Untersuchungen beziehen sich nur auf die Außenseite des Kronentors (Kap. 3.1). Zur Orientierung von weiteren historischen Aufnahmen werden aktuelle Bilder als „Füllmaterial“ verwendet, um mit kleineren Basisabständen zwischen den Bildern schrittweise alle Bilder hinzuzufügen. Somit können historische Bilder, welche am Rand liegen, mit in die Punktwolke einbezogen werden.

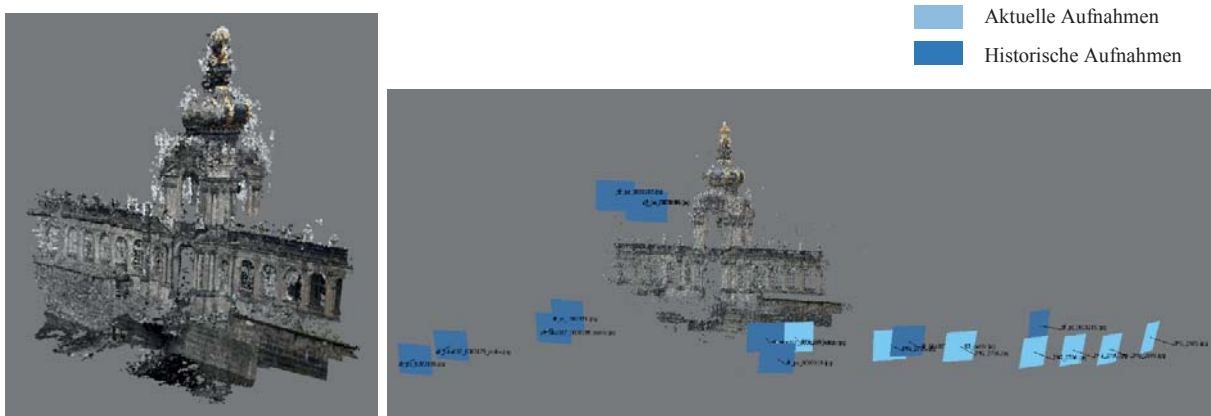


Abb. 7: Punktwolke erstellt aus historischen und aktuellen Bildern sowie Aufnahmestandpunkte

In Abb. 7 ist zu erkennen, dass der Detailgrad und die Anzahl der Punkte in der Punktwolke deutlich durch die aktuellen Aufnahmen gestiegen sind (von ca. 1,3 auf 3 Millionen Punkte in der „Dense Cloud“). Die historischen Punkte wurden durch aktuelle überlagert. Deutlich ist dies durch die bauliche Veränderung der Krone zu erkennen. Die historischen Aufnahmen zeigen die Goldbeschläge auf dem Dach nicht. Damit verfälschen die aktuellen Bilder die historische Punktwolke. Deshalb müssen diese Punkte herausgenommen werden, was im nachfolgenden Kapitel detaillierter beschrieben wird.

3.3 Generierung einer Punktwolke aus historischen Aufnahmen unter Nutzung aktueller Fotografien zur Unterstützung der Bildorientierung

Es soll nachfolgend eine Punktwolke erstellt werden, bei welcher die aktuellen Bilder nur zur Unterstützung der Bestimmung der Orientierung der historischen Bilder einbezogen wurden. Zum einen kann die zuvor berechnete Orientierung eines Bildes (Kap. 3.2) gespeichert und in das neue Projekt importiert werden. Dies bedeutet, dass im gesamten Workflow nur historische Aufnahmen verwendet werden. Zum anderen kann in dem bereits erstellten Projekt angegeben werden, dass die aktuellen Fotos bei der Generierung der „Dense Cloud“ nicht beachtet werden sollen. Resultierend wird in beiden Fällen eine Punktwolke erstellt, bei welcher alle Punkte ausschließlich aus historischen Aufnahmen stammen. Durch dieses Vorgehen konnten zwei weitere historische Fotos in den Bildverband mit integriert werden. Im Vergleich zur ersten Punktwolke konnten einige Lücken gefüllt bzw. auch an einigen Stellen der Detailgrad verbessert werden (Abb. 8). Auch die Anzahl an Punkten in der „Dense Cloud“ konnte von ca. 1,3 auf 1,7 Millionen gesteigert werden.



Abb. 8: Vergleich der Punktwolken basierend auf 11 (links) bzw. 13 (rechts) historischen Bildern

3.4 Vergleich der Punktwolken basierend auf historischen Bildern mit Punktwolken aus aktuellen Bildern sowie terrestrischen Laserscannerdaten

Abbildung 9 zeigt den Vergleich der historischen Punktwolke mit einer Punktwolke, welche aus 40 ausschließlich aktuellen Bildern erstellt wurde. In beiden Fällen sind die gleichen Einstellungen in PhotoScan verwendet worden. Es ist deutlich zu erkennen, dass die aktuelle Punktwolke nach visuellem Eindruck weniger verrauscht dargestellt und Details differenzierter abgebildet werden. Die dritte abgebildete Punktwolke wurde mit einem terrestrischen Laserscanner Riegl LMS-Z420i von insgesamt 6 rund um das Kronentor verteilten Standpunkten erzeugt und zeigt den im Vergleich höchsten Detailgrad und das geringste Rauschen.



Abb. 9: Vergleich der Punktwolken basierend auf aktuellen (links) bzw. historischen Bildern (Mitte) sowie terrestrischen Laserscannerdaten (rechts)

4 Fazit & Ausblick

Grundsätzlich ist es möglich aus historischen, digitalisierten Analogaufnahmen dreidimensionale Punktwolken und daraus triangulierte und texturierte 3D-Modelle zu erstellen. Der Beitrag zeigt, welche Anforderungen an die Fotos gestellt werden müssen bzw. wie der Structure-from-Motion-Workflow beeinflusst werden kann, um Punktwolken bzw. Modelle mit hoher visueller Qualität zu erstellen. Zu berücksichtigende Faktoren sind zum Beispiel die Aufnahmekonfiguration oder vorhandene Informationen über die Aufnahmetechnik. Aber auch die Auflösung und die Qualität der Aufnahmen spielen eine große Rolle. Außerdem werden im Beitrag alternative Möglichkeiten zur Erzeugung historischer 3D-Modelle diskutiert. Insbesondere wird auf die Integration aktueller und historischer Fotos in einer gemeinsamen SfM-Auswertung eingegangen. Diese Möglichkeit besteht natürlich nur, sofern das Objekt sich baulich über die verschiedenen Zeitepochen hinweg nicht signifikant verändert hat und aktuell noch existiert und auch erreichbar ist. Bei Verfügbarkeit sehr weniger geeigneter Bilder bzw. bei zu großen Abständen zwischen den Aufnahmepositionen wurden Teilmodelle des Objektes erstellt. Diese können gegebenenfalls in einer späteren Weiterverarbeitung zusammengefügt werden.

Im nächsten Schritt soll untersucht werden, inwieweit sich eine hochauflösendere Digitalisierung der verwendeten historischen Fotografien auf die Qualität der Punktwolken auswirkt. Dafür sollen möglichst objektive Qualitätsmetriken, u.a. für die Genauigkeit und Vollständigkeit definiert werden, mit Hilfe derer die Qualität verschiedener Punktwolken verglichen und analysiert werden kann.

Es soll ferner untersucht werden, ob bei ungenügender Qualität der vorhandenen historischen Fotos möglicherweise ein interaktiver, auf der Modellierung von Objektgeometrien basierender Ansatz mit Blick auf die Verwendung der Modelle im Rahmen der Geschichtswissenschaft und -didaktik zu besseren Ergebnissen führen kann. Außerdem bietet sich die Orientierung einzelner historische Fotos an, welche anschließend lediglich als Texturen für virtuelle 3D-Modelle verwendet werden.

5 Literaturverzeichnis

- AGARWAL, S., SNAVELY, N., SIMON, I., SEITZ, S. M. & SZELISKI, R., 2014: Building rome in a day. IEEE 12th international conference on computer vision, 72-79.
- AGISOFT, 2016: Agisoft PhotoScan User Manual. http://www.agisoft.com/pdf/photoscan-pro_1_2_en.pdf, letzter Zugriff am 20.1.2017.
- FALKINGHAM, P., BATES, K. & FARLOW, J., 2014: Historical Photogrammetry: Bird's Paluxy River Dinosaur Chase Sequence Digitally Reconstructed as It Was prior to Excavation 70 Years Ago. PLoS ONE **9**(4), 1-5.
- KERSTEN, T. P., KELLER, F., SAENGER, J. & SCHIEWE, J., 2012: Automated generation of an historic 4d city model of Hamburg and its visualisation with the GE engine. EuroMed 2012 - Int. Conference on Cultural Heritage, IOANNIDES, M., FRITSCH, D., LEISSNER, J., DAVIES, R., REMONDINO, F. & CAFFO, R. (Eds.), Lecture Notes in Computer Science (LNCS) **7616**, Springer-Verlag Berlin Heidelberg, 55-65.
- LOWE, D. G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision **60**(2), 91-110.
- REMONDINO, F., DEL PIZZO, S., KERSTEN, T. P. & TROISI, S., 2012: Low-cost and Open-Source Solutions for Automated Image Orientation – a Critical Overview. EuroMed 2012 - Int. Conference on Cultural Heritage, IOANNIDES, M., FRITSCH, D., LEISSNER, J., DAVIES, R., REMONDINO, F. & CAFFO, R. (Eds.), Lecture Notes in Computer Science (LNCS) **7616**, Springer-Verlag Berlin Heidelberg, 40-54.
- SANTAGATI, C., INZERILLO, L. & DI PAOLA, F., 2013: Image-based modeling techniques for architectural heritage 3D digitalization: Limits and potentialities. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **40**(5/W2), 555-560.
- SCHINDLER, G. & DELLAERT, F., 2012: 4d cities: analyzing, visualizing, and interacting with historical urban photo collections. Journal of Multimedia **7**(2), 124-131.
- STOJAKOVIC, V. & TEPAVCEVIC, B., 2009: Optimal methods for 3d modeling of devastated architectural objects. Proceedings of the 3rd ISPRS International Workshop 3D-ARCH **38**, 5.
- WAHBEH, W. & NEBIKER, S., 2016: Kombination unterschiedlicher Photogrammetrischer Ansätze zur Rekonstruktion zerstörter Kulturdenkmäler. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **25**, KERSTEN, T. (Hrsg.), 338-350.

Entwicklung eines virtuellen Museums für ein historisches Gebäude am Beispiel des Alt-Segeberger Bürgerhauses

SIMON DEGGIM¹, FELIX TSCHIRSCHWITZ¹ & THOMAS P. KERSTEN¹

Zusammenfassung: Ergänzend zum traditionellen Museumsbesuch bieten sich Virtuelle Museen an, um Informationen in neuartiger, komfortabler und besonders informativer Weise zu vermitteln. Zusammenhänge können visualisiert und interaktiv erfahrbar gemacht werden, um das Verständnis für ein bestimmtes Thema zu erleichtern. Am Beispiel des Museums Alt-Segeberger Bürgerhaus in Bad Segeberg wurde ein solches Programm mithilfe der Spiele-Engine „Unreal Engine“ umgesetzt. Der Benutzer erhält somit die Möglichkeit, sich durch ein auf Basis von Vermessungsdaten erstelltes 3D-Modell des Gebäudes zu bewegen und spielerisch Informationen wie z.B. die Baugeschichte des Gebäudes in animierter Form abzurufen. Weiterentwicklungen dieses Konzepts, z.B. für Virtual Reality Systeme wie die HTC Vive, bieten ein großes Potential zur besseren Informationsvermittlung.

1 Einleitung

Museen stehen vor der Aufgabe, Informationen und Zusammenhänge in einer Weise zu vermitteln, dass fachfremde Personen diese nicht nur in kurzer Zeit verstehen, sondern dass bestenfalls auch ein weitergehendes Interesse am gezeigten Themengebiet geweckt wird. Gemäß ihrem Bildungsauftrag müssen Museen permanent Möglichkeiten entwickeln und anbieten, komplexe Sachverhalte gleichzeitig informativ und unterhaltsam aufbereitet einem möglichst breitem Zielpublikum zugänglich zu machen, da den Lernenden wegen des spezifischen Vorwissens, den Vorerfahrungen, individuellen Interessen und Zielen eine aktive Rolle zugeordnet wird (REUSSNER 2007). Diese Grundgedanken werden heute zunehmend in sog. Serious Games umgesetzt, die Informationen in eine virtuelle Welt einbetten und durch einen Spielfluss unterhaltsam erlebbar machen (MORTARA et al. 2014).

Für den musealen Bereich wird die Umsetzung von Kultur und Informationstechnologie oft als Virtuelles Museum (VM) bezeichnet. VMs nutzen verschiedene Medien wie Text, Bild, Ton und animierte 3D-Modelle, um als eine interaktive Plattform zur informativen Ergänzung des Museumsbesuchs zu fungieren (SAMIDA 2002). Die Gestaltungsform variiert dabei von einfachen Webseiten (BAUER 2001) über auf Panoramen basierenden virtuellen Touren (KERSTEN & LINDSTAEDT 2017) bis zu interaktiven Apps für Smartphones oder Tablets (GÜTT 2010).

Verschiedene Arten eines VMs können dabei unterschiedliche Schwerpunkte setzen: Ein über das Internet abrufbares Programm würde die Möglichkeit bieten, einen virtuellen Museumsbesuch zeit- und ortsunabhängig durchzuführen sowie den eigentlichen

¹ HafenCity Universität Hamburg, Labor für Photogrammetrie & Laserscanning, Überseeallee 16, D-20457 Hamburg, E-Mail: [Simon.Deggim, Felix.Tschirschwitz, Thomas.Kersten]@hcu-hamburg.de

Museumsbesuch vor- und nachzubereiten, in dem über dieses Medium die Aufmerksamkeit des Besuchers angeregt wird und zudem weiterführende Informationen bereitgestellt werden. Die große Stärke liegt aber in der ergänzenden Präsentation und der Einbindung in die reale Ausstellung, um Inhalte zu vermitteln, die mit klassischen Ausstellungsmethoden nicht möglich sind und durch aktuelle Informationstechnologie erheblich zum Verständnis beitragen können. In diesem Beitrag werden anhand des konkreten Beispiels des Museums Alt-Segeberger Bürgerhaus Möglichkeiten aufgezeigt, wie Museen den technischen und medialen Anforderungen im 21. Jahrhundert auch mithilfe von detaillierten Geodaten und entsprechender Informationstechnologie gerecht werden können.

2 Das Alt-Segeberger Bürgerhaus

Das Alt-Segeberger Bürgerhaus war bereits Ende des 19. Jahrhunderts bekannt als das älteste Haus der Stadt Segeberg und ist heute eines der wenigen erhaltenen kleinstädtischen Bürgerhäuser Schleswig-Holsteins aus der beginnenden Frühen Neuzeit (Abb. 1). Bereits im ca. 1539 neu angelegten Ratsbuch wurde das Gebäude im sog. Grundhauerverzeichnis aufgeführt. Nachdem während der Grafenfehde 1533-1536/37 Segeberg im Juni 1534 fast vollständig zerstört wurde, wurde das Bürgerhaus 1541 zunächst eingeschossig als einfaches Hallenhaus in Fachwerkbauweise mit Ziegelausfachung auf dem Grundstück der heutigen Lübecker Straße 15 wiedererrichtet. Die Bauweise war notdürftig und teilweise wurde dabei Baumaterial benachbarter Trümmergrundstücke (z.B. im Dachstuhl) zweitverwendet, während das Grundgerüst des Fachwerkbaus aus Frischholz (Eiche) errichtet wurde (REIMERS & HINRICHSEN 2015). Vermutlich bereits vom mittelalterlichen Vorgängerbau stammte der Keller mit den Wänden aus Findlingen.



Abb. 1. Frontansicht von vor 1963 (links), von heute (Mitte) und Grundriss vom Erdgeschoss des Museums Alt-Segeberger Bürgerhaus im heutigen Zustand (rechts).

In den folgenden Jahrhunderten wurde das Haus etliche Male erweitert und umgebaut. Eine kurze Zusammenfassung der verschiedenen Bauphasen ist in KERSTEN et al. (2013) zusammengestellt. Nach seiner Grundsanierung 1963/64 beinhaltete es das „Heimatmuseum“ der Stadt Bad Segeberg. In seinen historischen Räumen wurden in den ersten Jahrzehnten Exponate

aus der kleinbürgerlichen Wohn- und Arbeitswelt des 19. und 20. Jahrhunderts gezeigt. Nach Übernahme der Trägerschaft durch die Volkshochschule Bad Segeberg im Jahre 2012 entstehen unter dem neuen Namen „Museum Alt-Segeberger Bürgerhaus“ sukzessive zwei neue Dauerausstellungen zu Themen „500 Jahre Entwicklung bürgerlicher Wohnkultur im Spiegel eines 470 Jahre alten Hauses“ und „800 Jahre Segeberger Stadtgeschichte – Von der mittelalterlichen Burgsiedlung zum modernen Kurort“.

3 Die Game-Engine Unreal

Eine Game-Engine ist eine Sammlung von Werkzeugen bzw. Komponenten zur Erstellung von Computerspielen. Notwendige Komponenten sind die Rendering-Engine zur Darstellung texturierter 3D-Modelle (Geodaten), ein Audiosystem zur Ausgabe von Tönen und eine Schnittstelle zur Programmierung der Spiele-Logik. Weitere gängige Komponenten erlauben die vereinfachte Entwicklung weiterer spieletypischer Anwendungen: a) Schnittstellen zur Netzwerktechnik, b) Steuerung des Nutzerblickwinkels und c) Berechnung von realistischer Interaktion von Objekten durch eine entsprechende Physik-Engine. Für die Entwicklung eines virtuellen Museums bieten Game-Engines viele notwendige Konzepte mit zahlreicher Funktionalität an, damit die Nutzer mit dem VM interagieren können.

Die Entwicklung von Game-Engines geschah in der Vergangenheit meist anhand eines zu entwickelnden Spiels und anschließender kostenpflichtiger Lizenzierung an externe Spieleentwickler. In den letzten Jahren haben die meisten großen Engine-Anbieter sich allerdings stärker auf die Weiterentwicklung der Engines konzentriert und bieten nun zusätzlich kostenlose Zugänge für Entwickler an. Beispiele für Game-Engines mit kostenloser Nutzungsmöglichkeit sind bspw. die Engine Unity von Unity Technologies, die CryEngine vom deutschen Entwicklungsstudio Crytek sowie die hier verwendete Unreal Engine von Epic Games. Einen aktuellen Überblick und Vergleich unterschiedlicher Game-Engines findet sich z.B. bei O'FLANAGAN (2014) und LAWSON (2016). Die Auswahl der geeigneten Engine für ein Projekt erfolgt anhand der bereitgestellten oben beschriebenen Komponenten, der Anpassbarkeit in die bereits vorhandenen Arbeitsabläufe sowie spezieller Vorlieben der (Spiele)-Entwickler. Im Rahmen dieser Arbeit wurde sich bspw. auch deshalb für die Unreal Engine entschieden, da mithilfe einer visuellen Programmiersprache Anwendungs- und Interaktionslogiken entwickelt werden können, den sog. Blueprints. Visuelle Programmierung wie mit den Blueprints erfordert nicht, maschinenkonformen Quelltext zu schreiben, sondern sie bietet Möglichkeiten für Nicht-Informatiker, alle relevanten Funktionen eines VM durch grafische Elemente zu programmieren. Der damit ebenfalls verbundene Geschwindigkeitsgewinn bei der Entwicklung ermöglicht die Generierung weiterer Szenarien sowie intensivere Nutzertests. Game-Engines sind daher sehr gut geeignet, um Virtuelle Museen zu entwickeln.

4 Modellierung

Das gesamte Museum Alt-Segeberger Bürgerhaus wurde 3D modelliert, so dass die Besucher das Gebäude außen und innen virtuell erkunden können, verknüpft mit der Möglichkeit, verschiedene Ausstellungsstücke im Museum näher betrachten zu können. Besonderer Fokus

wurde dabei auf die interaktive Visualisierung der umfangreichen Baugeschichte des knapp 500 Jahre alten Museumsgebäudes gelegt, um die komplexe Entwicklungsgeschichte des Gebäudes verstehen und nachvollziehen zu können.

Das Projekt wurde in drei wesentliche Entwicklungsphasen eingeteilt (siehe Abb. 2):

- A) Konzepterstellung und Test aller geplanten Funktionen
- B) Modellierung und Texturierung des Museums, seiner Inhalte und der Modelle zur Visualisierung der Baugeschichte
- C) Einbindung in ein Programm und Erstellung der Interaktionsmöglichkeiten

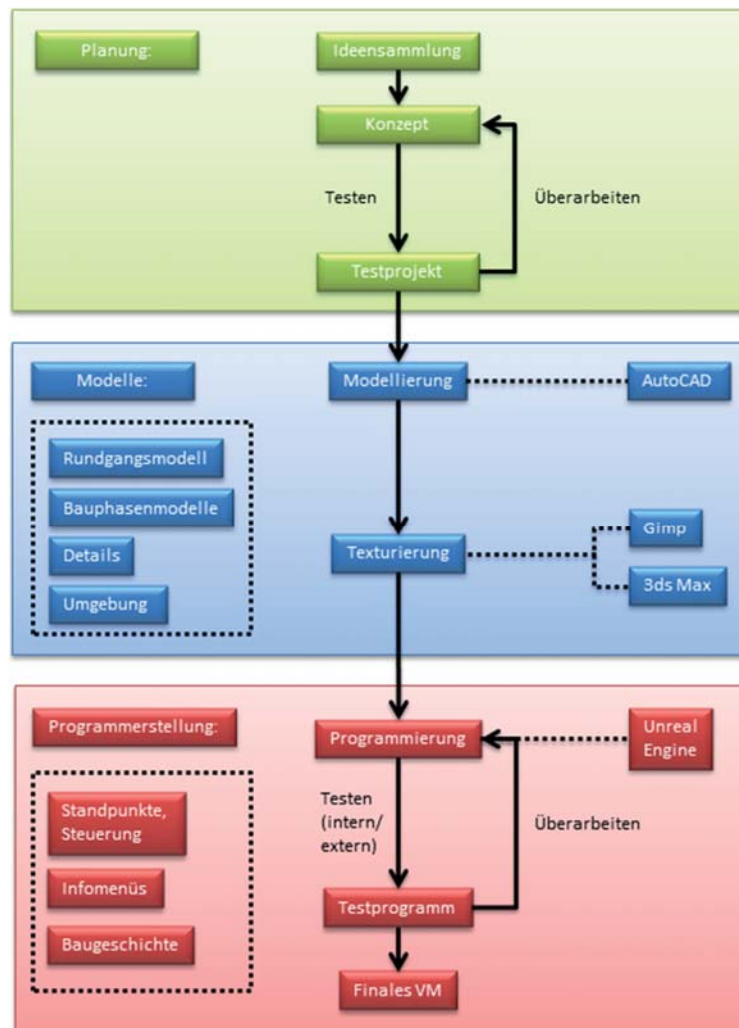


Abb. 2: Arbeitsschema zur Erstellung des VMs – A: Konzept (oben), B: Modellierung (Mitte) und C: Programmierung (unten)

Die bisher existierenden VMs bestehen zum einen aus digitalisierten Sammlungen und zum anderen größtenteils aus auf Panoramen basierenden virtuellen Touren mit wenig oder keinen Interaktionsmöglichkeiten. Um das Potential der digitalen Visualisierung weiter auszuschöpfen, wurde der Serious Games-Ansatz gewählt, d.h. die Verwendung aus Spielen bekannter Elemente

wie die weitgehend freie Bewegung und die Möglichkeit, seine Umgebung oder interessante Objekte näher betrachten zu können.

Zunächst wurden verschiedene Möglichkeiten der Steuerung und der Darstellung der Informationen in einem groben Versuchsmodell getestet und die geplanten Funktionen entsprechend umgesetzt, um später eine optimale Benutzung des Programms zu gewährleisten. Anschließend erfolgte eine 3D-CAD-Modellierung des Museumsgebäudes auf Basis einer bereits im Jahre 2011 von der HafenCity Universität Hamburg (HCU) durchgeführten Vermessung durch Photogrammetrie und terrestrischem Laserscanning (KERSTEN et al. 2013). Um den Innenraum lebendig zu gestalten, wurden außerdem die wichtigsten Ausstellungsstücke, Informationstafeln und Einrichtungsgegenstände modelliert und an ihre entsprechenden Stellen im Gebäude platziert. Zusätzlich wurden die sechs verschiedenen historischen Bauzustände des Alt-Segeberger Bürgerhauses in Zusammenarbeit mit dem Historiker NILS HINRICHSSEN (Leiter des Museums Alt-Segeberger Bürgerhaus) modelliert. Das Aussehen des Gebäudes ist vor allem für die frühen Bauphasen nur zu Teilen historisch-wissenschaftlich gesichert. Es erfolgte daher eine enge Abstimmung mit Bauforschungsergebnissen, die in den letzten Jahren zur Geschichte des Gebäudes durchgeführt wurden. Beispielsweise erfolgte eine dendrochronologische Analyse einzelner Bauhölzer, aus denen das Alter verschiedener Teile des Gebäudes bestimmt und den entsprechenden Bauphasen zugeordnet werden konnte (REIMERS & HINRICHSSEN 2015).

Zur Datenreduktion wurden die sechs Bauzustände zusammen modelliert, d.h. Objekte, welche in mehreren Bauphasen vorkommen, wurden nur einmal erstellt und in einer Datenbank hinterlegt, um vom Programm später entsprechend ihrer Zugehörigkeit aufgerufen werden zu können. Abb. 3 zeigt die ersten vier Bauzustände und ihre markantesten Änderungen aus demselben Blickwinkel. Schließlich wurde auch die Umgebung des Gebäudes auf Basis von terrestrischen Fotos und Informationen aus Google Earth rekonstruiert und als Stilmittel eingesetzt, um eine Einbettung des historischen Bauwerks in der Stadtumgebung zu gewährleisten.

Die Texturierung des Modells erfolgte mit der Software 3ds Max, in dem hauptsächlich auf vor Ort erstellte Fotos als Texturen zurückgegriffen wurde, aber auch bearbeitete, online frei verfügbare Texturen integriert wurden. Zur Verbesserung der Tiefenwirkung und Darstellung von Details wurden auch Bump- und Alphatexturen eingesetzt. Insgesamt wurden 239 Texturen für die Visualisierung verwendet.



Abb. 3: Verschiedene Bauzustände des Bürgerhauses. V.l.n.r.: 1541 - 1585, 1585 - ca. 1805, ca. 1805 - 1814 und 1814 - ca. 1890.

5 Implementierung des Virtuellen Museums

Der Hauptteil der Arbeit befasste sich mit der Programmierung der Benutzerbewegung im Museum, mit den Informationsabfragen und mit den entsprechenden Animationen in der Game-Engine. Die intuitive Bedienung des Programms war eine wesentliche Voraussetzung, damit auch PC-unerfahrenen Nutzern ein leichter Einstieg und Nutzung des VMs ermöglicht wird (Abb. 4). Die Steuerung erfolgt ausschließlich über die Maus und basiert auf vielen im Gebäude verteilten Standpunkten, welche direkt über eine Karte angesteuert oder als Kamerapfad durch die 3D-Umgebung angefahren werden können (Abb. 5). An diesen Standpunkten kann sich der Nutzer frei umsehen. Die Visualisierung der Baugeschichte wurde mit einem "Modell im Modell" realisiert (Abb. 6 und Abb. 7 rechts). Innerhalb des virtuellen Rundgangs kann ein separates Modell geöffnet werden, welches jeden der Bauzustände auf Wunsch anzeigt und animiert. Der Nutzer kann das Gebäude dabei durch virtuelle Drehung mit der Maus von allen Seiten betrachten. Zusätzlich lässt sich per Mausklick das Dach entfernen, die einzelnen Stockwerke werden per Animation auseinander gefahren und das Aussehen des Gebäudeinneren im jeweiligen Bauzustand wird angezeigt (Abb. 6). Jeder Raum kann nun einzeln ausgewählt werden, um sich Informationen über die Entwicklung der Gebäudenutzung anzeigen zu lassen. Außerdem ist es möglich, für jede Bauphase einen animierten Übergang zur jeweils nachfolgenden zu starten, welcher die baulichen Veränderungen zum nächsten Zustand zeigt und beschreibt. Diese Animationen machten eine Unterteilung aller 3D-Modelle in 387 Kleinobjekte erforderlich, um für jede Animation präzise die zu bewegendenden Objekte ansteuern zu können.



Abb. 4: Einstieg in die Anwendung des Virtuellen Museums Alt-Segeberger Bürgerhaus

Für die Ausstellungsstücke und Informationstafeln wurden Menüs erstellt, welche während des Rundgangs mithilfe von Infozeichen, die über oder neben den auszuwählenden Objekten platziert wurden, geöffnet werden können. Diese Menüs enthalten kurze Erläuterungen und

zumeist eine Abbildung, welche per Mausklick vergrößert dargestellt werden kann. Einige Informationen werden direkt über Detailkameraeinstellungen in der 3D-Umgebung übermittelt. Für solche Fälle wurden im ganzen Museum Kameras an den entsprechenden Stellen verteilt, welche ebenfalls über Infozeichen angewählt werden können.

Schließlich wurden Komfortfunktionen wie Tool-Tipps, eine Übersichtskarte und ein Hilfe-Menü erstellt. Zur Qualitätssicherung wurde das Programm von mehreren Personen mit unterschiedlicher PC-Erfahrung getestet und anschließend in Details angepasst.



Abb. 5: Einblicke in das Virtuelle Museum Alt-Segeberger Bürgerhaus – Navigationspfeile im Flur (links) und Standpunkt auf dem Dachboden.



Abb. 6: Menü für die Baugeschichte des Gebäudes mit Einblick in das Gebäudeinnere der dritten Bauphase (ca. 1805 - 1814).

6 Fazit & Ausblick

Das entwickelte Programm für den virtuellen Museumsbesuch enthält 13 im Museum verteilte Standpunkte und 52 Info-Menüs. Es hat eine Größe von 500 MByte und lässt sich als eigenständiges Programm auf Windows-Betriebssystemen ausführen. In der kommenden Museumssaison wird für dieses VM ein PC-Terminal im Alt-Segeberger Bürgerhaus aufgebaut, um es Besuchern zugänglich machen zu können. Es erlaubt – begleitend zur Ausstellung – eine multimediale Auseinandersetzung mit der Geschichte der Stadt und des Gebäudes. Es leistet somit einen wichtigen Bildungsbeitrag über die Stadtentwicklung von Bad Segeberg und über 500 Jahre Wohnkultur in Schleswig-Holstein anhand eines Gebäudes. Das komplett in 3D erstellte VM ist als informativer Bestandteil eines Museums in dieser Form in Deutschland einzigartig.

Die chronologische Visualisierung der Bauphasen hat sich außerdem als geeignet erwiesen, bisher nicht gesicherte oder nicht bekannte Abschnitte der Baugeschichte zu überprüfen bzw. zu verifizieren. Dabei konnten auch zusätzliche Erkenntnisse zur Baugeschichte gewonnen werden. Bisherige Programmvorfürungen haben gezeigt, dass besonders die Kombination aus Interaktivität und Visualisierung der Baugeschichte den besonderen Reiz des VMs ausmachen, da somit das Thema umfassend und gleichzeitig nachvollziehbar betrachtet wird. Die Animationen und die Möglichkeit, das Gebäude in seine Bestandteile zu zerlegen, regen dabei den Entdeckungsdrang des Nutzers an. Gerade diese Art der Informationsvermittlung, welche in dieser Form bisher nicht mit klassischen Ausstellungsmethoden möglich ist, hat bei den bisherigen Benutzern enormen Zuspruch erfahren. Ein virtuelles Museum setzt also idealerweise dort an, wo die verständliche und unterhaltsame Präsentation eines Themas mit einer konventionellen Ausstellung an ihre räumlichen, finanziellen oder didaktischen Grenzen stößt.

Im Labor Photogrammetrie & Laserscanning der HCU Hamburg bildet dieses Projekt den Ausgangspunkt für weitere Forschungsarbeiten im Bereich Virtueller Museen. So wurde bereits eine Version zur Benutzung des Programms mit dem Virtual Reality System HTC Vive umgesetzt (Abb. 7 links). Hierfür wurde die Standpunktsteuerung gegen eine Möglichkeit der freien Bewegung ausgetauscht. Zusätzlich können weite Strecken mithilfe einer Teleportfunktion überbrückt werden (Abb. 8). Auch die Baugeschichte lässt sich somit viel immersiver erleben, da der Besucher dank der freien Positionierung aus jedem beliebigen Blickwinkel das Modell betrachten und zuschauen kann, wie sich das Gebäude direkt vor seinen Augen und "zum Greifen nah" über die Jahrhunderte verändert. Die Objektauswahl und Menübedienung wird über einen "Laserstrahl" ermöglicht, der vom Motion Controller aus gesteuert wird (Abb. 7 Mitte).

Die aufsteigende Technologie Augmented Reality bietet ebenfalls großes Potential durch die Verknüpfung der Vorteile eines VMs mit dem realen Museumsbesuch. Sie ermöglicht außerdem mithilfe eines Smartphones oder Tablets den Museumsbesuch in situ, zum Beispiel an Standorten von (ehemaligen) historischen Stätten durch digitale Überlagerung des aktuellen Zustands mit einem historischen Bauzustand (CANCIANI et al. 2016).

Solche Anwendungen bzw. Systeme können auch außerhalb des Museumskontextes – beispielsweise für Produktvisualisierungen, auf Messen oder als Lernprogramme für Arbeitsabläufe (z.B. für die Feuerwehr) – zum Einsatz kommen.



Abb. 7: Virtual Reality als intensives Erlebnis – Der Bildschirm im Hintergrund zeigt einen Ausschnitt dessen, was der Benutzer in der VR Brille sieht (links). Menüführung mit grünem Motion Controller durch „Laserstrahl“ (Mitte) und Menüinformationen an rotem Handcontroller (rechts).

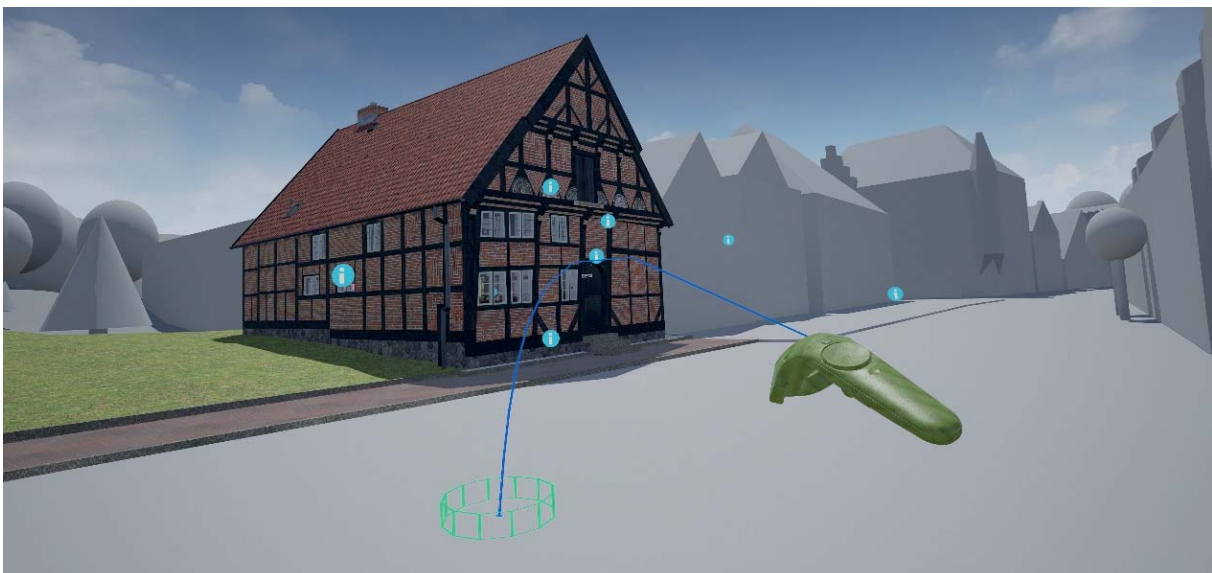


Abb. 8: Blick in die virtuelle Welt – Freie Bewegung durch Teleportation in der virtuellen Realität mit grünem Handcontroller.

7 Literaturverzeichnis

BAUER, T., 2001: Museen und Internet. *Museologie Online* **3**, 112-161.

CANCIANI, M., CONIGLIARO, E., DEL GRASSO, M., PAPALINI, P. & SACCONEA, M., 2016: 3D Survey and Augmented Reality for Cultural Heritage. The Case Study of Aurelian Wall at Castra Praetoria in Rome. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41** (B5), 931-937.

- GÜTT, I., 2010: Smartphone-Applikationen im Museumsbereich. Bachelorarbeit, Studiengang Museumskunde, Fachbereich Gestaltung, HTW Berlin, 74 S.
- HINRICHSEN, N., 2012: Der „Segeberger Rathausirsch“ – Museumsobjekte intensiv „befragt“ erzählen Stadtgeschichte. Unveröffentlichtes Manuskript und Vortragsmanuskript, Bad Segeberg, 20 S.
- KERSTEN, T., HINRICHSEN, N., LINDSTAEDT, M., WEBER, C., SCHREYER, K. & TSCHIRSCHWITZ, F., 2013: Baugeschichtliche 3D-Dokumentation des Alt-Segeberger Bürgerhauses durch Photogrammetrie und terrestrisches Laserscanning. Denkmäler3.de 2013 – Von low-cost bis high-tech: 3D-Dokumentation in Archäologie und Denkmalpflege, Przybilla, H.-J., Kersten, T., Boochs, F. (Hrsg.), LWL Industriemuseum Zeche Zollern Dortmund, 16.-18. Oktober 2013, 30-37.
- KERSTEN, T. & LINDSTAEDT, M., 2017: Mit Kamera, Maßstab und Laptop – Best Practice zur Erstellung von 3D-Modellen für eine Museumsausstellung am Beispiel vietnamesischer Kulturobjekte. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 26, Kersten, T. (Hrsg.), 35-44.
- LAWSON, E., 2016. Game Engine Analysis. <https://www.gamesparks.com/blog/game-engine-analysis/>, letzter Zugriff 12. Januar 2017.
- MORTARA, M., CATALANO, C. E., BELLOTTI, F., FIUCCI, G., HOURY-PANCHETTI, M. & PETRIDIS, P., 2014: Learning cultural heritage by serious games. *Journal of Cultural Heritage* **15**(3), 318-325.
- O’FLANAGAN, J., 2014. Game Engine Analysis and Comparison. <https://www.gamesparks.com/blog/game-engine-analysis-and-comparison/>, letzter Zugriff 12. Januar 2017.
- REIMERS, H. & HINRICHSEN, N., 2015: Das Alt-Segeberger Bürgerhaus. Vom Stadtwohnhaus des 16. Jahrhunderts zum Stadtmuseum im 21. Jahrhundert. Lutherstadt Wittenberg, Torgau und der Hausbau im 16. Jahrhundert, *Jahrbuch für Hausforschung*, Band **62**, 341-355.
- REUSSNER, E., 2007: Wissensvermittlung im Museum – ein überholtes Konzept? *Kultur und Management im Dialog* **5**, 20-23.
- SAMIDA, S., 2002: Überlegungen zu Begriff und Funktion des „virtuellen Museums“: Das archäologische Museum im Internet. *Museologie Online* **4**, 1-58.
- SAEFTEL, F., 1966: Das Alt-Segeberger Bürgerhaus: Baugeschichte und Wiederinstandsetzung. *Die Heimat* Nr. 10, 73. Jahrgang, Wachholtz Verlag, Neumünster.
- TSCHIRSCHWITZ, F., KERSTEN, T. & ZOBEL, K., 2014: Interaktive Visualisierung und Steuerung von 3D-Modellen und Punktwolken mit Low-Cost Systemen. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 23, Seyfert, E., Gülch, E., Heipke, C., Schiewe, J., Sester, M. (Hrsg.), Beitrag 173.

3D Rekonstruktion von Biberdämmen zur Simulation des Wasserrückhaltes an Fließgewässern

LISA KNOPP¹, KONRAD EDER¹, LUDWIG HOEGNER¹ & UWE STILLA¹

Zusammenfassung: Der nachfolgende Beitrag untersucht die Leistungsfähigkeit der Photogrammetrie zur 3D-Rekonstruktion von Biberdämmen und beschreibt einen geeigneten Ablauf für die Datenerfassung und Prozessierung. Als Datengrundlage dienen Multikopter-Aufnahmen sowie terrestrische Aufnahmen des Biberreviers Glonn/Erdweg, die im Frühjahr 2016 erstellt wurden. Im Zuge der photogrammetrischen Auswertung werden zunächst Kameraorientierung, Punktverdichtung, Orthophoto und ein Digitales Geländemodell berechnet. Daraus werden schließlich projektrelevante Informationen, wie die Höhen der Wasserniveaus, die Lage der Uferlinien sowie die Dammförmigkeit extrahiert. Die erreichten Genauigkeiten werden anhand verschiedener Qualitätskriterien beurteilt und liegen im Bereich weniger Zentimeter.

1 Einleitung

Hochwasserkatastrophen wie 2016 in Niederbayern zeigen, wie aktuell das Thema Hochwasserschutz ist und wie wichtig es bleibt, sich mit Möglichkeiten des Wasserrückhaltes auseinanderzusetzen. Genau dies ist das Ziel des 2015 gestarteten Projekts „Prozessbasierte Modellierung natürlicher sowie dezentraler Hochwasserrückhaltmaßnahmen zur Analyse der ereignis- und gebietsabhängigen Wirksamkeit“ des Bayerischen Staatsministeriums für Umwelt und Verbraucherschutz (DISSE 2016). Dabei werden nicht nur klassische Maßnahmen betrachtet, wie beispielsweise Rückhaltebecken, sondern auch „natürliche“ Retentionsmaßnahmen, wie Auenvegetation und Biberdämme.

Die Hochschule Weihenstephan-Triesdorf startete 2015 ein gesondertes Projekt, in dem insbesondere die Wirkung des Bibers auf den natürlichen Wasserrückhalt untersucht wird (ZAHNER 2015). Hierfür erfolgt zunächst eine Systematisierung und Typisierung der Biberdämme, bei der Geometrie (Lage, Dammhöhe, Durchmesser, etc.) und weitere technische Daten, wie z.B. Holzart und Durchlässigkeit, erfasst werden. Ausgehend von dieser Datengrundlage wird anschließend das Hochwasserverhalten von Biberdämmen hinsichtlich Stabilität, Dauerhaftigkeit und Statik untersucht. Hochwassersimulationen zeigen schließlich die Auswirkungen auf die Gewässerstruktur und Gewässerentwicklung. Um die Forschungen auf einer breiten Datengrundlage fundieren zu können, sollen mindestens 30 Biberdämme in Bayern erfasst werden.

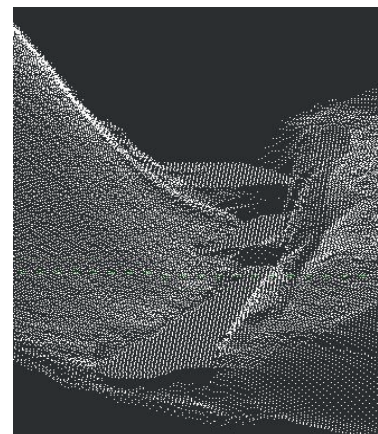


Abb. 1: Punktwolke des Staubereichs (überhöhte Ansicht)

¹ Technische Universität München, Photogrammetrie und Fernerkundung, Arcisstr. 21, 80333 München, E-Mail: [lisa.knopp, ludwig.hoegner, stilla]@tum.de, konrad.eder@bv.tum.de

Die bisherigen Verfahren zur Erfassung von Biberdämmen beruhen hauptsächlich auf Messungen mit Maßband und Nivellement (RUOFF 2015). Da dies jedoch sehr zeitaufwendig ist und lediglich einzelne Querprofile liefert, soll mit Hilfe der Photogrammetrie eine neue Verarbeitungskette für eine großflächige und detaillierte Erfassung entwickelt werden, um so geometrische Daten der Staubebereiche, inklusive deren Dämme und Uferbereiche, zu gewinnen. Diese Daten können dann in Form einer 3D-Punktwolke in entsprechende hydrologische Auswertesoftware integriert werden und dienen gemeinsam mit weiteren Umgebungsparametern als Grundlage für Hochwassersimulationen.

2 Datenerfassung

Als Untersuchungsgebiet für die Erarbeitung einer geeigneten Methodik wurde das Biberrevier Glonn / Erdweg im Nordwesten Münchens ausgewählt. Das Revier ist durch seine kaskadenartige Struktur sehr repräsentativ: Getrennt durch vier große Dämme erstreckt sich das Gebiet entlang von ca. 350 m über fünf Haupt-Wasserniveaus. Über ein Betonrohr mit 40 cm Durchmesser kann das Wasser am Ende des Gebiets abfließen. Abb. 2 zeigt das Projektgebiet. Dargestellt sind die vier großen Dämme (blaue und weiße Linien) inkl. Hauptdamm, die unterschiedlichen Wasserniveaus (mit 1 bis 5 nummeriert) sowie der Abfluss.

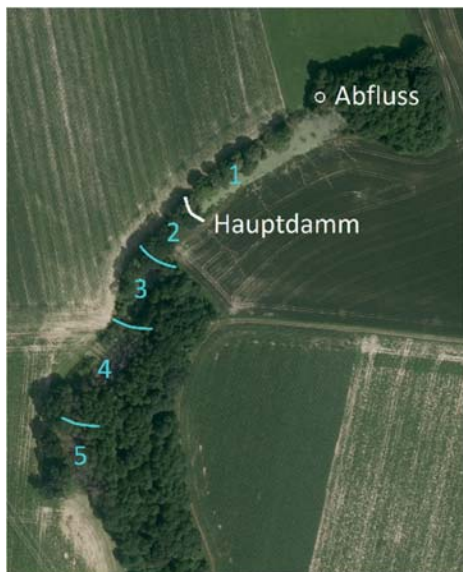


Abb. 2: Übersicht des Biberreviers Glonn / Erdweg

Da viele Biberreviere nur schwer zugänglich sind, wird zunächst versucht das Gebiet ausschließlich durch Luftbildaufnahmen zu erfassen. Für den späteren Übergang in Landeskoordinaten werden entlang des Projektgebiets Passpunkte angeordnet und mit einem GPS-Empfänger eingemessen. Zur Vermarkung werden vor allem rund zugeschnittene Styroporplatten verwendet, da diese leicht zu transportieren sind und ohne weiteres Zubehör mit einem Plastikdübel am Boden befestigt werden können. Zusätzlich wird in jedem Niveau eine Wassermarkierung für die Höhenbestimmung ausgelegt. Realisiert werden diese durch auf dem Wasser schwimmende Styroporplatten, die im Uferbereich mit einer Leine befestigt werden.

Im vorgestellten Gebiet erfolgte die Aufnahme mit Hilfe eines Multikopters und wurde in zwei unterschiedlichen Flughöhen in klassischer Blockgeometrie durchgeführt.

Die erste Aufnahme erfolgte Ende Februar bei einer

Flughöhe von ca. 50 m und starker Bewölkung. Die spätere Auswertung zeigte jedoch, dass sich die Bilder aus dieser Flughöhe auf Grund des hohen Waldanteils nicht eignen, um ausreichend eindeutige, homologe Punkte zu identifizieren. Die Befliegung wurde daher im Mai bei einer Flughöhe von ca. 100 m wiederholt. Zudem wurden einige Aufnahmen der Dämme vom Boden aus durchgeführt.

3 Photogrammetrische Auswertung

3.1 Erzeugung photogrammetrischer Produkte

Das Datenmaterial wurde zunächst ausführlich analysiert und in unterschiedlicher Software verarbeitet. Dabei wurde sowohl die manuelle, also auch die semi-automatische und automatische Verknüpfungspunktsuche angewandt. Als optimal erwies sich ein semi-automatisches Vorgehen, das im Folgenden näher erläutert wird.

Es werden zunächst drei Passpunkte manuell in den Bildern eingemessen, so dass eine vorläufige Kameraorientierung berechnet und die Szene als 3D-Punktwolke dargestellt werden kann (siehe Abb. 3). Die Markierung der übrigen Passpunkte erfolgt anschließend direkt in der Punktwolke. Sobald ein Passpunkt in je drei Bildern markiert ist, kann die Messung in den übrigen Bildern automatisch über Farbkorrelation und die bisher geschätzte Position erfolgen.

Werden nicht ausreichend Verknüpfungspunkte zwischen den Bildern gefunden, werden die Bilder gegebenenfalls in mehreren Blöcken prozessiert. Um dies zu vermeiden, müssen manuelle Verknüpfungspunkte hinzugefügt werden. Hierfür eignen sich vor allem die ausgelegten Marken, aber ebenso andere markante Stellen, wie vom Biber angenagte Bäume oder Astgabelungen.

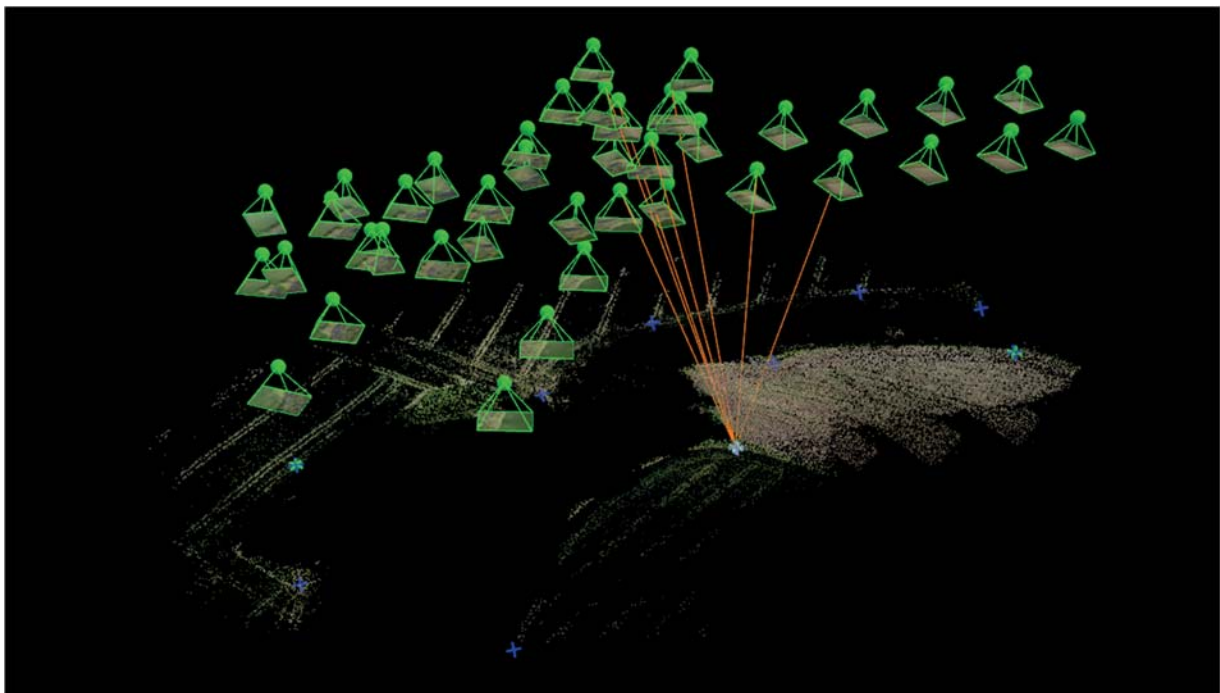


Abb. 3: 3D-Punktwolke nach Berechnung der Kameraorientierung

Nach einer automatischen Punktverdichtung werden anschließend ein Digitales Geländemodell (DGM) und ein Orthophoto generiert. Das Orthophoto wird optimiert, indem Saumlinien im Bereich der Dämme und Uferlinien vermieden und stattdessen in für die weitere Bearbeitung irrelevante Bereiche gelegt werden, wie beispielsweise Wald oder die Gewässermitteln. Das DGM wird zu einem späteren Zeitpunkt zusätzlich editiert (siehe Kap. 3.3).

3.2 Extraktion projektrelevanter Informationen

Ausgehend von den bisherigen Produkten können nun verschiedene für die o. g. Projekte relevante Informationen über das Biberrevier gewonnen werden. Hierzu zählen die Höhen der Wasserniveaus, die Uferlinien sowie die Dammform.

Zunächst erfolgt die Bestimmung der Wasserhöhen. Diese dienen bei Hochwassersimulationen zur Überprüfung des hydrologischen Modells: Die nach der Simulation erreichten Wasserhöhen des Staubereichs können mit den Gemessenen verglichen werden und erlauben so Rückschlüsse darauf, ob das Modell die Realität ausreichend widerspiegelt. Für die Messung der Höhengniveaus werden zunächst Referenzpunkte an der Wasseroberfläche ausgewählt und in den Bildern eingemessen. Hierzu dienen die bei der Befliegung ausgelegten Wassermarken, ebenso wie Wasserränder an Baumstämmen und sonstige aus dem Wasser auftauchende Gegenstände (hier z.B. ein Gartenstuhl). Sind alle Referenzpunkte gemessen, wird die Punktwolke optimiert, so dass Koordinaten für die neuen Referenzpunkte berechnet werden. Um eine ausreichende Genauigkeit zu erhalten, werden pro Wasserniveau mehrere Referenzpunkte gemessen und deren Höhenwerte anschließend gemittelt.

Ebenfalls von Interesse für die o. g. Projekte sind die Uferlinien, da diese die Staubereiche abgrenzen. Diese werden auf Grundlage des zuvor berechneten Orthophoto digitalisiert und als Vektordatei abgespeichert. Abb. 4 zeigt das Orthophoto mit den eingezeichneten Uferlinien (weiß) sowie den geschätzten Dammkronen (rot).

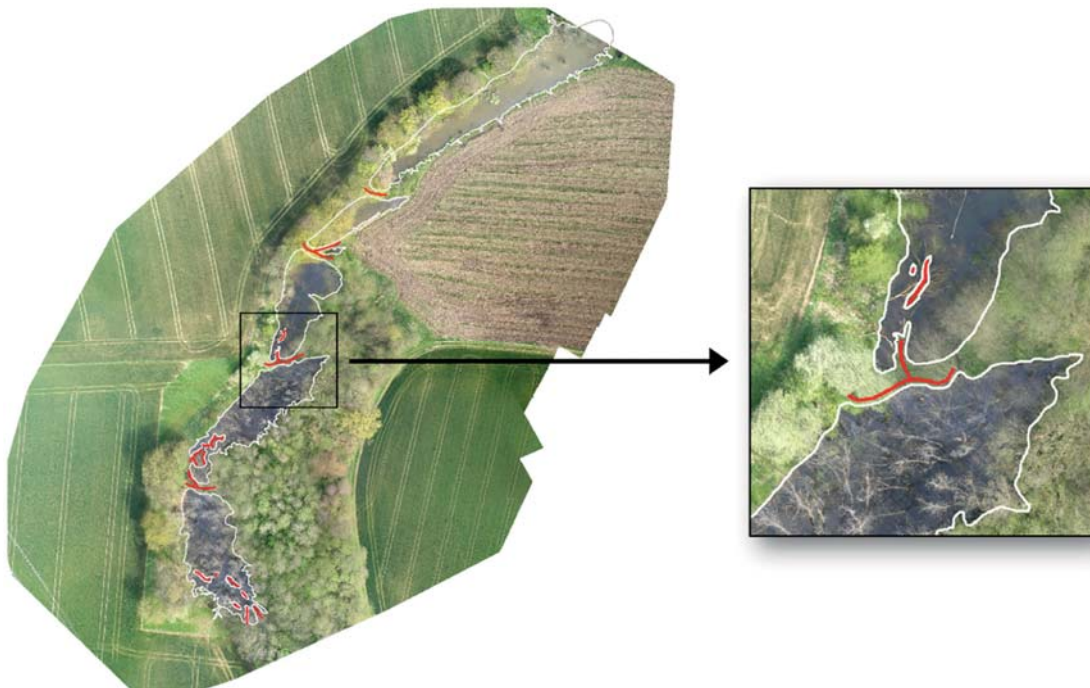


Abb. 4: Orthophoto mit digitalisierten Uferlinien und geschätzten Dammkronen

Von besonderer Bedeutung ist zudem die Form des Dammes, da sie entscheidend für das Verhalten bei Hochwasser und somit relevant für die entsprechenden Simulationen ist. Aus dem Orthophoto kann bereits die ungefähre Lage des Dammes abgegrenzt werden, Höheninformationen können

bisher jedoch nicht extrahiert werden. Grund hierfür ist, dass im Zuge der photogrammetrischen Auswertung im Bereich der Dämme nicht ausreichend Verknüpfungspunkte gefunden wurden und somit keine Punktwolke der Dämme vorliegt. Entsprechend stellen auch die daraus generierten DGM-Werte den Damm nicht ausreichend dar. Die Bilder werden daher gemeinsam mit den bereits berechneten Parametern in eine andere Auswertesoftware importiert. Dort können die Bilder im Stereomodus übereinandergelegt und so der Höhenverlauf der Dammkrone abgegriffen werden. Leider zeigt sich bei Kontrollmessungen, dass es in den Höhenwerten zu großen Abweichungen in dieser Software kommt. Eine absolute Bestimmung der Dammform ist somit nicht möglich. Auch relative Messungen sind nicht möglich, da die Messmarken im vorliegenden Maßstab nicht genau genug aufgesetzt werden können.

Die Verwendung von zusätzlichen terrestrischen Aufnahmen ist folglich zwingend notwendig. Versuche mit den Bodenaufnahmen zeigen, dass dadurch nach der Prozessierung sogar einzelne Hölzer deutlich zu erkennen sind und diese als 3D-Objekte modelliert werden können. Uferlinien und Dammkronen können in diesem Fall direkt in der Punktwolke digitalisiert werden (siehe orange Polylinie in Abb. 5), und enthalten somit nicht nur Lage-, sondern auch Höheninformation. Ebenso können terrestrische Aufnahmen für die Bestimmung der Wasserhöhe von Vorteil sein. Durch die bessere Auflösung sind zusätzliche Strukturen, wie beispielsweise kleine Blätter, auf der Wasserfläche zu erkennen. Diese wiederum eignen sich ausgezeichnet als Referenzpunkte für die Bestimmung der Wasserhöhe.

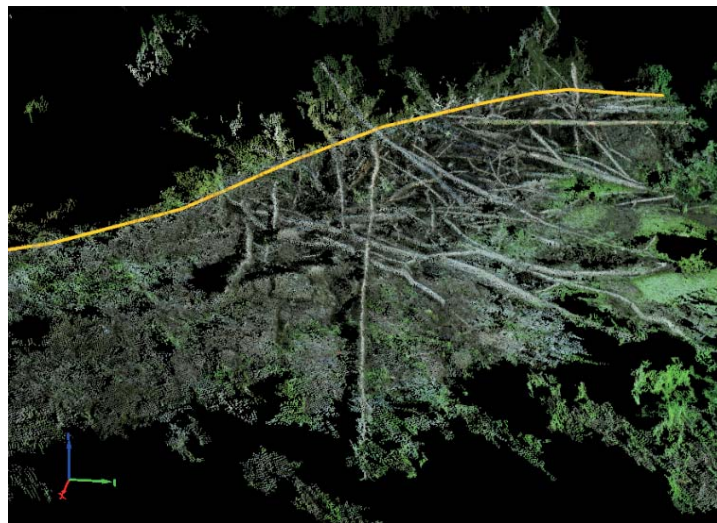


Abb. 5: Texturierte Punktwolke eines Biberdamms. Mit einer 3D Polylinie (orange) wird die Dammkrone digitalisiert.

3.3 Visualisierung

Bevor das DGM visualisiert werden kann, müssen noch einige manuelle Korrekturen durchgeführt werden: Wegen des starken Bewuchses im Bereich der Uferzone enthält das DGM Bäume und Sträucher, die entfernt werden müssen. Damit keine Löcher entstehen, wird der Datensatz mit dem amtlichen DGM (Landesamt für Digitalisierung, Breitband und Vermessung, München) kombiniert und der Verlauf anschließend geglättet. Zudem kann die Wasseroberfläche auf Grund mangelnder Verknüpfungspunkte nicht korrekt dargestellt werden. Die Wasserniveaus werden daher unter Zuhilfenahme der extrahierten Uferlinien und berechneten Höhen auf das korrekte Niveau gesetzt. Abb. 1 zeigt einen Ausschnitt der finalen Punktwolke, in dem die untersten drei Wasserniveaus (1-3) in überhöhter Ansicht dargestellt sind. Eine ebenfalls sehr anschauliche Darstellung liefert die Berechnung eines Reliefs.

4 Diskussion der Ergebnisse

Die erstellten Produkte dienen als Grundlage für die Simulation von Hochwasserszenarien in den Bibergebieten. Die Daten sollten daher eine Genauigkeit im Zentimeterbereich erreichen.

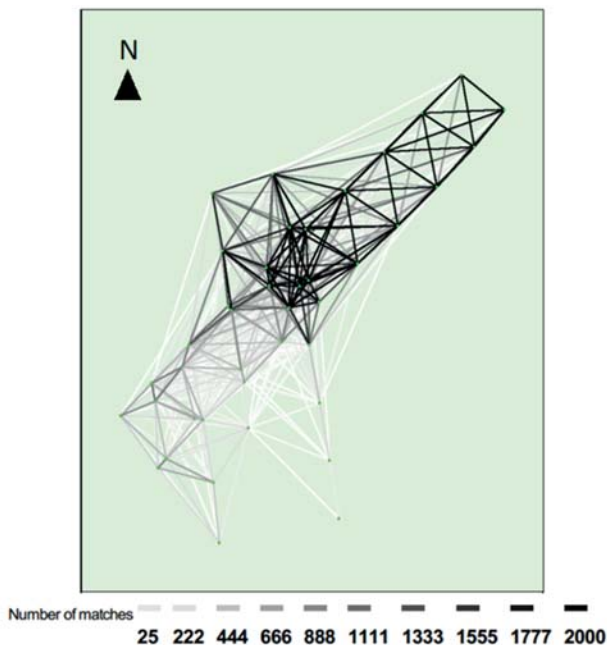


Abb. 6: Stärke des Matchings. Je dunkler die Linien sind, desto besser ist das Matching

Grundlegend hierfür sind geeignete Bilder, die prägnante Strukturen zeigen und ausreichend Überlappung aufweisen. Nur so können genügend 3D-Verknüpfungspunkte durch Matching generiert werden. Für eine hohe Qualität sollten mindestens 1.000 Verknüpfungspunkte pro Bild generiert werden (PIX4D 2016). Abb. 6 zeigt eine Aufsicht der Kamerapositionen und visualisiert die Qualität der Verknüpfung zwischen den Bildern. Je dunkler die Linie zwischen zwei Kamerapositionen, desto mehr verknüpfte 2D-Merkmalpunkte liegen vor und desto besser ist das Matching. Es ist deutlich zu erkennen, dass die Verknüpfung in den nördlichen, nicht bewachsenen Bereichen des Reviers sehr gute Qualität hat, hingegen in den höher liegenden, stark bewachsenen Bereichen eine sehr geringe Qualität.

Als Maß für die Genauigkeit der Georeferenzierung dient das quadratische Mittel (RMS). Im vorliegenden Projekt erreicht dies 1,1 cm und ist damit in etwa halb so groß wie die durchschnittliche Bodenpixelgröße von 2,13 cm. Für spätere Hochwassersimulationen ist insbesondere die relative Genauigkeit der Punktwolke relevant. Als Faustformel kann dabei von der ein- bis dreifachen Bodenpixelgröße ausgegangen werden (PIX4D 2016). Dies entspricht einer relativen Genauigkeit von ca. 2-6 cm. Für bessere Genauigkeiten ist eine Verkleinerung der Bodenpixelgröße nötig. Terrestrische Aufnahmen sind in diesem Fall vorteilhaft.

Die Auflösung von DGM und Orthophoto sind ebenfalls abhängig von der Bodenpixelgröße und den zugrunde liegenden 3D-Punkten. Ausgehend vom o.g. RMS liefert eine Digitalisierung im Orthophoto somit eine Lagegenauigkeit von mehreren Zentimetern bis Dezimetern. Durch eine 3D-Digitalisierung innerhalb der Punktwolke kann das Ergebnis gegenüber dem Orthophoto deutlich verbessert werden. Die Qualität des DGMs wird zudem maßgeblich dadurch beeinflusst, dass Bruchkanten fehlen und in den Uferbereichen kaum Stützpunkte vorliegen.

Die Bestimmung der Absoluthöhen der Wasserniveaus stellt den kritischsten Aspekt bei der Erstellung der finalen Punktwolke dar. In den freiliegenden Staubereichen 1 bis 3 (siehe Abb. 2) liegt die Genauigkeit der gemessenen Höhen bei Werten von 2 bis 7 cm bei Niveauunterschieden von ca. 0,5 – 1 m. In den stark bewaldeten Bereichen 4 und 5 erreicht jedoch die Standardabweichung teilweise bis zu 40 cm. Die Ergebnisse der Niveaus 4 und 5 sind für hydrologische Simulationen daher zu ungenau.

Auf Grund des natürlichen Umfelds der Biberdämme kommt es bei der Erfassung vermehrt zu Schwierigkeiten, die gemeinsam mit Lösungsansätzen im Folgenden erläutert werden.

Die größte Herausforderung stellt die Verdeckung durch Vegetation dar. Die Dämme liegen meist in bewaldeten Gebieten, sodass diese auf den Aufnahmen häufig durch Bäume verdeckt sind. Zudem sind die Dämme selbst oft mit hohen Gräsern und Sträuchern bewachsen, was die Bestimmung der Dammform erschwert. Die Aufnahmen sollten daher in den Wintermonaten November bis März erfolgen, in denen der Bewuchs noch nicht sehr ausgeprägt ist. Zusätzlich zur Blockbefliegung sollten Bodenaufnahmen aus unterschiedlichen Perspektiven vorgenommen werden, um eine freie Sicht auf Dämme und Uferlinien zu ermöglichen.

Das Biberrevier besteht in der Regel nicht nur aus einem Hauptdamm, sondern aus einer Gruppe mehrerer Dämme. Das Gebiet Glonn / Erdweg beinhaltet beispielsweise neben den vier großen Dämmen noch zahlreiche kleinere Dämme, die sich jeweils nur über einen Teil der Gewässerbreite erstrecken. Solche Strukturen erschweren die Bestimmung der Wasserhöhen und führen zu großen Standardabweichungen. Ausgehend von Luftbildern ist es zudem äußerst schwer zu beurteilen, ob es sich bereits um einen Damm handelt oder nur um belanglos im Wasser liegendes Gehölz bzw. eine weiter ins Wasser reichende Grasfläche. Eine genaue Dokumentation während der Aufnahmen vor Ort ist daher vorteilhaft.

Die Verdeckung durch Vegetation spielt auch bei der Wahl geeigneter Passpunkte eine Rolle. Diese sollten in der Nähe des zu untersuchenden Gebiets gut verteilt sein, damit es nicht zu Verdrehungen des Systems kommt. Zudem dürfen die Passpunkte auch nicht auf einer Raumgerade liegen, da dadurch das Normalgleichungssystem bei der Bündelblockausgleichung singulär wird. Insbesondere bei langgezogenen Bibergebieten sollte daher darauf geachtet werden, die Passpunkte auf beiden Uferseiten zu verteilen, um so schlecht konditionierte Gleichungssysteme zu vermeiden.

Die Messung der Referenzpunkte auf der Wasseroberfläche stellt grundsätzlich eine Herausforderung dar. Die Wassermarken bieten die einfachste Möglichkeit die Wasserhöhe zu messen. Jedoch sollte bei der Befliegung darauf geachtet werden, dass die Marken während der Befliegung nicht ihre Lage ändern (z.B. auf Grund von Wind). Weitere Referenzpunkte zu finden ist oftmals sehr schwierig. Bäume und andere Objekte, die aus dem Wasser ragen, scheinen gut geeignet. Jedoch ist oft schwer festzustellen, an welcher Stelle das Objekt das Wasser verlässt. Dunkle Flecken im Wasser sind ebenfalls nur unter Vorbehalt zu verwenden, denn es könnte sich dabei um Blätter oder andere Partikel auf der Wasseroberfläche handeln, aber auch um unter Wasser liegende Steine. Bei der Aufnahme des Gebiets sollte daher eine möglichst große Anzahl an Wassermarken eingesetzt werden, insbesondere in stark bewaldeten Bereichen.

Auch Überlappung und Flughöhe spielen eine wichtige Rolle, da diese entscheidend für die Durchführung der Bündelblockausgleichung sind. Die Distanz zur Geländeoberfläche sollte so gewählt werden, dass der Anteil prägnanter Strukturen groß genug ist. Besonders in Waldgebieten sind bei Aufnahmen aus zu geringem Abstand Strukturen oftmals nur unzureichend erkennbar. Bei einer Aufnahme aus größerer Distanz werden diese deutlicher, was die Verknüpfungspunktsuche erleichtert, jedoch gleichzeitig die Auflösung verschlechtert. Für bewaldete Gebiete empfehlen sich eine Überlappung von 60 % und eine Flughöhe zwischen 70 und 100 m bei Verwendung der Kamera Nex-7 mit 19 mm Festbrennweite. Bei offenem Gelände kann die Flughöhe reduziert werden.

Zuletzt darf bei allen Ergebnissen und Genauigkeiten nicht vergessen werden, dass sich die Biberdämme und Wasserniveaus durch die natürliche Aktivität des Bibers und Niederschläge in ständigem Wandeln befinden. Alle Aufnahmen sollten daher zur gleichen Zeit vorgenommen werden, um so eine in sich konsistente Abbildung zu gewährleisten.

5 Fazit & Ausblick

Zusammenfassend eignen sich die Methoden der Photogrammetrie gut für die 3D-Rekonstruktion von Biberdämmen. Aufnahmen aus der Luft und vom Boden erlauben es, auch nicht begehbare Dämme genau zu vermessen. Der zeitliche Aufwand ist deutlich geringer und die erzielte Punktdichte wesentlich höher als bei den bisherigen Verfahren mit Maßbandmessungen.

Aus den Messungen können neben der 3D-Punktwolke weitere wichtige Produkte extrahiert werden. Dazu zählen das Orthophoto, das als Grundlage für die Klassifizierung der Lebensraumtypen dienen kann, die Uferlinien, die Höhenniveaus in den einsehbaren Bereichen sowie die Lageabgrenzungen der Dämme. Die Genauigkeit der finalen Punktwolke liegt im Bereich weniger Zentimeter, was ein wichtiges Ergebnis für die hydrologischen Simulationen ist. Mit der bisherigen Arbeit wurde lediglich der Bereich des Dammes über Wasser rekonstruiert. Doch auch für die Bereiche unter Wasser kann die Photogrammetrie Lösungsansätze liefern: In Bereichen klaren und stillen Wassers können mittels Photo-Bathymetrie die Objektkoordinaten der Gewässersohle berechnet werden (MAAS 2015). Zusätzlich zu den passiven Kameraaufnahmen können auch aktive Lasersysteme dazu genutzt werden, um bei klarem Wasser Bodenbereiche sichtbar zu machen. Einen Mehrgewinn könnten zudem Thermalaufnahmen liefern, da mit ihrer Hilfe Temperaturunterschiede zwischen Totholz und fließendem Wasser visualisiert und somit Durchflüsse innerhalb des Biberdamms detektiert werden könnten.

6 Literaturverzeichnis

- DISSE, M., 2016: ProNaHo: Prozessbasierte Modellierung Natürlicher sowie Dezentraler Hochwasserrückhaltemaßnahmen zur Analyse der ereignis- und gebietsabhängigen Wirksamkeit (ProNaHo). <https://www.hydrologie.bgu.tum.de/index.php?id=176&L=0>, letzter Zugriff am 15.06.2016)
- MAAS, H.-G., 2015: On the Accuracy Potential in Underwater/Multimedia Photogrammetry. *Sensors* 2015, **15**(8), 18140-18152
- PIX4D, 2016: Pix4Dmapper Software Manual. <https://support.pix4d.com/hc/en-us#gsc.tab=0>, letzter Zugriff am 07.07.2016.
- RUOFF, K., 2015: Auswirkungen von Biberdämmen auf die Strukturvielfalt und Durchgängigkeit von Fließgewässern im östlichen Main-Kinzig-Kreis. Masterarbeit, Johann Wolfgang Goethe-Universität, Institut für Ökologie, Evolution und Diversität, Fachbereich Biowissenschaften, Frankfurt am Main.
- ZAHNER, V., 2015: Die Wirkung des Europäischen Bibers (*Castor fiber*) auf den natürlichen Wasserrückhalt an ausgewählten Fließgewässern Bayerns. Projektbeschreibung Hochschule Weihenstephan-Triesdorf, Fakultät für Wald und Forstwirtschaft, Weihenstephan.

BRDF Analysis of Different Land Cover Types Using Aerial and Satellite Imagery

ALEXANDER KÖHN¹

Abstract: Bi-directional reflectance distribution function (BRDF) effects are an important factor influencing the reflectance of the land surface, especially at large off-nadir angles. This work demonstrates the information gain from using multi-angular images from a low-cost aerial camera system as well as means to radiometrically calibrate these images facing unknown camera parameters. Land surface BRDF are estimated by inverting four semi-empirical kernel-driven models. A novel model is proposed utilising the same data sources as the other methods but featuring a more explicit geometrical modelling approach using the OpenGL library. Although some issues with the current model remain, an improved version may be a useful addition to conventional methods.

1 Introduction

Surface anisotropy describes the property of virtually all natural surfaces to reflect light differently strong in different directions. While spectral reflection variance is often due to chemical composition of the target surface, angular variance tends to primarily depend on textural properties (BARNSELY et al. 2004). In a simplified sense, rough surfaces will result in a more Lambertian scattering behaviour while very smooth surfaces (e.g. polished metal, water) have a strong directional component which is hereafter referred to as hotspot. Surface anisotropy is formalised as the bi-directional reflectance distribution function (BRDF) where the reflectance f_r for a given reflection angle (φ_r, θ_r) and an illumination angle (φ_i, θ_i) at wavelength λ is expressed by

$$f_r(\varphi_i, \varphi_r, \theta_i, \theta_r, \lambda) = \frac{dL_r(\varphi_i, \varphi_r, \theta_i, \theta_r, \lambda)}{dE_i(\varphi_i, \theta_i, \lambda)} = \frac{dL_r(\varphi_i, \varphi_r, \theta_i, \theta_r, \lambda)}{L_i(\varphi_i, \theta_i, \lambda) \cos \theta_i d\omega} [sr^{-1}],$$

i.e. the ratio of reflected radiance dL_r to irradiance dE_i of an infinitesimal area dA (NICODEMUS et al. 1977). BRDF is an important factor that needs to be corrected when working with wide-angle sensors such as AVHRR, MODIS and MERIS where anisotropic effects become especially pronounced in the off-nadir field-of-view (WANNER et al. 1995). Nonetheless, the additional information content from multi-angular imagery can be applied to classify land surfaces that are poorly distinguishable in the spectral domain (BARNSELY et al. 1997). BRDF is still widely ignored in land surface studies on a local scale, certainly in parts due to the prevalence of sensors featuring a narrow field of view with little or no tilting capabilities such as LANDSAT or Sentinel where BRDF effects are not particularly apparent (KOUKAL & SCHNEIDER 2010). More recently, commercial high-resolution platforms with considerable tilting capabilities emerged to be more responsive to customer needs (e.g. Ikonos, WorldView, Pléiades and SkySat). This

¹ Saarland University, Department of Computer Science, Campus, D-66123 Saarbrücken, E-Mail: s8alkoeh@stud.uni-saarland.de

trend demands a renewed level of awareness for angular effects but also provides new opportunities to investigate land surface BRDF.

The equation above implies that an infinite number of observations over the whole viewing hemisphere is required to accurately describe the BRDF of a given land surface. In reality, especially when using airborne or spaceborne sensors for measuring BRDF, interpolating over the data points is impractical due to the sparseness of observations and the tendency to angular bias of satellite orbits and flight paths (KOUKAL & ATZBERGER 2012). For that reason, special models have been developed to incorporate some assumptions on the target surface to more accurately predict the BRDF when facing sparse angular sampling. Hence, physical models generally achieve the most accurate descriptions of radiative transfer but are rather impractical for model inversion (CHEN et al. 2000; KOUKAL & ATZBERGER 2012). On the other hand, kernel-driven models attempt to describe reflectance as a linear combination of kernels, that are functions that simulate a specific scattering behaviour that contributes to the final BRDF. Inverting these models is relatively straightforward since they can be considered as an overdetermined system of linear equations which can be solved by minimising a linear least squares problem (LEWIS 1995; WANNER et al. 1995).

Hence, BRDF studies may benefit from a similar viewing geometry as common for photogrammetric purposes. Interestingly, there is little research on how such camera systems can be applied to sample BRDF which in turn may be used to normalise images on a physical basis. Moreover, height information from photogrammetrically derived digital surface models (DSM) can be used to explicitly formulate kernel functions for geometrical shadowing and diffuse shading which could improve BRDF model accuracy compared to conventional models while relying on the same input data.

2 Methodology

In this study, data from the aerial DLR 3K camera system and the Pléiades constellation are examined. The airborne sensor is used to answer the question whether a low-cost camera system can be used for BRDF studies. The goal here is to radiometrically correct the images and, by reproducing the methodology of BARNSELY et al. (1997), to generate multi-angular composite images displaying the information content. The second part features a comparison between four BRDF models including a novel one using high-resolution Pléiades data.

2.1 Datasets

The first dataset consists of aerial imagery acquired by the DLR 3K imaging system during a flight campaign on September 17th 2015 over the Kranzberger Forst near Freising, Germany. With one nadir-looking and two laterally mounted oblique-looking off-the-shelf DSLR cameras, it has a field-of-view of 110° in across-track direction and 31° in along-track direction (KURZ et al. 2012). The solar azimuth angles were almost perpendicular to the flight paths which rendered the oblique-looking images especially valuable for BRDF studies. This is due to the tendency of the solar plane to contain most information about surface anisotropy (KIMES 1983). However, since the respective flight paths over the area of interest were several tens of minutes apart, only

successive images from a single camera were used to prevent changes in the solar position to obscure the results.

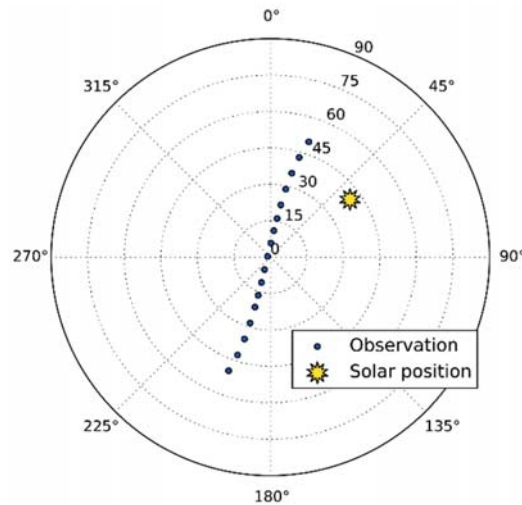


Fig. 1: Polar plot of the Pléiades viewing and illumination geometry

The second dataset comprises a set of 18 Pléiades images acquired over Melbourne, Australia on February 25th 2012 during a single overpass using the persistent surveillance imaging mode. This implies a good angular coverage approximately along the solar plane between zenith angles of -50.03° and $+49.8^\circ$ (Fig. 1). For all practical purposes, the solar incidence angle can be considered constant due to an overall time interval of 03:30 min between the first and the last image. Two small subsets with relatively homogeneous land cover were manually selected, one representing a suburban residential area while the other one was taken from a salt marsh.

2.2 Radiometric and geometric calibration

3K imagery is not designed for radiometric studies and is, by default, recorded in the JPEG format which implies a compression from 16-bit RAW to 8-bit grey values and, more important, that both camera response curve and white balance factor are not preserved which are needed for radiometric calibration. The whole conversion procedure can be expressed by

$$L = \frac{g(B)}{w_i e M}, \text{ with } e = \frac{\pi s t}{k^2},$$

where L is the scene radiance, g the inverted camera response curve, B the raw digital number, w_i the white-balance factor for band i , e the effective exposure, M the vignetting coefficient, s the sensor sensitivity (ISO), t the shutter speed and k the aperture value (D'ANGELO 2007).

The camera response curve was estimated by plotting grey values of simultaneously taken RAW and JPEG photos against each other and taking the most abundant RAW value for each 8-bit JPEG value which was then saved to a look-up table (Fig. 2, right). Since weight balance appears to be least pronounced in the green band (D'ANGELO 2007), only this band was transformed. Finally, the vignetting coefficient of the camera lens was determined by finding the weights to a radial polynomial (D'ANGELO 2007) using the physically-based photo stitching software hugin on panorama photos of one of the 3K cameras. This resulted in individual vignetting curves for all f-numbers between 3.2 and 9.9 (Fig. 2, left).

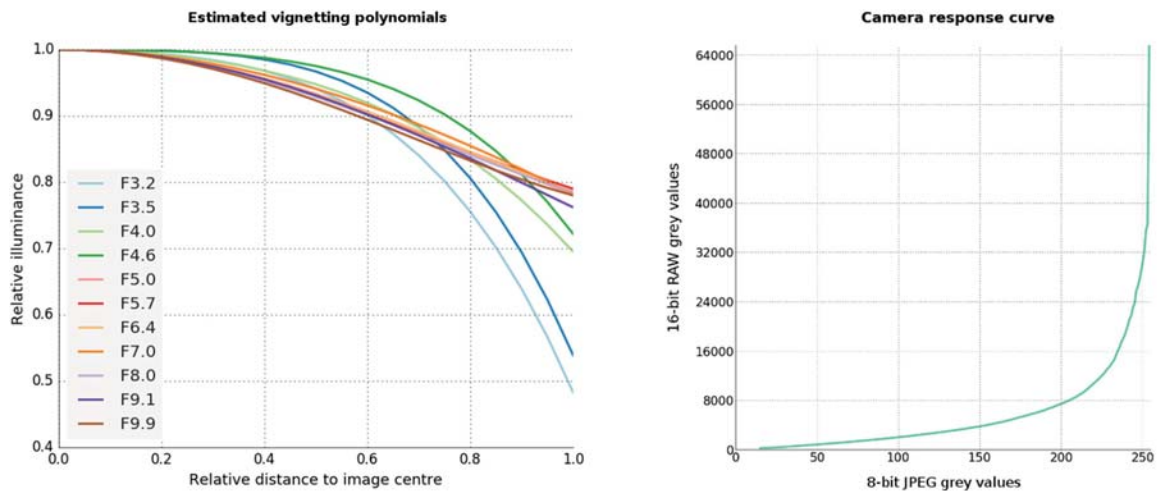


Fig. 2: Radiometric calibration results: estimated vignetting curves (left) and camera response curve (right)

Like BARNSELY et al. (1997), 3K imagery was not atmospherically corrected in these experiments which, however, will be necessary in future studies if those data are used for BRDF modelling. Geometric calibration was performed using highly accurate rational polynomial coefficients (RPC) and a 90 m SRTM DEM. Note that perfect orthorectification is not desired here since perspectival deformations are important contributors to the BRDF of a land surface. This will, on the other hand, compromise the visualisation of the anisotropy of obstructions, i.e. canopies and roof tops. The Pléiades data used ATCOR 2 as atmospheric correction procedure (RICHTER & SCHLÄPFER 2005) since this method assumes a planar surface which is critical for later modelling. Unfortunately, the Pléiades DIMAP V2 format has a default band combination Red-Green-Blue-NIR which led to a mix-up in the correction procedure and ultimately implies that the modelling results have no real meaning (R. RICHTER, personal communication, 2016). Nonetheless, this does not compromise the validity of the assessment of different BRDF models since all images are still comparable in relative terms. The atmospherically corrected images were then projected onto a 90 m SRTM DEM in case of the three conventional methods. However, the design of the fourth model required an orthorectified input which is why a photogrammetric DSM was used instead.

2.3 BRDF model descriptions

All of the implemented models are kernel-driven. The first one is the Walthall model which is a simple empirical description of the BRDF of canopies and one of the first kernel-driven models (WANNER et al. 1995). By contrast, the semi-empirical Roujean model formulates a volumetric scattering kernel, dubbed RossThick, a geometric kernel to simulate macroscopic shadowing and an isotropic kernel for Lambertian reflectance (WANNER et al. 1995). The third tested model is the RossThick-LiSparse (RTLs) model which adapts the general design of the Roujean model but exchanges the geometric Roujean kernel by the LiSparse kernel which in turn is based on the Li-Strahler model of mutual shadowing (WANNER et al. 1995). Both geometric kernels are analytical descriptions of the shadowing effects of randomly scattered primitives on a horizontal plane.

Since BRDF partly depends on the macroscopic geometry of the scene, utilising height information allows to explicitly model shadows and diffuse shading instead of relying on abstract geometrical models as used in the Roujean and RTLS models. The proposed modification of the RTLS model therefore utilises simple Lambertian flat shading and a fast shadow computation technique called shadow mapping while adopting the volumetric RossThick kernel. The rendering has been implemented using OpenGL 4.5.

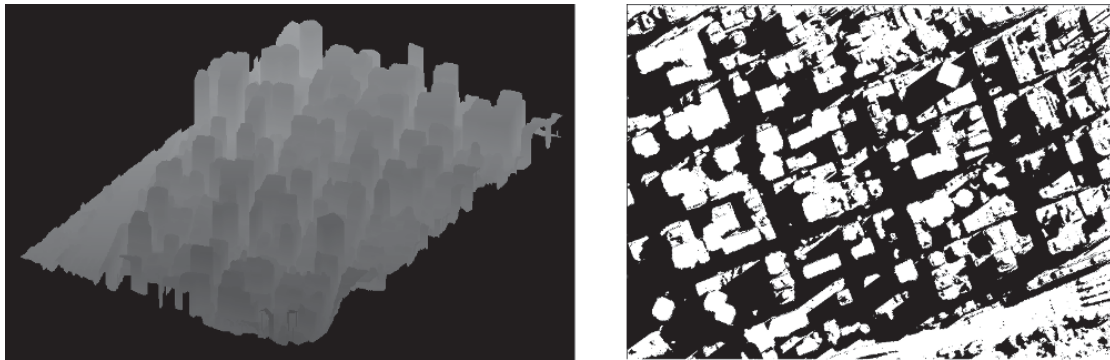


Fig. 3: The principle of shadow mapping: depth map (left) and resulting binary shadow map in nadir view (right).

Shadow mapping (WILLIAMS 1978) is an algorithm where for each pixel a depth value with respect to the light source is computed and saved to a so-called depth map (Fig. 3, left). When transforming the vertices of the geometry from world to light source coordinates, the depth map is queried to decide whether a pixel is visible from the light source. A negative result of this test implies that the pixel must be shadowed. This results in a binary map of hard shadows (Fig. 3, right). Even though, minor modifications would allow to render soft shadows as well, this is unnecessary as the model requires only an averaged value over the larger area of interest.

The original isotropic kernel as used in the Roujean and RTLS model assumes a horizontal plane that diffusely reflects incident light. Since it is invariant of the viewing direction, the kernel function is usually set to unity. This, however, changes if a DSM is used to compute the diffuse reflection for a given illumination angle. An important design decision whose implications will be discussed later was to omit occluded areas for a given viewing direction on the orthorectified image which was also realised by shadow mapping. Since shadows are areas not visible from the position of the sun, one can equally simulate occluded areas by the same algorithm by treating the simulated position of the sensor as light source.

A per-pixel BRDF computation is not feasible at this resolution since this would require a perfect image registration and an even better resolved DSM. Instead, reflectance values over a larger area were averaged which also forgives unrealistic model assumptions in flat shading and hard shadows due to shadow mapping.

2.4 Statistical model assessment

Thorough BRDF model assessment requires substantial effort as simultaneous ground measurements at the time the sensor captures the scene would have to be taken (LUCHT et al. 2000). This, of course, exceeds the capacities of most studies which is why less involving statistical methods are chosen. However, splitting the data into a training, test and verification set

is equally unfavourable due to the sparseness of the data one usually deals with. This makes a representative model assessment rather difficult. Nonetheless, one can assess the goodness of the fit using the coefficient of determination (R^2) or the symmetric mean absolute percentage error (sMAPE). To generate comparable results to KOUKAL & SCHNEIDER (2010), a similar assessment methodology was chosen by computing the sMAPE. A good fit, however, does not necessarily indicate good model performance but could rather imply overfitting to the data, so these results should be handled with care.

3 Results and Discussion

An exemplary result of the angular composites is shown in Fig. 4. The cropland is hardly distinguishable in each of the three images on the left which are selected from five successive images. The angular composite clearly distinguishes between the two crop types and emphasises irregularities. By contrast, forest canopies (upper left corner) and more mature crops exhibit a rather noise-like pattern in the composite due to perspectival deformations.

The angular composites show that the radiometric calibration appears to be very successful in recovering reflectance values and that simple aerial images can be used in radiometric studies. Other results have emphasised crop stains which could prove useful in applications like precision farming, damage assessment or aerial archaeology.

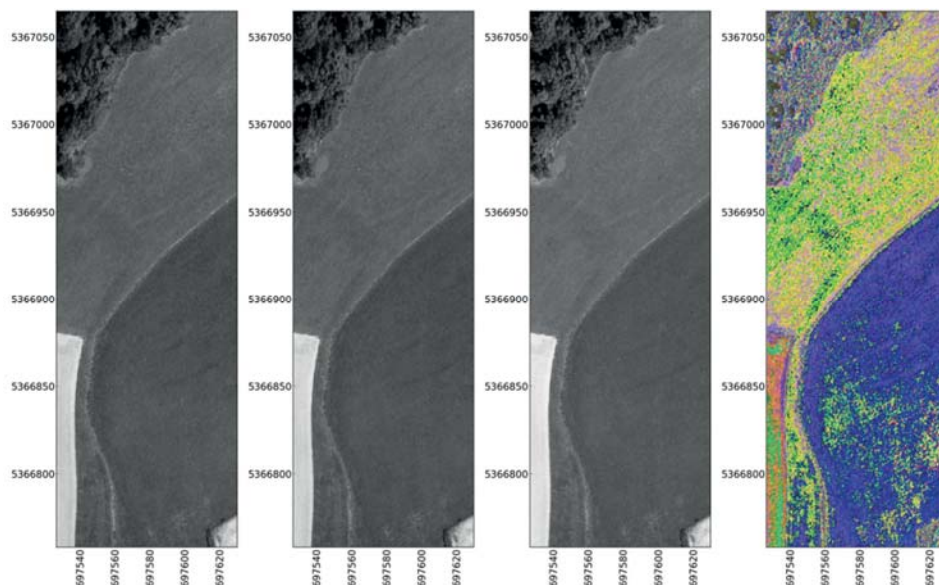


Fig. 4: Three left-looking 3K images (1-3) and the resulting composite image (far right)

A selected BRDF model result of averaged grey values over a residential area is shown in Fig. 5. The three conventional methods performed rather similarly in all selected areas of interest with the RTLS model featuring the most pronounced hotspot which is an observation consistent with KOUKAL & SCHNEIDER (2010). On the other hand, the proposed model produced several severe discrepancies when compared to the other models (Fig. 5, upper row) while agreeing in other cases (Fig. 5, lower row).

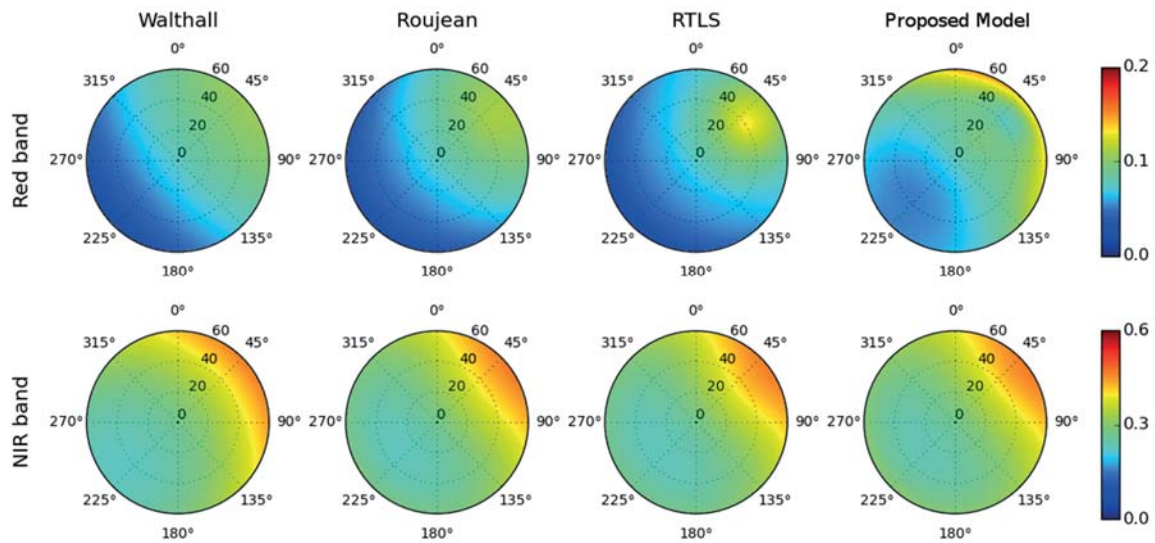


Fig. 5: Simulated BRDF for a suburban residential area in Melbourne, Australia.

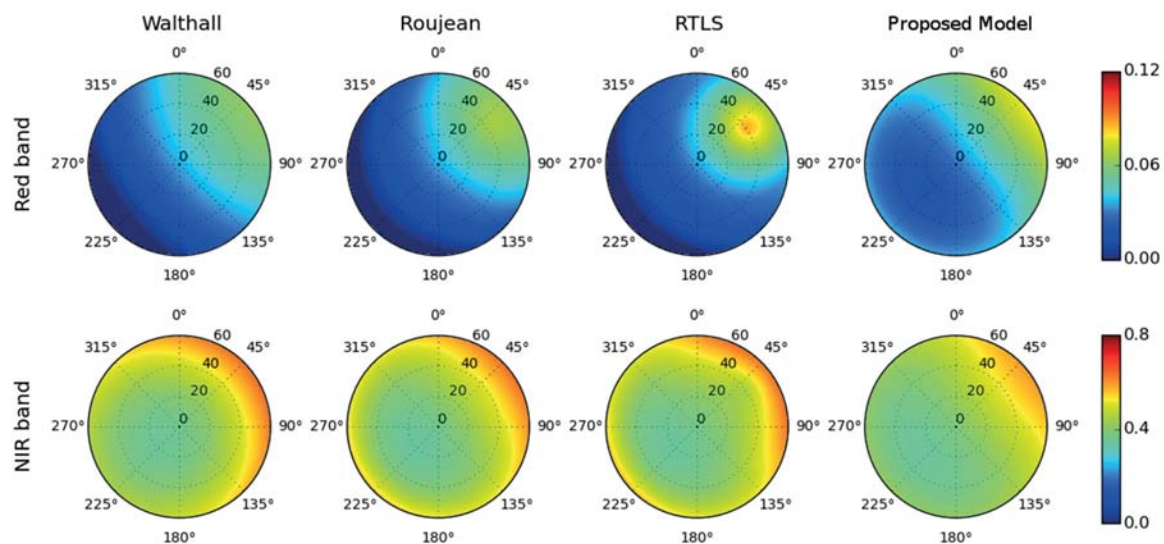


Fig. 6: Simulated BRDF for a salt marsh in Melbourne, Australia.

The results of the salt marsh subset (Fig. 6) highlight the same observations; while the simulated NIR band in the proposed model roughly agrees with the other models, the red band fails to simulate a pronounced hot spot and closer resembles the Walthall model than RTLS. The parameter weights are given in Tab. 1 and 2, respectively. In Tab. 1, the proposed model features a proportionally larger dependence on the geometric and isotropic kernels with respect to the other models.

sMAPE results are visualised in Fig. 7. The statistical results agree with the qualitative assessment with RTLS achieving the smallest error while the proposed model produced the overall highest error over all tested datasets. This was especially apparent in the salt marsh scene (Fig. 7, right).

The disagreement is probably due to masking out occluded areas in the orthorectified image which causes a model bias towards elevated areas at increasing viewing zenith angles. Since vegetation reflects less than the underlying soil in the visible domain (KIMES 1983), the bias towards canopy reflectance is most pronounced in off-nadir directions. Conversely, vegetation is much brighter in the near-infrared domain than its background (KIMES 1983) which also explains the opposite trend found in the respective simulated band.

Tab. 1: Simulated BRDF model weights for the residential area.

Band	Walthall			Roujean		
	k_1	k_2	k_3	k_{iso}	k_{vol}	k_{geo}
Blue	-0.008623	0.039471	0.075886	0.136766	0.059169	0.046739
Green	0.016004	0.057929	0.118468	0.188311	0.169166	0.031151
Red	0.025829	0.058531	0.134191	0.201849	0.199392	0.018721
NIR	0.066232	0.114765	0.275328	0.400207	0.435017	0.016606
Band	RTLS			Proposed Model		
	k_{iso}	k_{vol}	k_{geo}	k_{iso}	k_{vol}	k_{geo}
Blue	0.154369	0.002417	0.043449	-3.385180	0.200868	4.746887
Green	0.201605	0.127868	0.030144	-1.249163	0.224814	1.892070
Red	0.211819	0.170172	0.019618	-0.023008	0.201927	0.237596
NIR	0.409712	0.407628	0.017904	1.059879	0.385021	-1.025671

Tab. 2: Simulated BRDF model weights for the salt marsh.

Band	Walthall			Roujean		
	k_1	k_2	k_3	k_{iso}	k_{vol}	k_{geo}
Blue	-0.009328	0.035356	0.036076	0.082385	0.045860	0.044575
Green	0.030135	0.050245	0.088621	0.133115	0.190759	0.006552
Red	0.042327	0.054977	0.117857	0.165518	0.237521	-0.005591
NIR	0.164078	0.091132	0.365564	0.427633	0.663488	-0.130618
Band	RTLS			Proposed Model		
	k_{iso}	k_{vol}	k_{geo}	k_{iso}	k_{vol}	k_{geo}
Blue	0.095982	-0.001173	0.039016	0.021817	0.112419	0.015625
Green	0.134719	0.184725	0.005435	0.073197	0.197799	0.031250
Red	0.163918	0.243186	-0.004814	0.057714	0.225182	0.093750
NIR	0.387774	0.801343	-0.114339	0.220637	0.459911	0.250000

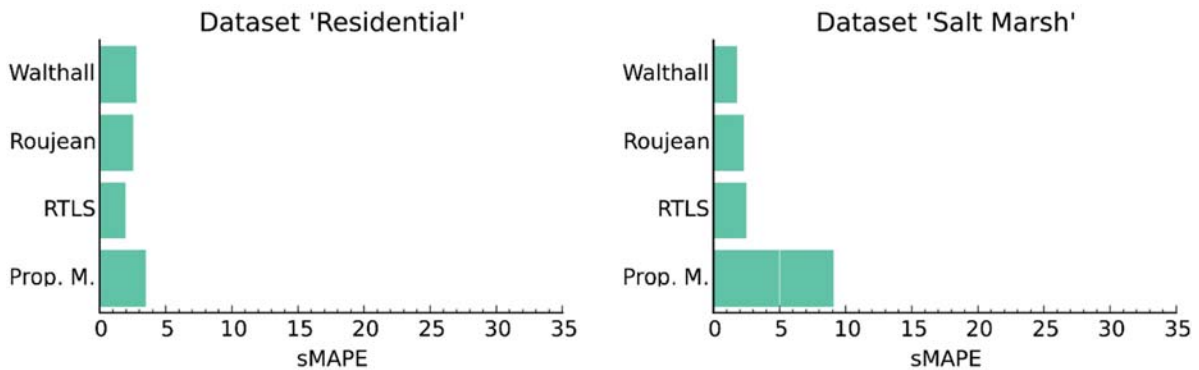


Fig. 7: sMAPE assessment results for both subsets.

These erroneous results might be prevented by using perspectival projections onto a plane instead of occlusion masks which would provide a more realistic simulation of geometric shadowing.

4 Conclusions

In this work, BRDF effects were analysed both qualitatively and quantitatively. First, aerial 3K images meant for photogrammetric purposes were radiometrically calibrated in order to create multi-angular composite images which allowed to assess the surface anisotropy of the given scene. It was possible to visualise the multi-angular information content that can enhance classification approaches solely based on spectral bands. In the second part, four different semi-empirical BRDF models were applied to 18 multi-angular Pléiades images. Among those models, a novel one is proposed based on photogrammetrically derived DSM and computer graphics rendering techniques. The model assessment generally confirms the observations by KOUKAL & SCHNEIDER (2010). Unfortunately, a design flaw in the proposed model caused a bias towards higher elevated areas which is why this model yielded the worst test results.

Future work will clearly focus on fixing these issues and develop the model further. It is planned to test the model on completely artificial datasets to allow a more conclusive accuracy assessment. Another target is to provide a quantitative accuracy assessment of the radiometric calibration procedure for aerial images.

5 Acknowledgements

The author would like to thank Thomas Krauß (German Aerospace Center, Remote Sensing Technology Institute, Oberpfaffenhofen) for his useful advice. Special thanks go to Yunshan Li for her support throughout the generation of this work and help in formatting an earlier version of this paper.

6 Literature

- BARNESLEY, M., ALLISON, D. & LEWIS, P., 1997: On the Information Content of Multiple View Angle (MVA) images. *International Journal of Remote Sensing* **18**(9), 1937-1960.
- BARNESLEY, M. J., SETTLE, J. J., CUTTER, M. A., LOBB, D. R. & TESTON, F., 2004: The CHRIS/PROBA Mission: A Low-Cost Smallsat for Hyperspectral Multiangle Observations of the Earth Surface and Atmosphere. *IEEE Transactions of Geoscience and Remote Sensing* **42**(5), 1512-1520.
- CHEN, J. M., LI, X., NILSON, T. & STRAHLER, A., 2000: Recent Advances in Geometrical Optical Modelling and its Applications. *Remote Sensing Reviews* **18**(2-4), 227-262.
- D'ANGELO, P., 2007: Radiometric Alignment and Vignetting Calibration. Proceedings of the 5th International Conference on Computer Vision Systems, March 21-24, 2007, Bielefeld University, Germany.
- KIMES, D. S., 1983: Dynamics of Directional Reflectance Factor Distributions for Vegetation Canopies. *Applied Optics* **22**(9), 1364-1372.
- KOUKAL, T. & ATZBERGER, C., 2012: Potential of Multi-Angular Data Derived from a Digital Aerial Frame Camera for Forest Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **5**(1), 30-43.
- KOUKAL, T. & SCHNEIDER, W., 2010: Analysis of BRDF Characteristics of Forest Stands with a Digital Aerial Frame Camera. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **38**(7A), 100-105.
- KURZ, F., TÜRMER, S., MEYNBERG, O., ROSENBAUM, D., RUNGE, H., REINARTZ, P. & LEITLOFF, J., 2012: Low-Cost Optical Camera Systems for Real-Time Mapping Applications. *Photogrammetrie Fernerkundung Geoinformation* **2012**(2), 159-176.
- LEWIS, P., 1995: The Utility of Kernel-Driven BRDF Models in Global BRDF and Albedo Studies. Proceedings of the International Geoscience and Remote Sensing Symposium **95**, July 10-14, Firenze, Italy, 1995, 1186-1188.
- LUCHT, W. & ROUJEAN, J. L., 2000: Considerations in the Parametric Modelling of BRDF and Albedo from Multiangular Satellite Sensor Observations. *Remote Sensing Reviews* **18**(2-4), 343-379.
- NICODEMUS, F. E., RICHMOND, F. C., HSIA, J. J., GINSBERG, I. W. & LIMPERIS, T., 1977: Geometrical Considerations and Nomenclature for Reflectance. US Department of Commerce, National Bureau of Standards, Washington DC.
- RICHTER, R. & SCHLÄPFER, D., 2005: Atmospheric/Topographic Correction for Satellite Imagery. German Aerospace Center (DLR), Weßling, Germany.
- WANNER, W., LI, X. & STRAHLER, A., 1995: On the Derivation of Kernels for Kernel-Driven Models of Bidirectional Reflectance. *Journal of Geophysical Research: Atmospheres* **100** (D10), 21077-21089.
- WILLIAMS, L., 1978: Casting Curved Shadows on Curved Surfaces. *ACM Siggraph Computer Graphics* **12**(3), 270-274.

Sensitivitätsstudie zur Bestimmung der Meeresoberflächentemperatur aus Messungen des Satelliten TET-1

NORA METTIG¹, RALF REULKE¹, JÜRGEN FISCHER², RENÉ PREUSKER² & ECKEHARD LORENZ¹

Zusammenfassung: Das Ziel dieser Arbeit ist die Entwicklung und Anwendung einer Methodik zur erstmaligen Bestimmung der Meeresoberflächentemperatur (engl. Sea Surface Temperature) aus Messungen des Satelliten TET-1 im infraroten Spektralbereich (3,8 μm und 8,8 μm). Für die Abschätzung des Einfluss der Atmosphäre wurden zunächst Strahlungstransportsimulationen durchgeführt und die stärksten Einflussfaktoren analysiert. Anschließend wurde die Sea Surface Temperature mittels einer Look-Up Table und eines Optimal Estimation Prozessors bestimmt. Die Ergebnisse können mit einer Unsicherheit von 1,7-2,0 K angegeben werden. Abschließend wurden die bestimmten SSTs mit einem Datensatz der GRSST verglichen und eine systematische Abweichung von +1-2 K festgestellt.

1 Einleitung und Problemstellung

Die Bestimmung der Meeresoberflächentemperatur oder Sea Surface Temperature (SST) gehört zu den grundlegenden Messungen in der Meteorologie, Klimatologie und Ozeanografie. Die SST enthält Informationen über Meeresströmungen und Ozeanwirbel (MILLERO 2005), der Anstieg der Meerestemperaturen ist ein wichtiges Indiz für die Klimaerwärmung (BARALE 2010) und die SST ist die Schnittstelle zwischen Ozean und Atmosphäre sowie ein Parameter für die Kopplung dieser beiden Systeme. Für globale und regionale Wettermodelle ist die SST ein entscheidender Eingangsparameter. Für die Datenassimilation der Modelle ist es wichtig, sowohl die SST als auch deren Genauigkeit möglichst präzise zu kennen.

Die Bestimmung der SST mit Hilfe von Satelliten bietet viele Vorteile: Während beispielsweise Bojen oder Schiffe nur punktuelle Messungen liefern können, ermöglichen Satelliten eine große räumliche Abdeckung eines Gebietes. Es ist möglich größere räumliche Strukturen, wie die turbulente Durchmischung zweier Ozeanströmungen, zu erkennen. Die Satellitenmessungen erzeugen zeitlich und räumlich hoch aufgelöste, globale und relativ kontinuierliche Datensätze aus denen die SST abgeleitet werden kann.

Der Nachteil der Satellitenmeteorologie besteht darin, dass relevante Größen nur indirekt bestimmt werden können. Der gesuchte Wert muss aus den gemessenen Daten abgeleitet, „invertiert“, werden. Eine weitere Schwierigkeit bei der exakten Bestimmung besteht im Einfluss der Atmosphäre. Ein Signal von der Erdoberfläche, das vom Satelliten gemessen wird, muss die Erdatmosphäre erst durchqueren bevor es von der Kamera erfasst werden kann. Den größten Einfluss auf das gemessene Signal bei etwa 10 μm nehmen die 10 km mächtige Troposphäre und

¹ Deutsches Zentrum für Luft- und Raumfahrt, Institut für Optische Sensorsysteme, Rutherfordstraße 2, D- 12489 Berlin-Adlershof, E-Mail: nora.mettig@dlr.de

² Freie Universität Berlin, Institut für Weltraumwissenschaften, Carl-Heinrich-Becker-Weg 6-10, D-12165 Berlin

der darin enthaltene Wasserdampf (EMERY et al. 1994). In den Tropen kann die Abweichung zwischen gemessener SST und berechneter SST ohne Berücksichtigung der Strahlungsprozesse in der Atmosphäre bis zu 10 K betragen (MAUL & SIDRAN 1973).

Ziel dieser Arbeit soll die erstmalige Berechnung der SST aus den Messungen des Satelliten TET-1 (REILE 2013) des DLRs, sowie die Abschätzung des Einflusses von verschiedenen atmosphärischen und technischen Parametern auf die SST, sein. Mit Hilfe von Strahlungstransportsimulationen mit MODTRAN® werden die erwarteten spektralen Strahldichten am Satellit für unterschiedliche Zustände der Atmosphäre berechnet und daraus eine Look-Up Table erstellt. Mittels dieser und einer Optimal Estimation (siehe Kapitel 2: Methodik) kann dann die SST abgeschätzt werden. Eine Besonderheit ist die hohe räumliche Auflösung von bis zu 160 m bei Infrarotaufnahmen mit TET-1, welche deutlich höher ist als jene anderer aktueller SST-Produkte. Abschließend wird die geschätzte SST mit Hilfe eines globalen täglichen GHRSSST-Produktes verifiziert.

2 Methodik

Sowohl die Sensitivitätsstudie zur Atmosphärenkorrektur, als auch die Bestimmung der SST erfolgt auf Grundlage von Strahlungstransportsimulationen mit MODTRAN®5. Es wurden zunächst unterschiedliche Zustände der Atmosphäre simuliert, um den Einfluss der Atmosphäre auf die Messungen von TET-1 abschätzen zu können. Mit diesen Daten konnte eine *Look-Up Table* (LUT) erstellt werden. Zusammen mit der LUT und einem *Optimal Estimation* Prozessor wurden schließlich die SST für verschiedene TET-1 Aufnahmen bestimmt.

2.1 Erstellung der *Look-Up Table*

Die *Look-Up Table* besteht aus einzelnen Simulationen mit MODTRAN®5 für die beiden Wellenlängenbereiche der Infrarotkameras des Satelliten TET-1. Es werden anhand der Strahlungstransportalgorithmen für jede einzelne Wellenlänge eines Spektralbereiches eine vertikale Strecke zwischen dem Boden und einer angegebenen Höhe von 500 km simuliert. Dabei ist es möglich die Spektraldichten im terrestrischen und solaren/lunaren Bereich zu berechnen und dabei Mehrfachstreuung zu berücksichtigen. Für alle betrachteten atmosphärischen Zustände wurde die Albedo der Ozeanoberfläche mit 0,02 (Ozean) angegeben. Die in der Simulation enthaltenen Aerosoleffekte sind durch maritime Aerosolprofile, welche sich an die ausgewählten Atmosphärenmodelle anpassen, beschrieben.

Neben diesen gleichbleibenden Grundeinstellungen von MODTRAN®5 wurden diverse Parameter der Simulation variiert, um die verschiedenen Zustände der Atmosphäre zu erzeugen. Der natürliche Bereich der SST wurde mit 5K-Schritten von 270 K bis 305 K abgedeckt. Wasserdampf, als ein essentieller Parameter in der Atmosphäre, soll ebenfalls innerhalb der gesamten natürlichen Variabilität verändert werden. Die Tabelle 1 zeigt neben den Werten für den Wasserdampfgehalt in der Atmosphäre sämtliche in den Simulationen variierte Parameter. Diese sind der Aufnahmeblickwinkel in °, die Sichtweite (VIS) in km, der Ozonwert in $\text{g}\cdot\text{cm}^{-2}$ und Dobson Units (DU) sowie die unterschiedlichen verwendeten Atmosphärenmodelle.

Insgesamt ergeben sich aus diesen Variationen 70 000 mit MODTRAN®5 durchgeführte Simulationen. Für die weitere Bearbeitung wurde eine siebendimensionale *Look-Up-Table* mit

den 70 000 Einträgen erstellt. Diese LUT kann zum einem für die Auswertung der Sensitivität bezüglich der variierten Parameter verwendet werden und bildet zum anderen die Grundlage der *Optimal Estimation*.

Tab. 1: Variable Parameter der Strahlungstransportsimulationen.
Es ergeben sich 70 000 durchgeführte Simulationen für verschiedene atmosphärische Zustände.

Parameter	Variationsbereich
SST	270 – 305 K
Wasserdampf	0,01 – 7 g·cm ⁻²
Ozon	100 – 500 DU
Sichtweite	1 – 70 km
Aufnahmewinkel	0 – 40 ° Abweichung von Nadir
Atmosphärenmodell	5 Modelle, Angepasst an Klimazonen (tropisch - polar)

2.2 *Optimal Estimation* zur Bestimmung der Sea Surface Temperature

Für die Berechnung der SST ist, wie bei der Bestimmung von allen meteorologischen Parametern aus gemessenen Strahlungsgrößen, eine Invertierung notwendig. Dabei wird versucht aus der gemessenen Strahldichte möglichst genau auf die SST rückzuschließen und störende Parameter, welche das Signal maßgeblich mit beeinflussen, weitgehend einzugrenzen. Mit dem Invertierungsalgorithmus (engl. *Retrieval Algorithm*) werden aus den vorher simulierten atmosphärischen Zuständen, einem gegebenen a priori Wissen zu den gesuchten Parametern, den Strahlungsmessungen sowie Angaben über Ungenauigkeiten die gesuchten Größen SST und Wasserdampfgehalt ermittelt.

Im ersten Schritt werden die einzelnen, in der LUT enthaltenen, Simulationen interpoliert und zu Funktionen, Vorwärtsmodell genannt, zusammengefasst. Mit diesem Vorwärtsmodell kann bei bekannten Atmosphärischen Größen (SST, Wasserdampf, Ozon...) die zu erwartende messbare Strahlung am Satelliten bestimmt werden. Ziel ist es jedoch mit Hilfe der *Optimal Estimation* diesen Schritt umzukehren und aus der gemessenen Strahldichte die SST zu bestimmen. Dafür sind neben den Satellitenmessungen noch Vorwärtsmodellparameter notwendig, welche zur Spezifizierung des Vorwärtsmodells dienen. Diese Parameter sind die in Kapitel 2.1 beschriebenen Variablen wie Sichtweite, Ozon oder Atmosphärenmodell.

Das Herzstück der *Optimal Estimation* bildet die Kostenfunktion $J(\vec{x})$, welche es zu minimieren gilt. Eine ausführliche Beschreibung des *Optimal Estimation* Prozesses findet sich bei RODGERS (1976) und RODGERS (2000). Die Iterationsschritte \vec{x}_{i+1} werden wiederholt bis ein Konvergenzkriterium erfüllt, oder die maximale Anzahl an Schritten erreicht ist. Die Kostenfunktion lautet:

$$J(\vec{x}) = (\vec{x} - \vec{x}_a)^T \mathbf{S}_a^{-1} (\vec{x} - \vec{x}_a) + (\vec{y} - f(\vec{x}))^T \mathbf{S}_e^{-1} (\vec{y} - f(\vec{x})) \quad (1)$$

Der erste Summand gibt die quadratische Abweichung zwischen dem gesuchten Ergebnisvektor \vec{x} und dem Vorwissen \vec{x}_a (GFS Globalmodell) gewichtet mit der Genauigkeit des Vorwissens \mathbf{S}_a^{-1} wieder. Der zweite Summand beschreibt die quadratische Abweichung von Messung \vec{y} und Vorwärtsmodell $f(\vec{x})$ gewichtet mit der Genauigkeit der Messung \mathbf{S}_e^{-1} .

Die Iterationsschritte berechnen sich mit Hilfe des Newton-Iterationsverfahrens:

$$\vec{x}_{i+1} = \vec{x}_i - (\mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_e^{-1} \mathbf{K}_i)^{-1} [\mathbf{K}_i^T \mathbf{S}_e^{-1} \cdot g(\vec{x}_i) - \mathbf{S}_a^{-1} \cdot (\vec{x}_a - \vec{x}_i)] \quad (2)$$

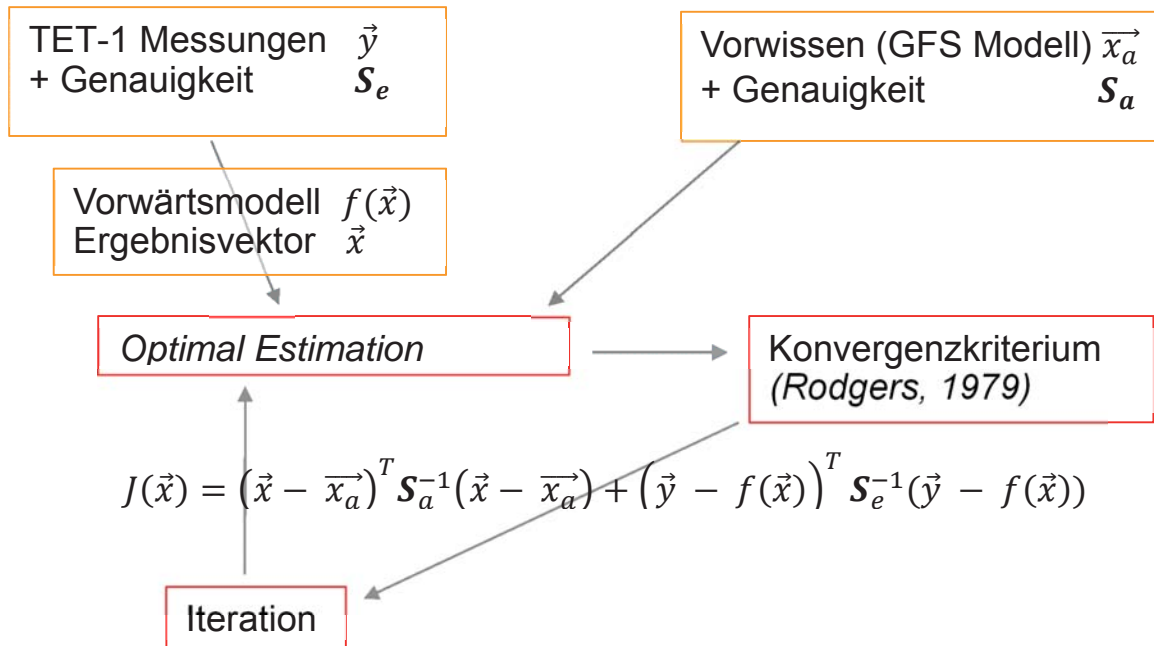


Abb. 1: Schema *Optimal Estimation* Prozess. Eingangsdatensätze sind die TET-1 Messungen der Strahldichte und Temperatur und Wasserdampf des GFS Globalmodells und deren Genauigkeiten. Basierend auf dem zuvor erstellten Vorwärtsmodell (*Look-Up Table*) wird die *Optimal Estimation* durchgeführt, indem die Kostenfunktion $J(\vec{x})$ minimiert wird. Diese Schritte werden so lang iteriert, bis das Konvergenzkriterium (nach Rodgers) erfüllt ist.

3 Daten

3.1 Messungen von TET-1

Die verwendeten Satellitendaten stammen aus dem L2 Produkt der TET-1 Messungen und beinhalten die Strahldichten der beiden infraroten Kanäle MWIR (3,4 – 4,2 μm) und LWIR (8,5 – 9,3 μm). Es wurden ausschließlich Nachtaufnahmen verwendet, um die Problematik der solaren Einstrahlung im Kanal MWIR vernachlässigen zu können. Das geodätische Referenzsystem der Daten ist WGS84 und die Messungen sind untereinander co-registriert. Die Auflösung am Boden beträgt ca. 150 m bei einer Genauigkeit der Georeferenzierung von etwa 500 m. Die Schwadbreite beträgt 178 km und die Länge der aufgenommenen Szene kann je nach ausgewählter Aufnahmezeit variieren. Die Messungen der Strahldichte liegen in der Einheit $\text{W}/(\text{sr m}^2 \mu\text{m})$ und einer Genauigkeit von 0,005 $\text{W}/(\text{sr m}^2 \mu\text{m})$ (MWIR) bzw. 0,05 $\text{W}/(\text{sr m}^2 \mu\text{m})$ (LWIR) vor.

Damit ergibt sich eine *Error-Covarianve*-Matrix \mathbf{S}_e von:

$$\mathbf{S}_e = \begin{pmatrix} 0,005 \frac{W}{sr\ m^2\ \mu m} & 0 \\ 0 & 0,05 \frac{W}{sr\ m^2\ \mu m} \end{pmatrix} \quad (3)$$

3.2 Globalmodell GFS

Das *Global Forecast System* (GFS) ist ein von den *National Centers for Environmental Prediction* (NCEP) betriebenes globales Wettervorhersagemodell, welches aus gekoppelten Atmosphären-, Ozean-, Boden-, und Seeismodellen besteht. Es beruht auf zahlreichen Atmosphären- und Bodenparametern wie Temperatur, Wind, Niederschlag, Feuchte und Ozonkonzentration. Für die Auswertung werden zwei der über zweihundert Ausgabeparameter als Vorwissen verwendet: Temperatur am Boden (*TMPS surface*) und in der Atmosphäre enthaltenes, niederschlagsfähiges Wasser (*PWAT*). Die Auflösung am Boden beträgt $0,5^\circ$ lon/lat und die Unsicherheit der beiden Parameter wird angegeben mit:

$$\mathbf{S}_a = \begin{pmatrix} 2 \left(\frac{g}{cm^2}\right)^2 & 0 \\ 0 & 10\ K^2 \end{pmatrix} \quad (4)$$

4 Ergebnisse

4.1 Sensitivitätsstudie Atmosphärenkorrektur

Ausgangspunkt für die Sensitivitätsstudie sind die mit MODTRAN® durchgeführten Strahlungstransportsimulationen. Es soll der Einfluss von verschiedenen atmosphärischen und technischen Parametern auf die Bestimmung der SST untersucht werden. Auf den Parameter Wasserdampf wird in besonderem Maße eingegangen, da dieser, wie nach Literaturrecherchen zu erwarten war, den größten Einfluss auf die SST nimmt.

Die Abbildung 2 verdeutlicht die Abhängigkeit der SST von dem in der Atmosphäre enthaltenen Wasserdampf bei verschiedenen Atmosphärenmodellen im Kanal LWIR (8,5 - 9,3 μm). Ein anderer gegenübergestellt werden die am Boden vorrausgesetzte SST und die simulierte, theoretische am Satelliten messbare, *Brightness Temperature* (BT). Die BT liegt je nach Temperaturbereich in allen Atmosphärenmodellen ca. 5 – 10 K niedriger als die SST. Dies entspricht der Größenordnung der Absorption aller atmosphärischen Parameter. Der Effekt der Wasserdampfabsorption kann aus dem Abstand der Geraden untereinander bestimmt werden. So ergibt sich beispielsweise im Mid-Latitude Summer bei 285 K (BT) eine Unsicherheit von 7 K bei unbekanntem Wasserdampfgehalt. Wie zu erwarten treten größere Abweichungen in Atmosphärenmodellen mit höherer maximaler Wasserdampfkonzentration auf. Im Modell des Mid-Latitude Winter beträgt die Unsicherheit maximal 4 K, während in der tropischen Atmosphäre bis zu 10 K Unsicherheit auftreten können. In den wärmeren Atmosphären kommt es außerdem zu einer Überschneidung der Geraden. Während oberhalb des Schnittpunktes die

Unsicherheit durch Wasserdampfabsorption entsteht, dominieren im unteren Bereich Effekte der Emission. Ist bei niedrigen Temperaturen in diesen Modellen (Tropical, Mid-Latitude und Sub-Arctic Summer) viel Wasserdampf vorhanden, emittiert dieser mehr Strahlung als er absorbiert. Allerdings ist dieses Ergebnis der Simulation nur bedingt plausibel, da bei niedrigeren Temperaturen nicht unbegrenzt Wasserdampf in der Atmosphäre vorhanden ist, sondern dieser als Niederschlag ausfällt. Am Schnittpunkt der Geraden wiegen sich Absorption und Emission auf und die Unsicherheit durch Wasserdampf verschwindet. Im spektralen Bereich des Kanals MWIR tritt ebenfalls eine Abhängigkeit der SST vom Wasserdampf auf, diese ist jedoch etwas schwächer ausgeprägt. Andere untersuchte Parameter (siehe Tabelle 1) wurden ebenso untersucht, erzeugen jedoch geringere Effekte.

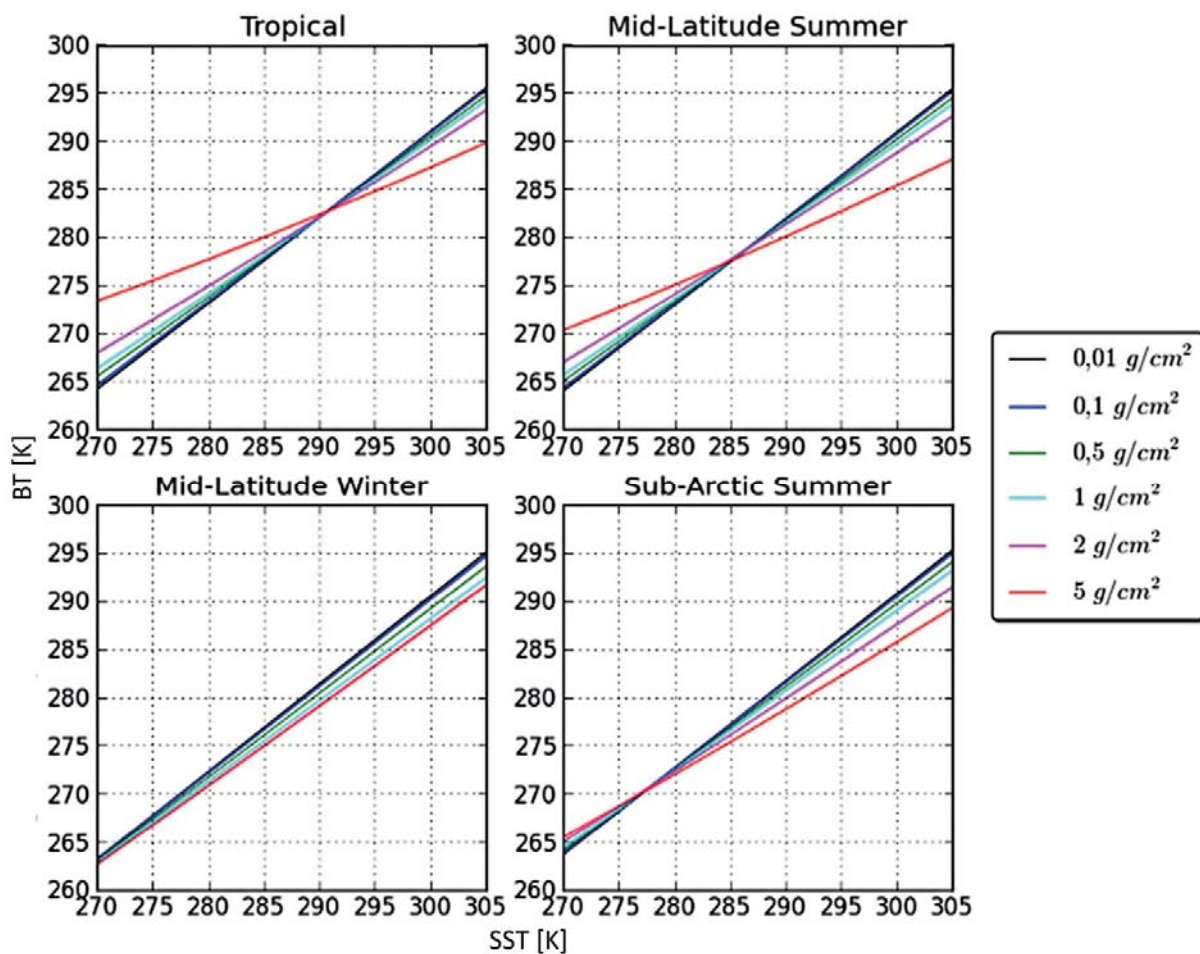


Abb. 2: Abhängigkeit der SST von dem in der Atmosphäre enthaltenen Wasserdampf in $g \cdot cm^{-2}$ bei verschiedenen Atmosphärenmodellen im Kanal LWIR (8,5 - 9,3 μm). Gegenübergestellt sind die theoretisch am Satellitenmessbare *Brightness Temperature* (BT) und die SST am Boden.

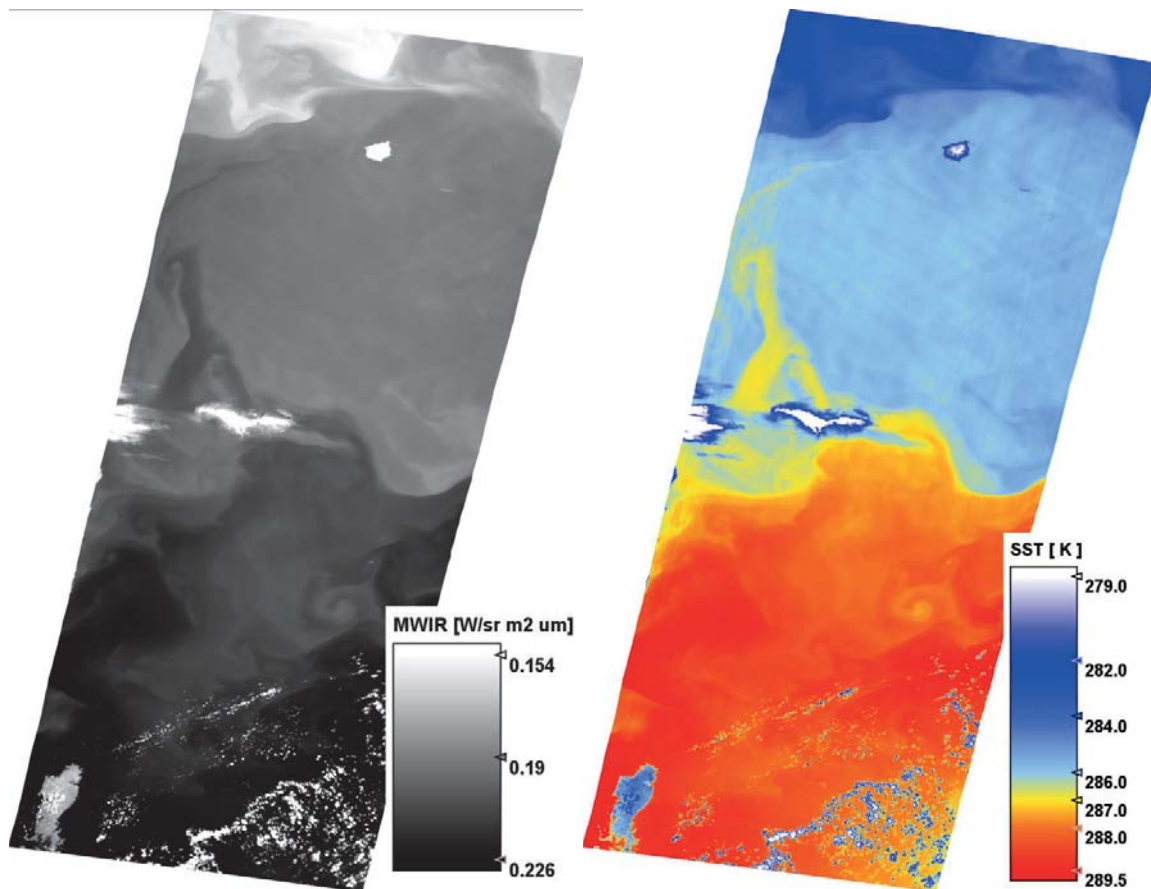


Abb. 3: Links: Strahldichten $[W \cdot (sr \cdot m^2 \cdot \mu m)^{-1}]$ des L2-Produkt von TET-1 im Kanal MWIR. Rechts: Das Ergebnis der mit *Optimal Estimation* berechnete SST [K].

4.2 Bestimmung der Sea Surface Temperature

Die Bestimmung der SST erfolgt wie in Kapitel 2.2 beschrieben mit Hilfe des *Optimal Estimation* Prozessors unter Einbeziehung der *Look-Up Table*. Es sollen hier die Ergebnisse aus den Daten der Szene vom 02.03.2016 um 16:44 Uhr UTC über dem Japanischen Meer genauer betrachtet werden (siehe Abb. 3).

Die Abbildung 3 (rechts) zeigt die Schätzung der SST mit einem relativ großen Gradienten von bis zu 7 K zwischen Nord und Süd. Im südlichen Teil des Ozeans beträgt die SST ca. 288 – 289 K während im nördlichen Teil kühlere Strömungen mit einer Temperatur von ca. 282K vorliegen. Sichtbar werden im südlichen Bereich turbulente Strukturen, welche sich auch in Richtung Norden ausdehnen. Diese Strukturen sind zum Teil kleinskalig und können aufgrund der hohen Auflösung der IR Kameras von TET-1 dargestellt werden. Auffallend kühlere Objekte

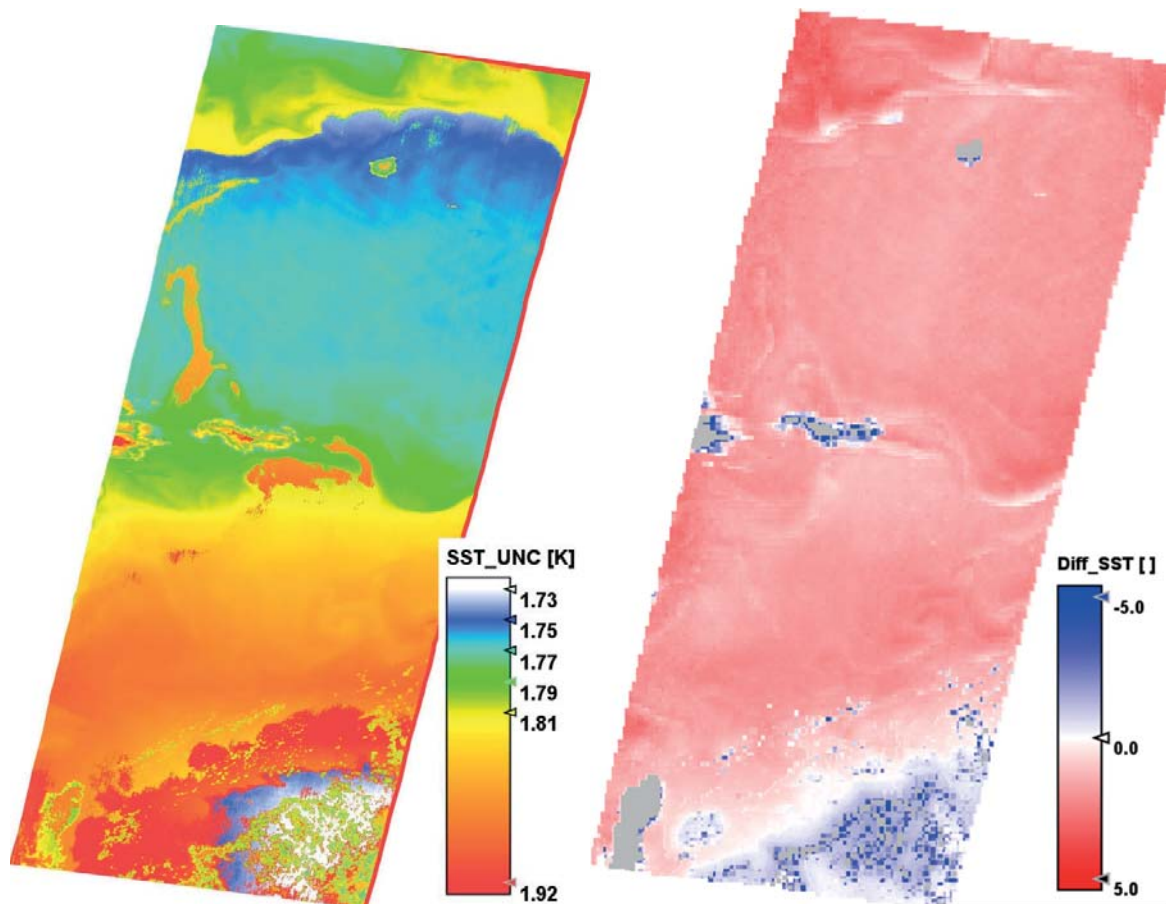


Abb. 4: Links: Genauigkeit der SST in K aus der *Retrieval-Error-Covariance*-Matrix. Rechts: Validierung der Ergebnisse. Differenz zwischen errechneter SST und Vergleichsprodukt (L-4 SST-Produkt GHRSSST)

können in dieser Szene anhand ihrer Oberflächentemperatur als Wolken oder Landflächen interpretiert werden. Da zu den TET-1 Aufnahmen momentan keine Wolkenmaske verfügbar ist, lässt sich deren Lokalisation und Umfang in diesem Fall nur aus den Temperaturunterschieden ableiten. Im Süden scheinen kleinskalige Wolkenstrukturen aufzutreten und im mittleren Bereich der Aufnahme können zwei größere kompaktere Wolken erkannt werden. Die im Norden der Szene auftretenden kühleren Gebiete sind wahrscheinlich durch Wasserdampf in der Atmosphäre entstanden. Dieser lässt die darunter liegenden Wasserfläche kälter erscheinen.

Im Rahmen der Masterarbeit werden außerdem die Unsicherheiten, welche sich aus der *Retrieval-Error-Covariance*-Matrix ableiten lassen, der Anteil der Messung an der Lösung, bestimmt aus der *Averaging-Kernel*-Matrix, die Kostenfunktion, die Anzahl an Iterationsschritten sowie die Konvergenz des Prozessors bewertet. Die Unsicherheit bzw. Genauigkeit der SST ist in Abbildung 4 (rechts) dargestellt. Es handelt sich dabei um die Werte der Hauptdiagonalen der *Retrieval-Error-Covariance*-Matrix. Diese berücksichtigt sowohl die Unsicherheit der Messung, als auch die des eingegangenen Globalmodells anteilig. Die angegebene Unsicherheit in dieser Szene beträgt $\pm 1,75$ bis $\pm 1,9$ K und ist damit größer als bei

momentan verfügbaren operationellen SST Produkten. Ein Grund hierfür ist allerdings im verwendeten Globalmodell zu sehen. Für dieses wurde mit einer Genauigkeit von ± 3 K angenommen, da dessen Auflösung sehr gering ist und innerhalb der Modellpixel starke Temperaturgradienten auftreten können. Die Verwendung eines höher aufgelösten Modells als Eingangsdatensatz könnte die Unsicherheit des Ergebnisses der SST reduzieren. Auffällig bei der Betrachtung der Unsicherheit ist der Bereich im Südosten der Szene. Hier scheint sich unter den kleinskaligen Wolkenstrukturen die Unsicherheit der SST zu verringern. Die Betrachtung der Kostenfunktion und des Konvergenzkriteriums (nicht abgebildet) zeigen jedoch, dass der Optimal Estimation Prozessor hier die Iteration nicht beenden kann und zu keinem Ergebnis kommt. Damit können diese Bereiche nicht ausgewertet werden.

Abschließend wurde eine Verifikation der Ergebnisse mit einem L-4 SST-Produkt der Gruppe GHRSSST durchgeführt (siehe Abbildung 4, links). Dafür wurde pixelweise die Differenz zwischen den errechneten Werte der SST und dem Vergleichsprodukt der GHRSSST gebildet. Hierbei zeigt sich eine systematische Abweichung von ca. $+2$ K bei allen untersuchten Szenen. Diese Abweichung verteilt sich größtenteils gleichmäßig über die gesamte Szene, wie in Abbildung 4 (links) zu sehen ist. Negative Differenzen treten nur im Bereich der Bewölkung auf, in denen der Optimal Estimation Prozessor zu keinem sinnvollen Ergebnis kommen kann. Einen leichten Einfluss scheint ebenfalls der Wasserdampfgehalt in der Atmosphäre auszuüben. Am nördlichen Rand der Szene mit erhöhtem Wasserdampfgehalt ist die positive Abweichung etwas stärker ausgeprägt.

5 Zusammenfassung

Die Sensitivitätsstudie zur Bestimmung der Meeresoberflächentemperatur zeigt eine weitere Anwendungsmöglichkeit der Messungen des Satelliten TET-1 auf. Mittels der hoch aufgelösten Daten ist es möglich die SST auf einer kleinen Skala zu bestimmen und so zukünftig beispielsweise Küsten- und Binnengewässer genauer zu untersuchen. Die zuvor durch Strahlungstransportsimulationen durchgeführte Sensitivitätsstudie ermöglicht einerseits die Abschätzung des Einflusses verschiedener atmosphärischer und technischer Parameter auf die SST und bildet andererseits die Grundlage der für die *Optimal Estimation* notwendige *Look-Up Table*. Besonders sensitiv reagiert die zu berechnende SST auf den Wasserdampfgehalt in der Atmosphäre. Es können im Kanal LWIR unter tropischen Bedingungen bis zu 10 K Abweichung zwischen SST am Boden und messbarer Brightness Temperature am Satelliten auftreten. Da andere Parameter die SST ebenfalls beeinflussen ist eine genaue Kenntnis der atmosphärischen Bedingungen zum Zeitpunkt der Messung nötig um möglichst exakte Ergebnisse zu erzielen.

Die mit Hilfe der *Optimal Estimation* bestimmte SST ist räumlich hoch aufgelöst und kann kleinskalige Turbulenzen wiedergeben. Die relative große Unsicherheit von $1,7 - 2$ K resultiert größtenteils aus dem als Vorwissen verwendeten Globalmodell. Als problematisch für das Verfahren haben sich große Wasserdampfkonzentration und unterschiedliche Bewölkung im Allgemeinen herausgestellt. Hier könnte zukünftig eine Wolkenmaske entwickelt und verwendet werden.

Die Verifikation der Ergebnisse mit einem SST-Datensatz der GHRSSST zeigt eine pixelweise systematische Abweichung von bis zu $+2$ K im Vergleich zur ermittelten SST. Diese

Abweichung liegt innerhalb des angegebenen Unsicherheitsbereichs, aber sollte dennoch genauer untersucht werden. Die Ursache dieser systematischen Abweichung in ist Rahmen der Masterarbeit nicht auszumachen. Es könnte sich hierbei beispielsweise um eine unvollständige und zu ungenaue Beschreibung der Atmosphäre handeln, oder um einen systematischen Fehler in der radiometrischen Kalibration der Satellitendaten.

6 Literaturverzeichnis

- BARALE, V., GOWER, J. F. R. & ALBEROTANZA, L., 2010: Oceanography from Space. Springer, Heidelberg
- EMERY, W. J., YU, Y. & WICK, G. A., 1994: Correcting Infrared Estimations of Sea Surface Temperature for Atmospheric Water Vapor Attenuation. *Journal of Geophysical Research: Ocean* **99**(C3), 2156-2202.
- FISCHER, C., KLEIN, D., KERR, G., STEIN, E., LORENZ, E., FRAUENBERGER, O. & BORGE, E., 2015: Data Validation and Case Studies using the TET-1 Thermal Infrared Satellite System. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **40**(7), 1177
- MAUL, G. A. & SIDRAN, M., 1973: Atmospheric Effects on Ocean Surface Temperature sensing from the NOAA Satellite Scanning Radiometer. *Journal of Geophysical Research* **78**(12), 1909-1916.
- MILLERO, F. J., 2005: *Chemical Oceanography*, 3rd Edition, CRC Press, Boca Raton (Fla).
- REILE, H., LORENZ E. & TERZIBASCHIAN, T., 2013: The FireBird Mission – A Scientific Mission for Earth Observation and Hot Spot Detection. *Small Satellites for Earth Observation. Digest of the 9th International Symposium of the International Academy of Astronautics.* Wissenschaft und Technik Verlag.
- RODGERS, C. D., 1976: Retrieval of Atmospheric Temperature and Composition from Remote Measurements of Thermal Radiation. *Reviews of Geophysics* **14**(4), 609-624.
- RODGERS, C. D., 2000: *Inverse Methods for Atmospheric Sounding: Theory and Practice.* World Scientific, Singapur.

Aufbau einer Geodateninfrastruktur am Beispiel Sokodé, Togo

KERSTIN MEYER¹

Zusammenfassung: Wie kann ich Karten günstig und schnell erstellen? Wofür brauche ich Karten und was sind deren Schwächen? Für Kommunen in Ländern des Globalen Südens, die bislang kaum kartographisch erfasst sind, wird mit dieser Arbeit eine Methode vorgestellt, um Karten einfach auf Grundlage von OpenStreetMap zu erstellen. Die Open-Source-Applikation OSMTracker für Android OS sowie JOSM und QGIS als Desktop-Anwendungen sind hierfür sehr geeignet, um aktuelles Kartenmaterial zum Aufbau einer Geodateninfrastruktur (GDI) zu erstellen. In Sokodé wurde mit dem Aufbau einer GDI begonnen und ein Netzwerk aus Akteuren geschaffen, die nun in der Lage sind, diese eigenständig fortzuschreiben und zu nutzen. Durch die Integration lokaler Akteure konnte so am Ende ein akzeptiertes Crowdsourced Katastermodell entstehen.

1 Einleitung

„Land is a basic need for all kinds of urban activities. Industries, housing, recreation, traffic (transport) all have their (conflicting) claims on land“ (TURKSTRA 1998). In vielen Ländern des Globalen Südens erfolgte die Verstädterung in Form einer „inversen Urbanisierung“ (BRILLEMBOURG TAMAYO et al. 2005). Sie folgt der Formel „Besetzung – Konstruktion – Infrastruktur – Grundbucheintrag“ (BERNT et al. 2007). Den Regierungen und dem Privatsektor gelingt es häufig nicht, genügend günstigen Wohnraum und Daseinsvorsorgefunktionen wie technische Infrastruktur oder Bildungs- und Gesundheitseinrichtungen zur Verfügung zu stellen. Da es nur wenige und oft nur prekäre und informelle Arbeitsplätze mit geringen Löhnen gibt, ziehen viele Familien zunächst in eine Lebensumwelt mit physischen Mängeln und auf Gelände, das ihnen nicht gehört. Illegale Stromanschlüsse, mangelnde und fehlerhafte Wasserversorgung, hygienische Missstände bei der Entsorgung, wilde Deponien sowie erhöhte Infektionsgefahr sind nur ein paar der daraus resultierenden Probleme. Es entstehen dadurch so genannte „Zwei-Klassen-Städte“, die meist im Zentrum ein modernes Ideal einer egalitären, geplanten und rationalen Stadt haben und von informellen Quartieren umgeben sind (BERNT et al. 2007). Diese komplexen städtischen Gegebenheiten zwingen Planer heute dazu, zielorientiert und schnell Lösungen zu finden. Partizipative Verfahren zur Sammlung von (Geo-)Daten und deren gemeinsame Analyse sind dafür ein probates Mittel. Die verwendeten Technologien müssen kostengünstig und einfach anwendbar sein, um überhaupt eine Art Grundversorgung an Planung zu gewährleisten (RAUT & RAUT 2015). Es sind die Bewohner/innen, die am besten wissen, wie viele Menschen etwa im Quartier wohnen, wie viele Gebäude es gibt, in welchem Zustand diese sind, wo ein Wasseranschluss vorhanden ist oder welche Häuser von Überschwemmungen bedroht sind (BERNT et al. 2007).

Ziel der Arbeit war es, ein geeignetes Open Source Verfahren zu entwickeln, das auf lokaler Ebene dazu verwendet werden kann, aktuelles Kartenmaterial schnell mithilfe von Smartphones und Laptops zu erstellen. Als Pilotkommune diente die Stadt Sokodé (Togo).

¹ TU Kaiserslautern, Computergestützte Planungs- und Entwurfsmethoden,
Pfaffenbergstraße 95, 67663 Kaiserslautern, Deutschland, E-Mail: Meyer.Kerstin1@t-online.de



Abb. 1: Aufbau der Arbeit (Quelle: Eigene Darstellung).

Abb. 1 stellt den Ablauf innerhalb der Arbeit dar, der langfristig in eine nationale GDI münden soll. Diese wiederum ist die Grundlage, um in Entscheidungsprozessen und Beteiligungsverfahren einen Überblick über bestimmte Themenbereiche zu bekommen und die o. g. Probleme zu lösen. Die Akzeptanz der Bevölkerung ist dabei elementar. „Land maps in general and cadastral maps in particular are designed to make the local situations legible to an outsider. For purely local purposes, a cadastral map was redundant“ (SCOTT 1998).

2 Geodateninfrastruktur

Der Begriff *Spatial Data Infrastructure* (SDI) entstand in der sog. Clinton-Order von 1994 und ist heute mit dem Wort Geodateninfrastruktur gleichzusetzen. Kurz darauf wurde das Thema international aufgenommen und Leitfäden für den Aufbau einer GDI veröffentlicht (SEIFERT 2008). Eine GDI besteht aus vernetzten Geodatenbanken und Funktionen zur Verwendung der Daten sowie einem Bereich institutioneller, organisatorischer, technologischer und wirtschaftlicher Ressourcen, die die GDI weiter entwickeln und aufrechterhalten. Dabei sollen die Geoinformationen verantwortungsvoll behandelt werden. Wie in Abb. 2 zu sehen ist, sind Nutzer, Netzwerk, Regeln, Standards, Dienste und Daten grundlegende Bestandteile einer GDI. Mechanismen, Regelwerke und Standards sind gleichermaßen bedeutend. Fehlt eines der Bestandteile, so ist die GDI in der Praxis nicht mehr funktionsfähig und in der Theorie nicht mehr als solche definiert (BERNARD et al. 2004 und SEIFERT 2008).

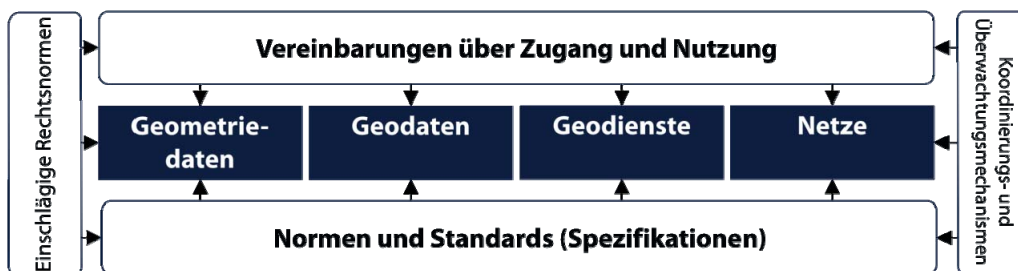


Abb. 2: Technische Komponenten und Rahmenbedingungen einer Geodateninfrastruktur (Quelle: Eigene Darstellung nach SEIFERT 2008: 14).

Mobile Geodatenerfassung zum Aufbau einer GDI findet dort statt, wo klassische Vermessung, Digitalisierung der analogen Datenbestände, Photogrammetrie etc. an ihre Grenzen kommen. Wenn Geodaten verschiedener Herkunft erzeugt werden, muss auf ein einheitliches Bezugs- und Koordinatensystem geachtet werden. Dies geschieht durch Georeferenzierung oder Geocodierung. Dabei wird zwischen Koordinatenumrechnung, d. h. ohne Bezugssystemwechsel, und Koordinatentransformation für verschiedene Bezugssysteme unterschieden (RESNIK & BILL 2009.). Durch die verbesserte Informations- und Kommunikationstechnologie entstehen internetbasierte Geodaten (3. Generation der Geoinformation). Die vierte Generation wird durch die Vernetzung und Interaktion der Akteure mit installierten GPS-Geräten in den Smartphones

und Computern unterstützt (Abb. 3). Kommentare, Fotos und Videos können direkt einer Online-Karte hinzugefügt und anderen Nutzer/innen zur Verfügung gestellt werden.

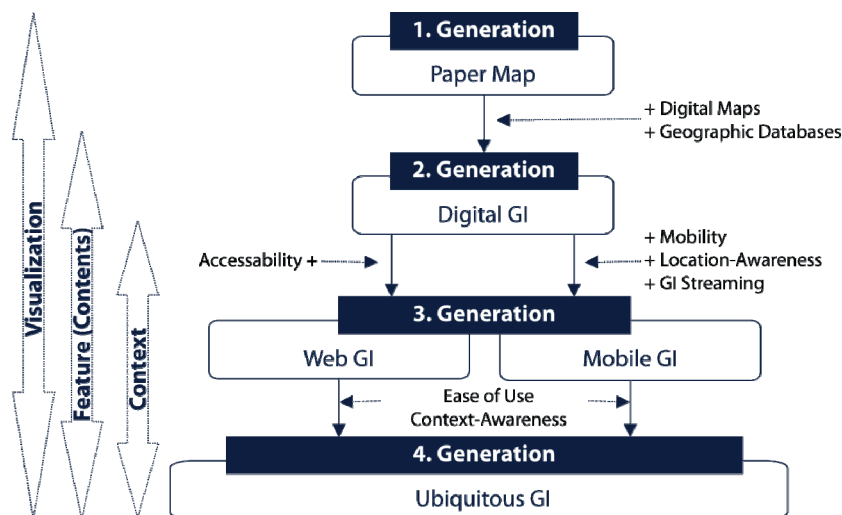


Abb. 3: Entwicklung der Geoinformation (GI) (Quelle: Eigene Darstellung nach SEIFERT 2008: 7)

„Push technology, can be used to keep people constantly but unobtrusively aware and remain engaged in discussions. Thus, the new concept of ‘Mashups’ is emerging in Urban and Regional Planning through Virtual technology” (RAUT & RAUT 2015).

2.1 Neue digitale Tools zur Erfassung von Geobasisdaten

„Eine Konservierung nicht mehr zeitgemäßer Strukturen [bei der Erfassung von Geobasisdaten] verlängert nur den Leidensprozess und behindert die innovative Entwicklung, hält sie aber nicht auf“ (ZIEGLER 2008). Nutzergenerierte Geoinformationen umfassen freiwillig bzw. bewusst sowie unfreiwillig durch den Anwender erfasste Geodaten. Im Bereich der freiwillig erfassten Geodaten tauchen Projekte wie *Volunteered Geoinformation* (VGI), Crowdsourcing oder *Citizen Science* auf. VGI basiert auf Internetkarten und kartographischen Informationen, die über komplexe Web-Mapping-Systeme an die Nutzer weitergegeben werden. (BILL et al. 2014). Schon im Jahr 2004 hatte der Brite Steven Coast das Ziel, eine freie Weltkarte zu schaffen, die zur Navigation und zur Erzeugung eigener graphischer Darstellungen dienen sollte. Es ging ihm dabei vor allem um den Open Source-Gedanken: Daten sollten nicht nur kostenlos betrachtet, sondern im Gegensatz zu den privaten Anbietern v. a. ohne urheberrechtliche Beschränkungen verfügbar und nutzbar gemacht werden. Alle, die etwas zur Datenbank von OpenStreetMap (OSM) beitragen, sind Dateneigentümer/innen und stellen die Daten seit 2012 unter der Open Database Licence (ODbL) nach dem Prinzip „create – adapt – share“ zur Verfügung. Bei OSM mit seinen Nachfolgeprodukten handelt es sich somit um eine Open Source-Anwendung, welche wie folgt definiert wird: Quellcode von Software, der öffentlich verfügbar, frei kopierbar, modifizierbar und nachnutzbar gemacht wird. „[OSM] is like the Wikipedia of maps. OSM is also more than a map, functioning as a data repository; users can collect and store almost any information that relates to the physical and built environment“ (ITOWORLD LTD. 2016). Durch vernetzte, mobile, digitale Geoinformationen sowie neue Open Source-Anwendungen zur Erfassung dieser wird ein Crowdsourcing Katastermodell, wie in Abb. 4 zu sehen ist, möglich.

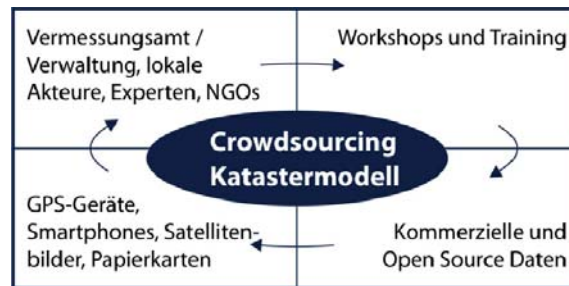


Abb. 4: Crowdsourcing Katastermodell (Quelle: Eigene Darstellung nach BASIOUKA et al. 2015: 334).

Im Folgenden werden die für die Arbeit relevanten, angewandten Methoden des Remote-Mappings, der Vor-Ort-Kartierung und anschließenden Bearbeitung der Daten kurz erläutert.

2.2 Remote-Mapping

Remote-Mapping bedeutet Kartieren aus der Ferne. Dabei geht es v. a. um das Abzeichnen von Satelliten- oder Luftbildern sowie gedruckten Karten oder Bildern. Die OSM-Editoren JOSM, iD und Potlach ermöglichen das Laden von verschiedenen, freien Satelliten- oder Luftbildern in den Hintergrund der Karte, wodurch diese einfach abgezeichnet werden können. Liegt das Luftbild falsch, sollte dies in JOSM oder Potlach anhand von GPS-Daten korrigiert werden (RAMM & TOPF 2010). Um bestimmte Gebiete schnell und problemlos zu kartieren, wurde von den Organisationen Missing Maps und dem Humanitarian OpenStreetMap Team (HOT) der HOT Tasking Manager entwickelt. Es handelt sich dabei um eine Homepage, wodurch Projekte weltweit bekannt gemacht, nach Priorität geordnet und die Freiwilligen organisiert werden. Bei der Bekanntgabe eines Projektes über diese Homepage ist es notwendig, einige Empfehlungen zu beachten, da bei einer globalen Community jedes Individuum zur Erfüllung der Aufgabe anders vorgeht und Fehler entstehen können. Tagging Schemata sollten verbessert werden, indem z. B. Fotos von ähnlichen Projekten oder Beispiele den Mappern zeigen, worauf sie beim Abzeichnen achten müssen. Abb. 5 zeigt den Ausschnitt der Stadt Sokodé zu verschiedenen Zeitpunkten während des Projektes im HOT Tasking Manager.

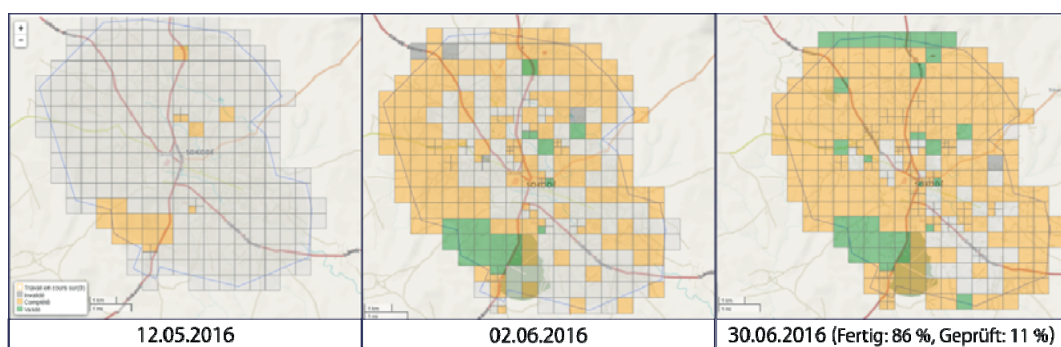


Abb. 5: Fortschritt des Missing Maps-Projektes von Sokodé im HOT Tasking Manager (Quelle: Eigene Darstellung. Karten- und Datengrundlage: HOT 2016).

2.3 Vor-Ort-Kartierung

Im Anschluss bzw. parallel zum Remote-Mapping bietet es sich an, vor Ort in den jeweiligen Gebieten mit interessierter Bevölkerung ebenfalls zu kartieren, Points of Interest (POIs)

hinzuzufügen und das bereits Kartierte zu überprüfen. Dazu empfiehlt es sich eine Mapping Party zu organisieren, die wie folgend beschrieben ablaufen könnte und in Abb. 6 zu sehen ist:

1. Treffen, an dem Neulinge kurz eingewiesen werden.
2. Verteilung von Aufgaben/Arbeitspaketen.
3. Individuelle Bearbeitung der Aufgaben/Arbeitspakete (z. B. anhand von FieldPapers): Positionsbestimmung des GPS-Gerätes/Smartphones, Wegpunkte am Gerät setzen, Attribute wie Name und Art des Gebäudes inkl. Straßenseite, der Brücke etc. notieren.
4. Nachbearbeitung der gesammelten GPS-Tracks z. B. mit JOSM: JOSM starten, GPX-Datei mit Track-Daten in JOSM öffnen, für den angezeigten GPS-Track die OSM-Dateien vom Server laden, Daten und Attribute ergänzen, neu erfassen und abspeichern.
5. Hochladen der Daten zum OSM-Server (RAMM & TOPF 2010).

In Deutschland wurden vor dem Einsatz in Sokodé bereits einige verschiedene mobile Anwendungen zur Geodatenerfassung getestet. Vor Ort hat sich am Ende der OSMTracker für Android als derzeit beste Applikation für das Smartphone erwiesen. Er ist sehr kompatibel mit OpenStreetMap und einfach in der Bedienung. GPS Tracks können aufgezeichnet und Wegpunkte (*Tags*, Sprachaufnahmen, Fotos) mit einem Klick gesetzt werden. Die Daten können als GPX gespeichert, direkt in OSM geladen oder in JOSM verwendet und abgezeichnet werden.

3 Fazit

Ein gutes Landmanagement zeichnet sich vor allem durch ein gutes Landverwaltungssystem aus, das akkurat und präzise, nachweisbar und verlässlich arbeitet und wodurch am Ende verbindliche Karten entstehen. Weitere Charakteristiken sollten eine flexible Handlungsweise, die integrativ sowie flächendeckend Lösungen unter Beteiligung der Bevölkerung findet, sowie die Erweiterbarkeit des Systems und die Finanzierbarkeit sein (ENEMARK et al. 2014).



Abb. 6: Schulung im Rathaus Sokodé (Quelle: Eigene Aufnahme am 24.05.2016)

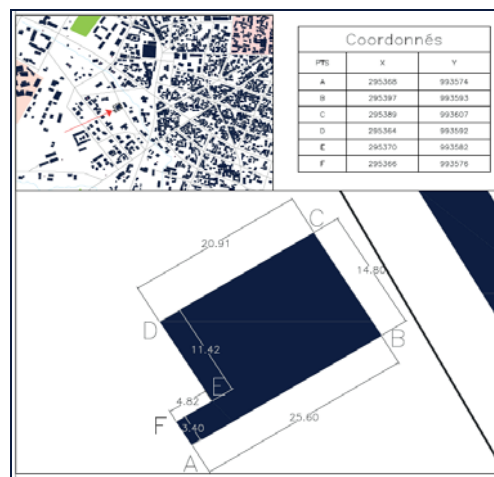


Abb. 7: Bemaßung des Rathauses Sokodé in AutoCAD (Quelle: Eigene Darstellung mit AutoCAD 2015. Datengrundlage: OSM 2016)

Bei der hier vorgestellten Methode zum Aufbau einer GDI z. B. für die Weiternutzung in einem Kataster (s. Bsp. in Abb. 7), handelt es sich um eine kostengünstige Vorgehensweise, die es erlaubt, die Daten schnell und leicht zu nutzen, zu erfassen, zu editieren und zu speichern.

Außerdem kann jeder über die Bearbeitung informiert werden und auch daran teilnehmen. Die Karten können anschließend in diversen Desktop-Programmen weiterverwendet werden. Probleme, die bei dieser Methode auftauchen können, sind, dass z. B. Laien unbewusst Informationen falsch darstellen oder Datenmaterial verändern. Attributdaten können zudem nur begrenzt erweitert werden (BASIOUKA et al. 2015). LAARAKKER und DEVRIES haben 2011 dafür den Begriff „opencadastré“ gewählt, woraus 2014 der Begriff durch NAVRATIL (2013) und DEVRIES et al. (2014) zu „neo-cadastré“ abgewandelt wurde (BASIOUKA et al. 2015). Zukünftig können Länder des Globalen Südens von dieser Methode profitieren und aufgrund der besseren Kartengrundlage bspw. die notwendigen Infrastrukturplanungen angepasst durchführen.

4 Literatur

- BASIOUKA, S., POTSIU, C. & BAKOGIANNIS, E., 2015: OpenStreetMap for cadastral purposes: an application using VGI for official processes in urban areas. *Survey review*, Bd. 47 (344), 333–341.
- BERNARD, L., CROMPVOETS, J. & FITZKE, J., 2004: Geoinfrastrukturen - ein Überblick., *Geodaten-Infrastruktur. Grundlagen und Anwendungen*, Bernard, L., Fitzke, J. & Wagner, R. M. (Hrsg.), Wichmann, Heidelberg, 1–8.
- BERNT, M., DANILJUK, M. & HOLM, A., 2007: Globalisierung und Spaltungen in den Städten. Informelle Urbanisierung, Selbstorganisation und "Sozialismus des 21. Jahrhunderts". *Partizipative Stadtentwicklung in den Barrios von Caracas. PROKLA 149 - Westfälisches Dampfboot - Zeitschrift für kritische Sozialwissenschaft*, 37. Jahrgang (4), 561–577.
- BILL, R., BERNARD, L. & BLANKENBACH, J., 2014: Geoinformationssysteme, *Das deutsche Vermessungs- und Geoinformationswesen 2015. Themenschwerpunkte Gesellschaftliche Verankerung und institutionelles Gefüge; Aufgabenfelder und Wirkungsbereiche; Technische Netzwerke und Transfer; Forschung und Lehre*, Kummer, K., Kötter & Eichhorn (Hrsg.), Wichmann, Berlin, 325–374.
- BRILLEMBOURG TAMAYO, A., FEIREISS, K. & KLUMPNER, H. (Hrsg.) 2005: *Informal city. Caracas case*. Prestel, Munich, New York.
- HOT, 2016: Tasking Manager - Projekt #1874. URL: <http://tasks.hotosm.org/project/1874> (10.11.2016).
- ITOWORLD LTD., 2011: OSM Tools. URL: <http://itoworld.com/product/open-data-tools/> (29.06.2016).
- RAMM, F. & TOPF, J., 2010: *OpenStreetMap: Die freie Weltkarte nutzen und mitgestalten*. Lehmanns.
- RAUT, S. K. & RAUT, P. B., 2015: Building Inclusive Smart Sustainable Cities through Virtual Environment., *Plan together right now overall. From Vision to Reality for Vibrant Cities and Regions*, Schrenk, M., Popovich, V. V., Zeile, P., Elisei, P. & Beyer, C. (Hrsg.), *Beiträge zur 20. internationalen Konferenz zu Stadtplanung, Regionalentwicklung und Informationsgesellschaft. CORP - Competence Center of Urban and Regional Planning*, Wien, 105–109.
- RESNIK, B. & BILL, R., 2009: *Vermessungskunde für den Planungs-, Bau- und Umweltbereich*. Wichmann Herbert.
- SCOTT, J. C., 1998: *Seeing like a state. How certain schemes to improve the human condition have failed*. Yale Univ. Press, New Haven.
- SEIFERT, M., 2008: *Wissenschaftlicher Beitrag für den Aufbau einer Geodateninfrastruktur zur Lösung von Aufgaben des E-Government*. ETH Zürich Inst. f. Geodäsie u. Photogrammetrie.
- TURKSTRA, J., 1998: *Urban Development and Geographical Information. Spatial and temporal patterns of urban development and land values using integrated geo-data*, Villavicencio, Colombia. o. V.
- ZIEGLER, P., 2008: *Infrastrukturen - Voraussetzungen für ein funktionierendes Gemeinwesen*. DVW e.V. - Gesellschaft f. Geodäsie GI u. Landmanagement (Hrsg.), *Geodateninfrastruktur - ein Beitrag zur Verbesserung unserer Lebensbedingungen: Beiträge zum 84. DVW-Seminar am 26. November 2008 in Potsdam (Schriftenreihe des DVW)*. Wißner-Verlag, 1–7.

Determination of the Freeze/Thaw Surface State from ERS-2 Backscatter Data

ISABELLA PFEIL¹, CHRISTOPH REIMER² & WOLFGANG WAGNER³

Summary: The freeze/thaw cycle of the Earth's surface determines the timing and the length of the vegetation growing season. Applications like permafrost monitoring and climate studies require information on the freeze/thaw state, however, conventional methods such as temperature records from in-situ stations are often severely hampered by the low density of weather stations. In order to obtain a global dataset of the freeze/thaw cycle, satellite remote sensing methods provide a good means since their observations of albedo, temperature or backscatter can be used to infer surface freeze/thaw state. The aim of this work was to derive freeze/thaw state from backscatter measurements by the scatterometer on-board the European Remote Sensing satellites. The results were validated against different soil models and in-situ networks.

1. Introduction

About 50 million km² of the terrestrial northern hemisphere experience seasonal surface freezing and thawing processes. Globally, about two thirds of the Earth's landmasses are subject to this phenomenon. The freeze/thaw cycle of the Earth's surface determines the timing and the length of the vegetation growing season and has a high impact on the land-atmosphere carbon dioxide exchange. It affects seasonal snowmelt and associated soil thaw, runoff generation and flooding, ice break up in large rivers and lakes, and trace gas dynamics (KIMBALL et al. 2001). For applications like permafrost monitoring and climate studies, information on the freeze/thaw state of the surface is highly valuable.

It is therefore important to monitor the changes between the surface states. Conventional methods such as using temperature records from in-situ stations are often severely hampered by the low density of weather stations, especially in Canada and Siberia. In order to obtain a global dataset of the freeze/thaw cycle, satellite remote sensing methods provide a good mean since their observations of albedo, temperature, snow coverage or backscatter can be used to infer surface freeze/thaw state.

The Remote Sensing research group at TU Wien has developed methods to retrieve global freeze/thaw states of the Earth's surface from backscatter measurements obtained from microwave scatterometers (NAEIMI et al. 2012). The algorithm for the retrieval of the surface state was originally developed for data from the Advanced Scatterometer (ASCAT), covering the period from 2007 until present. Since geoscientific studies require data from different periods, it's desirable to have long time series available.

¹ TU Wien, Centre for Water Resource Systems, Karlsplatz 13, A-1040 Vienna,
E-mail: isabella.pfeil (at) geo.tuwien.ac.at

² EODC Earth Observation Data Centre for Water Resources Monitoring GmbH,
Gußhausstraße 27-29/CA 02 06, A-1040 Vienna

³ TU Wien, Department for Geodesy and Geoinformation, Gußhausstraße 25-29/120, A-1040 Vienna

The aim of this work was to investigate if the ASCAT surface state algorithm can also be applied on data from the scatterometer (ESCAT) on-board the European Remote Sensing (ERS) satellites in order to obtain prolonged freeze/thaw time series, despite the lower amount of available observations because of limitations in the observation geometry and technical problems during the mission. The algorithm requires a certain amount of observations under different conditions in order to derive a surface state, which made the data availability the largest factor of uncertainty when starting the work on the algorithm adaptation.

In the following, the algorithm behind the surface state determination will be explained, along with results for different climate and land cover regions. The results were validated against soil and surface temperature data from the Global Land Data Assimilation System (GLDAS) and in-situ networks, as well as against arctic freeze/thaw soil state from the National Snow and Ice Data Center (NSIDC).

2. Determination of the Surface State

The calculation of the surface state flag is done using a threshold-analysis method. In a first step, backscatter values normalized to a common incidence angle are compared to the ERA Interim surface temperature dataset in order to derive the so-called freeze/thaw parameters. In most cases, the relationship between the two datasets follows a certain trend: In winter, when the soil moisture is lower and the soil is frozen, backscatter values are lower than in summer, where a high soil moisture dominates most landscapes. However, this behavior varies depending on climatic conditions and land cover. If possible, a logistic function is fit to all observations between -10°C and $+10^{\circ}\text{C}$, from which the freeze/thaw parameters can be derived.

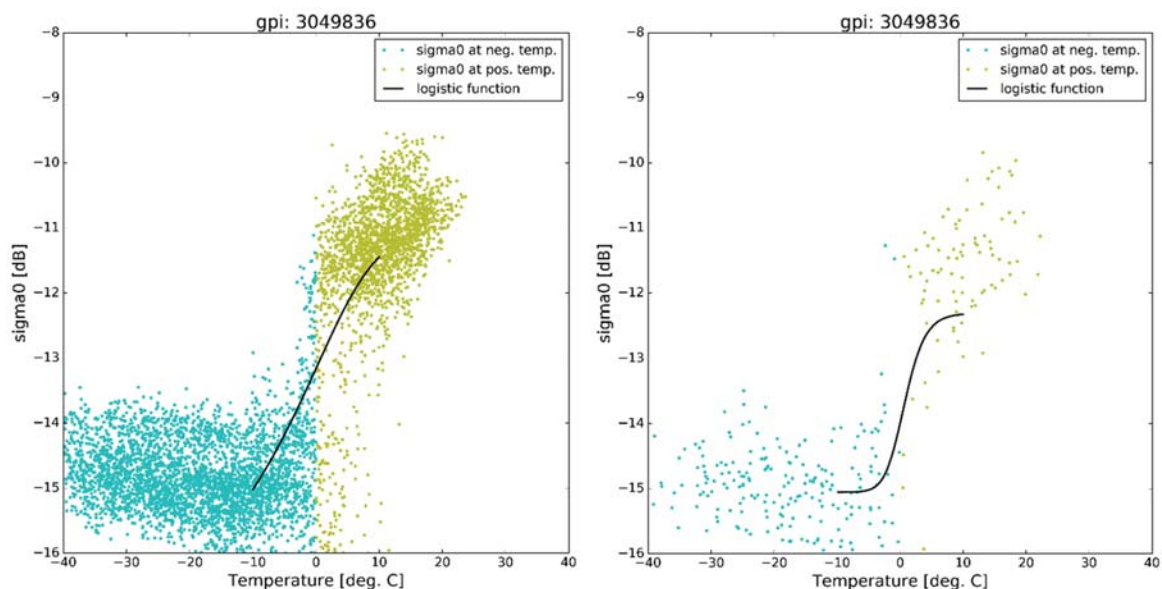


Fig. 1: Behavior of normalized backscatter against ERA Interim surface temperature for a grid point in Alaska, Happy Valley. (left: ASCAT, right: ESCAT)

A typical example for the backscatter-temperature relationship at higher latitudes is shown in Fig 1. The chosen grid point is located in the

Happy Valley, Alaska, at 69.16° northern latitude and 148.84° western longitude. This area shows low backscatter values in winter due to frozen soil, and high backscatter values due to vegetation growth and high soil moisture in summer. The two images show all ASCAT backscatter measurements from 2007-2013 (left) and all ESCAT measurements from 1997-2003 (right), respectively. The difference in the amount of observations is clearly visible. The logistic curves are displayed in black in the figure.

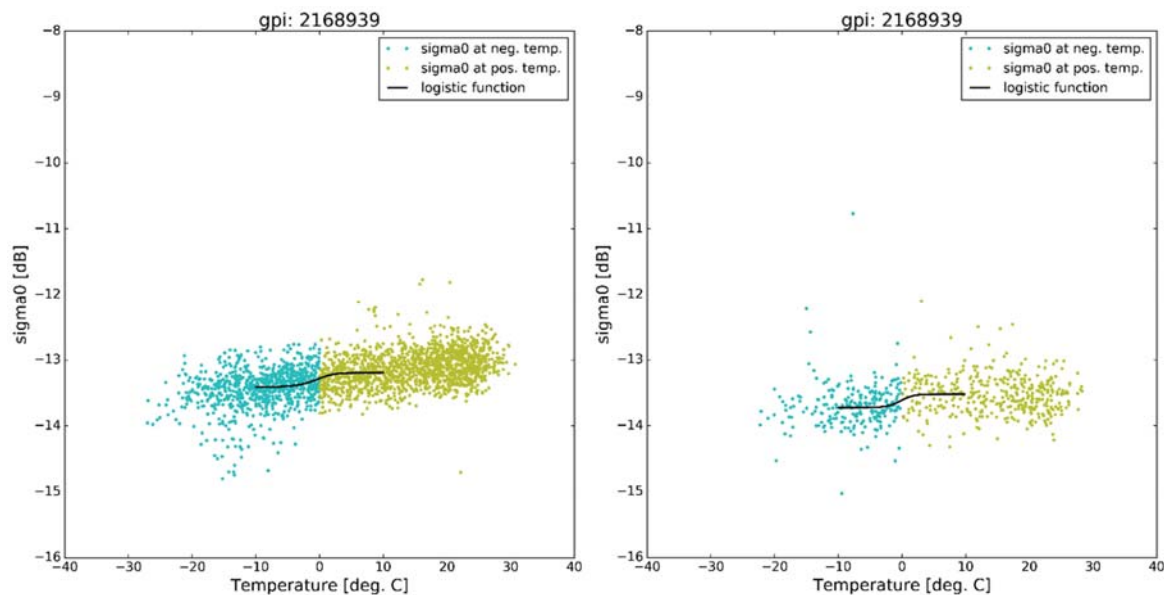


Fig. 2: Behavior of normalized backscatter against ERA Interim surface temperature for a grid point in China, Mazong Shan. (left: ASCAT, right: ESCAT)

An example where the discrimination of frozen and unfrozen soil is almost impossible is a grid point in Mazong Shan, in the Gobi desert (41.76° northern latitude, 97.25° eastern longitude, Fig. 2). This area is characterized by an extreme climate with very cold winters and very hot summers, with temperatures ranging from -40°C to $+40^\circ\text{C}$. The soil is extremely dry and there is no vegetation cover due to very low precipitation. Since the backscatter values from dry and frozen soil are about the same level, surface state changes are not clearly observable in this region.

If it is possible to fit a logistic curve, a set of parameters can be derived that are later used in the threshold-analysis. Those parameters include the mean backscatter level in summer and winter, the freeze/thaw threshold, which is found by calculating the inflection point of the logistic function, and the days of the year when the transition between winter and summer and summer and winter happens.

Those freeze/thaw parameters are then used to set up decision trees for the determination of the surface state flag (SSF). The SSF can take on the following values, which are displayed in the given color in the following:

- unfrozen (yellow)
- frozen (cyan)
- snowmelt/water on surface (green)

- permanent ice (purple)
- unknown (grey)

Fig. 3 shows the decision tree, which checks the observed backscatter value (σ_{40} in the figure) with six thresholds that are previously defined using the freeze/thaw parameters. Based on the season in which the backscatter value was observed, the decision tree returns the SSF.

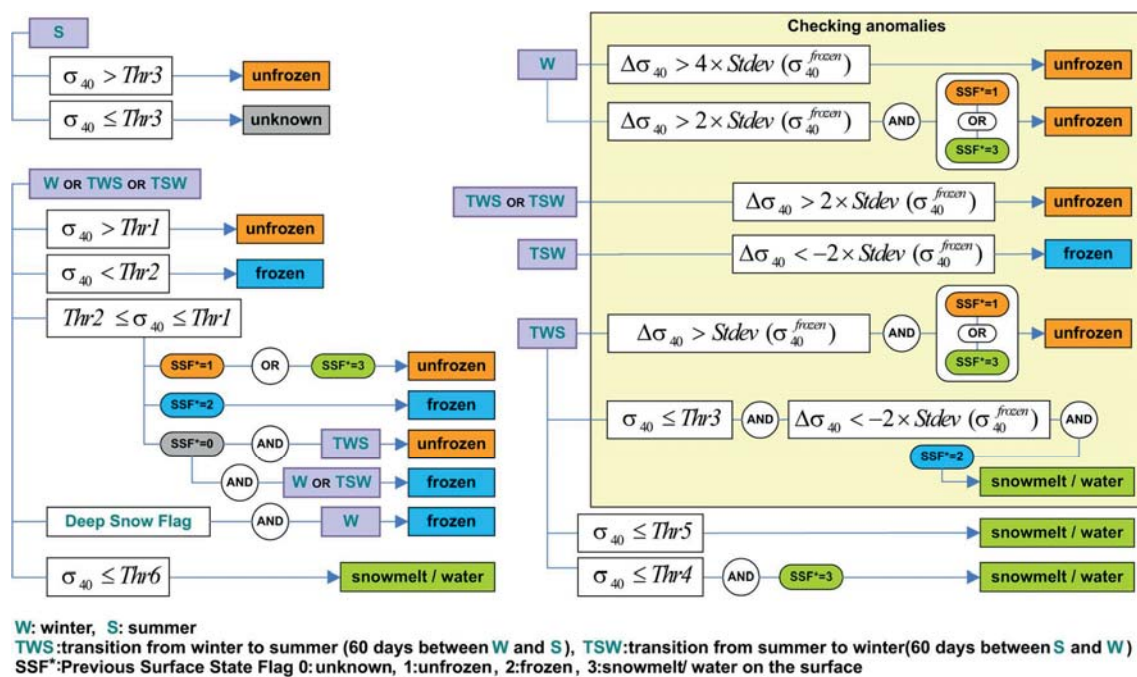


Fig. 3: Decision tree which returns the SSF for an observed backscatter value normalized to 40° incidence angle (σ_{40}) (NAEIMI et al. 2012)

The first run of the SSF algorithm with ESCAT backscatter data produced incorrect results. The main reason for these results were differences between the ASCAT and ESCAT input data format. After some adaptation of the code that does the data handling, the output for ESCAT SSF time series looked promising and the analysis and validation of the results was started.

3. Results of the Surface State Flag Determination

Fig. 3 shows example time series of surface states for a grid point in Russia, derived from ASCAT (left) and ESCAT (right), respectively. The grid point is located at 54.58° northern latitude and 36.73° eastern longitude in an area dominated by cold continental climate without a dry season, covered mostly with croplands. The SSF time series show very good agreement with the corresponding backscatter time series and the minimum and maximum temperature curves from the Met Office Hadley Centre (CAESAR et al. 2006). Different climate and land cover regions were selected to compare the ESCAT surface state flags with those retrieved from ASCAT backscatter data. The overall outcome shows very satisfying

results, contradicting the expectation that the low data availability might prevent a successful determination of the surface state from ESCAT data.

In order to obtain an impression not only for one grid point at a time but for a larger area, the SSF time series are plotted region-wise, meaning that the SSF time series of all grid points of one $5^\circ \times 5^\circ$ region (“cell”) are displayed above each other.

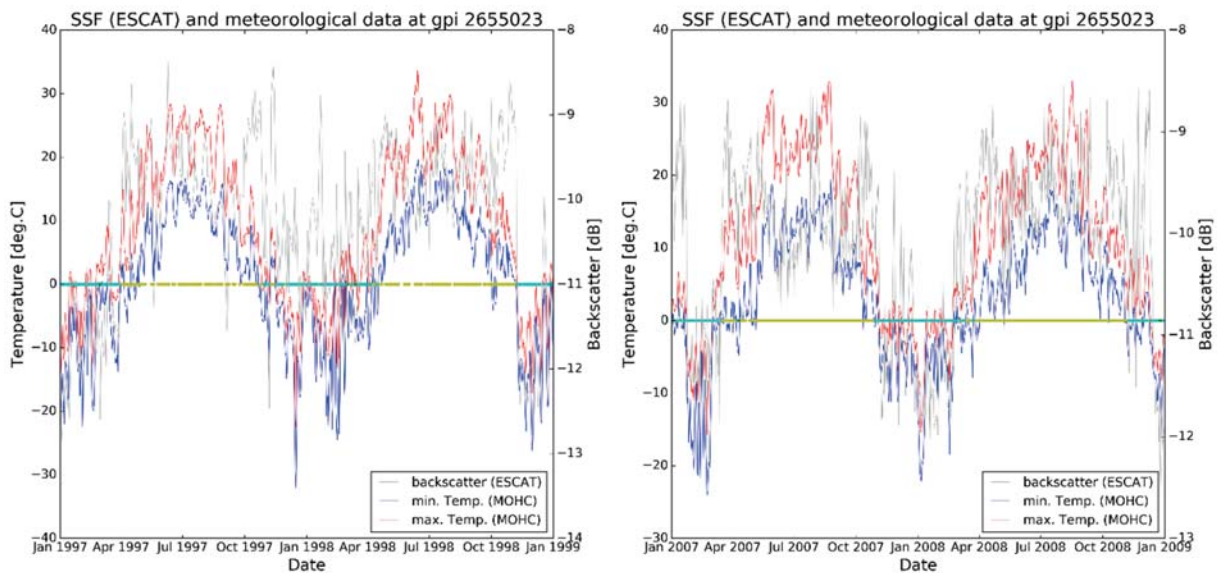


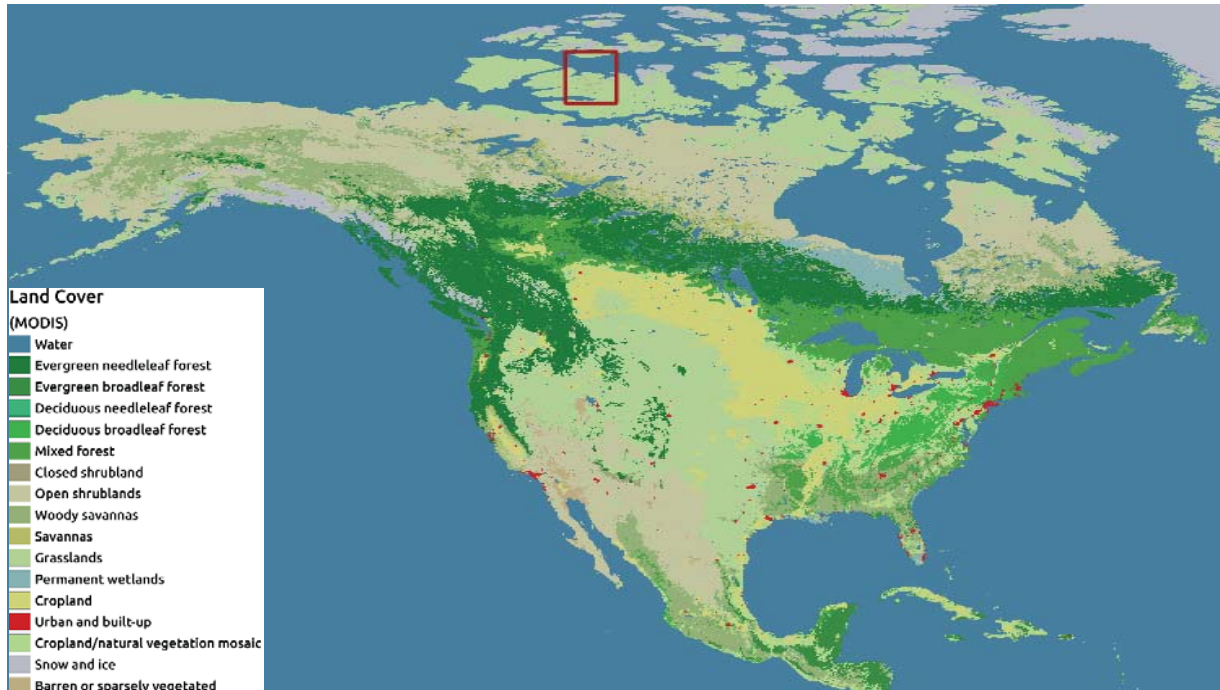
Fig. 4: SSF time series from ASCAT and ESCAT data compared to the corresponding backscatter data from ASCAT and ESCAT and minimum/maximum temperature data from the Met Office Hadley Centre

Fig. 5 and Fig. 6 show the location and results for a region in northern Canada ($70\text{--}75^\circ$ northern latitude, $110\text{--}115^\circ$ western longitude; cell 500). The land cover in this area mainly consists of grasslands, only some smaller areas show barren or sparsely vegetated land. According to the Koeppen-Geiger climate map, the region is dominated by polar Tundra climate (PEEL et al. 2007).

The surface is frozen for most time of the year, only during summer there is an unfrozen period of about two-three months. A few grid points show only frozen conditions throughout the year. Before every unfrozen period, most of the grid points show a melting period of about two weeks. In some parts of the area, the frozen period is interrupted by short-term melting events. The Surfaces State Flags derived from ASCAT and ESCAT data show very similar patterns. In the ESCAT results (Fig. 6, left), the data gap due to a failure of the gyroscopes is visible from 2001 onwards.

Fig. 8 shows the location of a selected region in temperate climate in northwestern Germany, the Netherlands and Belgium ($50\text{--}55^\circ$ northern latitude, $5\text{--}10^\circ$ eastern longitude; cell 1360). This region experiences warm summers and no dry season, and is characterized by heterogenous land cover, including mixed forest, croplands, natural vegetation mosaics, and urban areas. The heterogenous land cover reflects as vertical changes in the two plots shown in Fig. 7, that means that there are visible differences in the surface state between the different grid points. The frozen period lasts about two months in this region. This long duration is unexpected due to the influence

of the Gulf Stream on the climate of the area surrounding the North Sea. However, the SSF algorithm is designed to flag untrusted soil moisture values, which are retrieved from backscatter



values that might have been measured over frozen soils. The risk of flagging too many values is accepted in order to provide a reliable soil moisture product.

Fig. 5: Location of cell 500

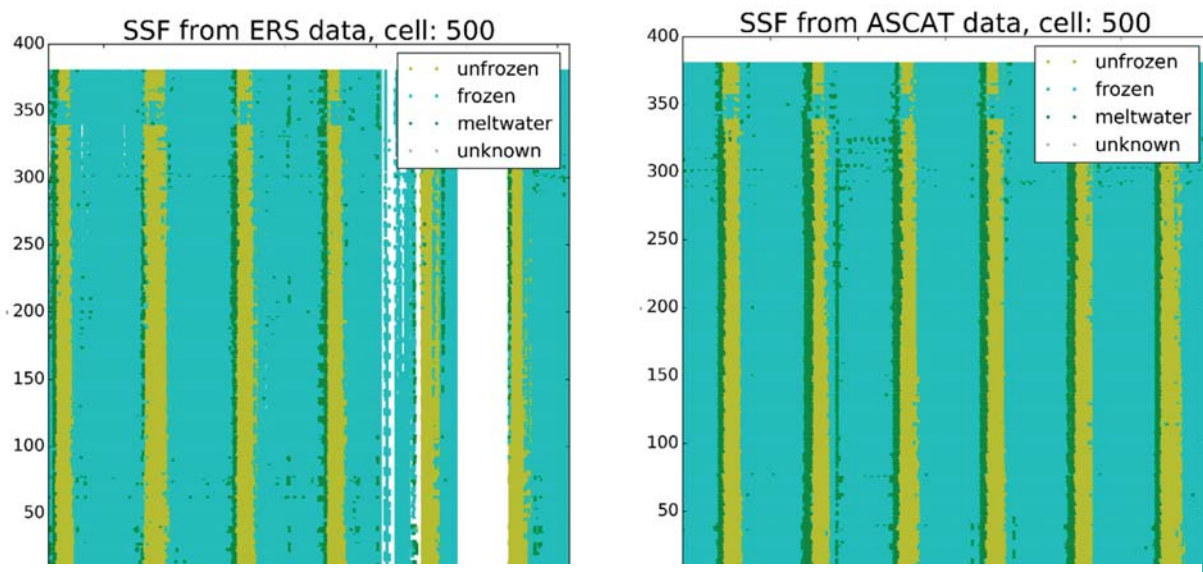


Fig. 6: SSF time series plot for cell 500, determined from ESCAT (left) and ASCAT (right) backscatter values. This region is located in polar Tundra climate in an area dominated by grasslands

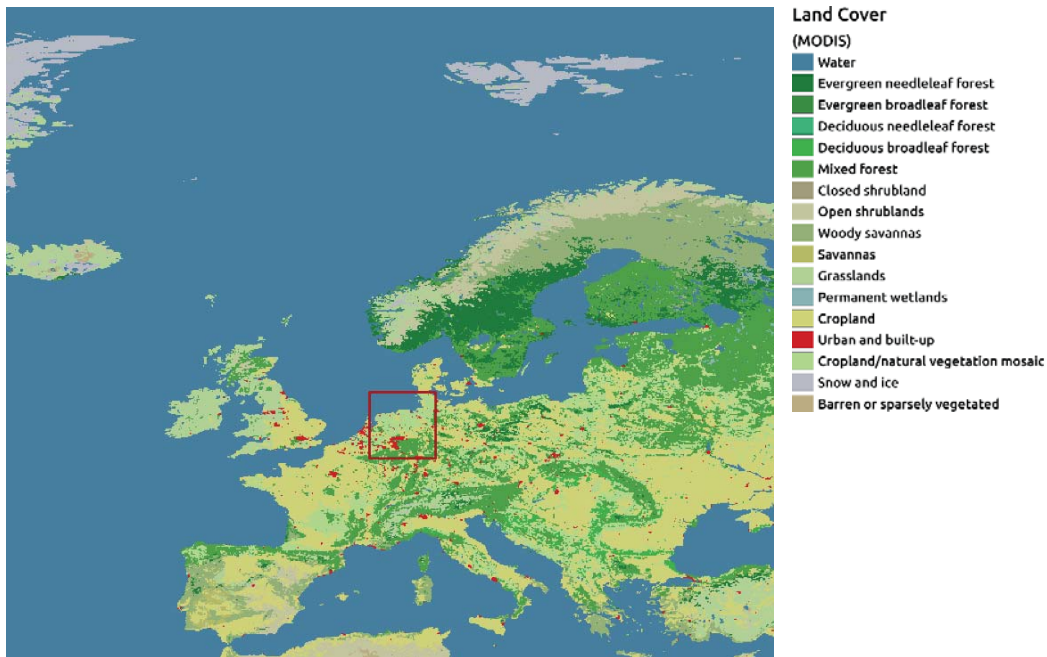


Fig. 8: Location of cell 1360.

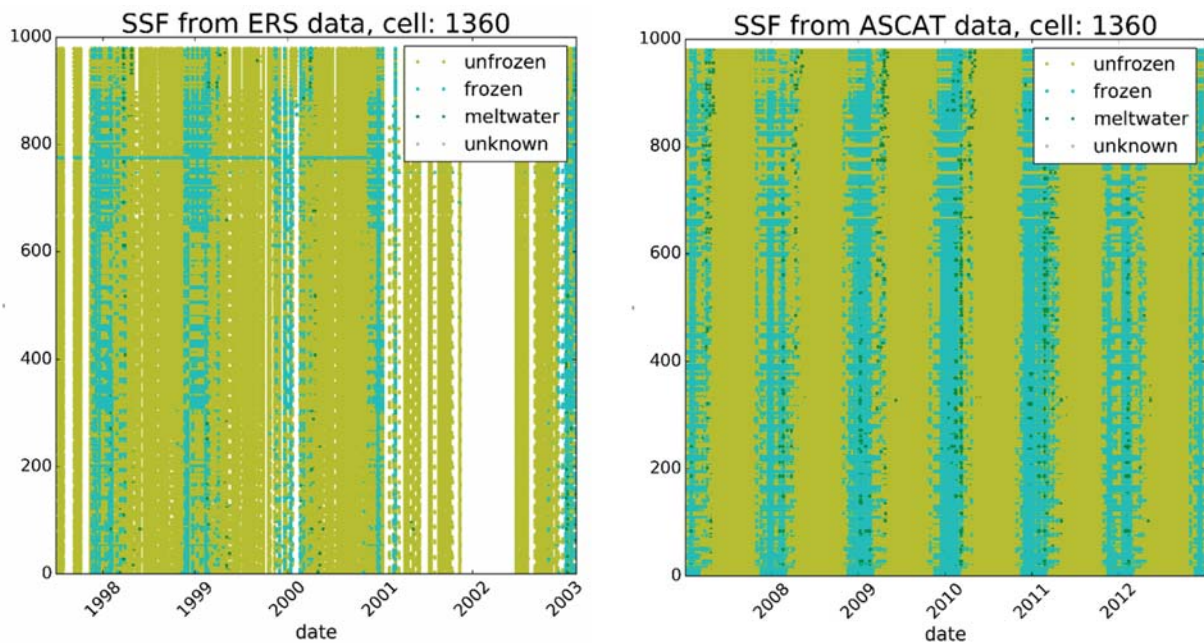


Fig. 7: SSF time series plot for cell 1360, determined from ESCAT (left) and ASCAT (right) backscatter values. This cell is located in temperate climate, in an area with very heterogeneous land cover

4. Validation and Discussion of the Results

The ESCAT surface state flags were validated against soil and surface temperature data from the Global Land Data Assimilation System (GLDAS) and in-situ networks, as well as against arctic freeze/thaw soil state from the National Snow and Ice Data Center (NSIDC). All validations show very good coherence between the datasets.

The validation against GLDAS data was furthermore divided into different periods, namely summer, winter, and the two transition periods between summer and winter (TZ1) and winter and summer (TZ2, Fig. 9 and Fig. 10). Blue and yellow bars show the percentage of corresponding observations in both datasets, either frozen or unfrozen surface. Red and green bars mean contradicting observations: red shows the percentage of SSF: frozen and GLDAS: unfrozen, green shows the percentage of SSF: unfrozen and GLDAS: frozen observations. Fig. 9 shows that the SSF determination works best in summer and winter in polar climate, when the freeze/thaw conditions are not as complex as in spring and fall, when the Earth's surface is exposed to multiple thawing and refreezing events in short time intervals. However, the performance of the algorithm is still sufficiently high in the transition periods. Similar results were observed for most of the test regions.

The validation of the SSF in temperate climate against GLDAS is shown in Fig. 10. This region is influenced by the Gulf Stream, and the soil is not permanently frozen during winter. Thawing and refreezing periods alternate depending on the weather situation.

The transition periods are defined as the period ± 30 days around the transition day. In some regions, especially at lower latitudes or close to the sea, this definition might not be suitable, since winter is not a season of permanent frozen conditions there (Fig. 10). Especially in those regions, a quite large number of unfrozen states is classified during winter, which suggests re-defining the transition periods. Considering this in the algorithm might deliver improved results for the surface state flag.

Building on the successful adaptation of the SSF algorithm for ESCAT data, the research on the algorithm shall be further continued. An assessment of the reliability of the computed SSF values could be valuable for different applications, by providing a confidence flag for each surface state flag. SSF values in summer and winter will be given higher reliability values than SSF values in the transition periods.

The SSF algorithm is based on ERA-Interim model temperature data. Since a model can never fully represent the true natural conditions, it would be very interesting to replace the model data with suitable land surface temperature observations, which might be available in the near future.

Thanks to this successful adaptation, the surface state time series can be extended to the end of the twentieth century. Since the freeze/thaw states of the Earth's surface are becoming more and more important for the understanding of our environment and the counteraction against climate change, this time series extension provides important data for numerous research areas.

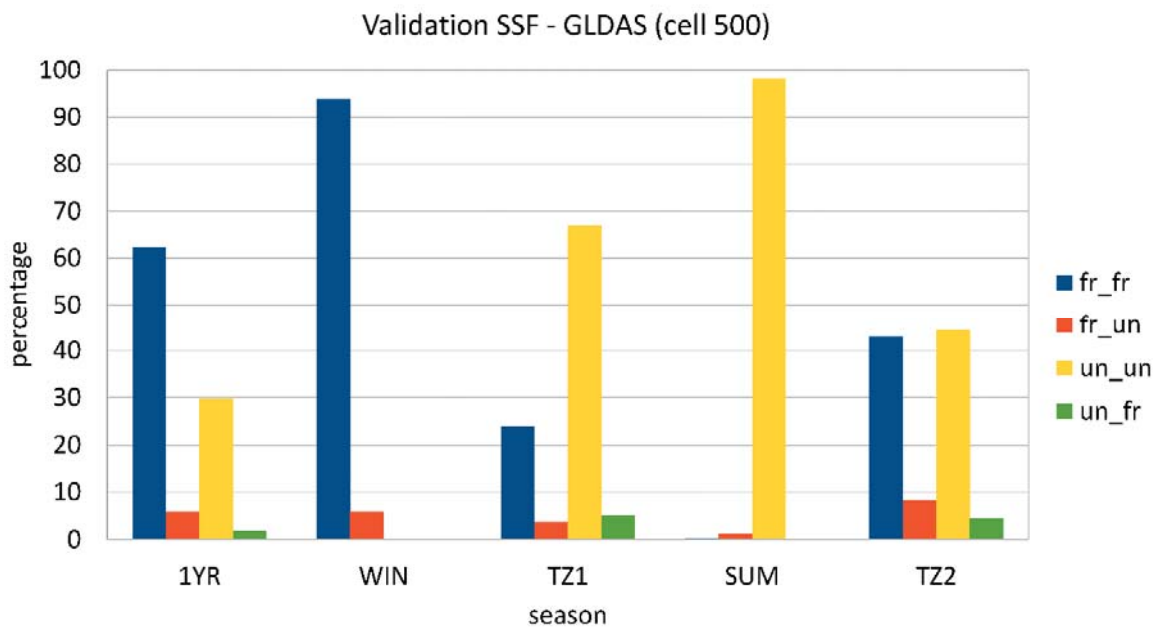


Fig. 9: Results of the validation of ESCAT SSF and GLDAS, divided into seasons. Cell 500 is situated in northern Canada, in polar Tundra climate. The region is covered by grasslands

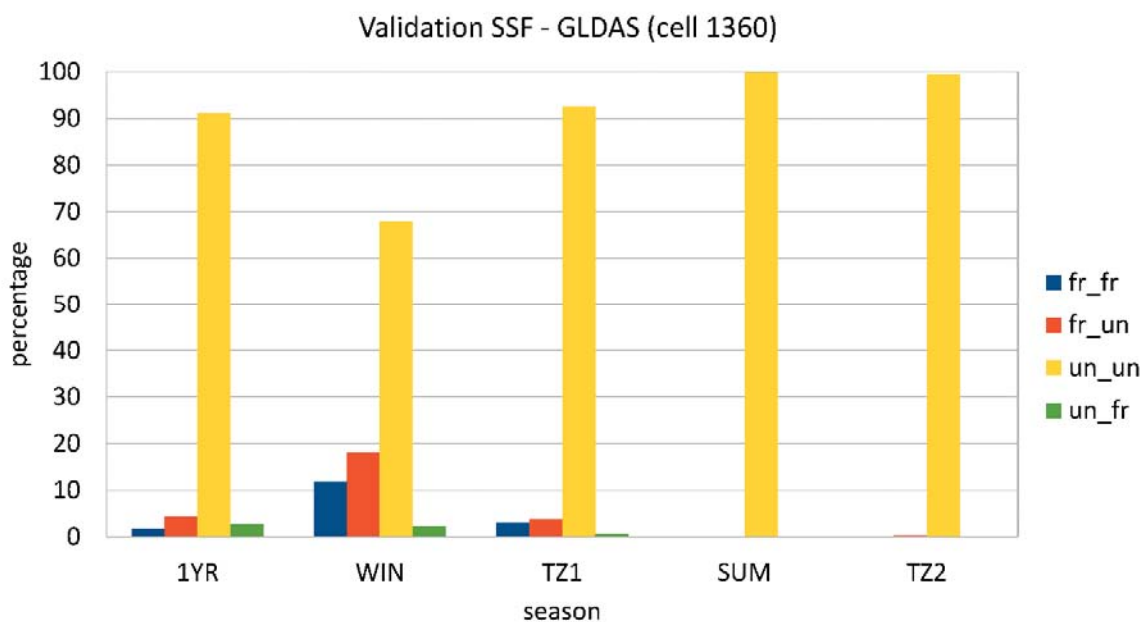


Fig. 10: Results of the validation of ESCAT SSF and GLDAS, divided into seasons. Cell 1360 is situated in Europe, in temperate climate influenced by the Gulf Stream. The land cover in this region is heterogenous; it includes mixed forests, croplands, and urban areas

5. References

- CAESAR, J., ALEXANDER, L. & VOSE, R., 2006: Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set. *Journal of Geophysical Research* **111**(D5), 30.
- KIMBALL, L., McDONALD, K., KEYSER, A., FROLKNING, S. & RUNNING, S., 2001: Application of the NASA scatterometer (NSCAT) for determining the daily frozen and nonfrozen landscape of Alaska. *Remote Sensing of Environment* **75**(1), 113-126.
- NAEIMI, V., PAULIK, C., BARTSCH, A., WAGNER, W., KIDD, R., PARK, S.-E., ELGER, K. & BOIKE, J., 2012: ASCAT Surface State Flag (SSF): Extracting information of surface freeze/thaw conditions from backscatter data using an empirical threshold-analysis algorithm. *IEEE Transactions on Geoscience and Remote Sensing* **50**(7), 2566-2582.
- PEEL, M. C., FINLAYSON, B. L. & MCMAHON, T. A., 2007: Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences Discussions* **4**(2), 439-473.

Entwicklung einer Tourenplanungssoftware (SNOWMAPS) auf Basis von abgeleiteten Karten aus Höhendaten

MATTHIAS PILZ¹ & KONRAD RAUTZ¹

Zusammenfassung: Für Ski- und Schneeschuhtourengeher sowie Freerider wird in Zusammenarbeit mit dem steirischen Lawinenwarndienst eine Tourenplanungssoftware erstellt. Als Planungsgrundlage dienen dem Benutzer neben topografischen Grundkarten und Orthofotos auch aus Airborne-Laserscanning (ALS) Höhendaten abgeleitete Hangneigungs-, Expositions- sowie Bewuchskarten (ALS-derived Attribute Maps). Die einzelnen Komponenten der Software werden in ihrer Funktionalität erklärt. Insbesondere wird die Bedeutung der Topografie für die Einschätzung der Lawinengefahr erläutert. Ab dem Winter 2016/17 steht SNOWMAPS kostenlos allen Benutzern zu Verfügung.

1 Einleitung

Ziel von SNOWMAPS ist es, Ski- bzw. Schneeschuhtourengeher ein professionelles Werkzeug zur Planung ihrer Aktivitäten abseits gesicherter Pisten zur Verfügung zu stellen. SNOWMAPS soll dabei als zweites Hilfsmittel neben dem aktuellen Lawinenlagebericht verwendet werden. Das Hauptaugenmerk liegt dabei neben der freien Zugänglichkeit, welche durch die Browser-Software realisiert wird, auf der Bereitstellung und Aufbereitung hochqualitativer Geodaten. Hierbei liegt der Fokus bei der bestmöglichen Auswahl und Aufbereitung dieser Daten für lawinenrelevante Beurteilungen. Weiters wird großer Wert auf einfache und intuitive Benutzerinteraktion gelegt.

Langfristig soll durch die breite Verwendung von SNOWMAPS die Zahl der Lawinenunfälle reduziert werden. Denn eine gute Tourenplanung sorgt für den entscheidenden Sicherheitsvorsprung auf einer Ski-, Schneeschuh- bzw. Freeridetour! In weiterer Folge wird nur noch von Skitouren gesprochen werden.

2 Problemstellung

2.1 Grundaufbau

Die technische Umsetzung der Tourenplanungssoftware SNOWMAPS erfordert eine komplexe Zusammenarbeit von Auszeichnungssprachen, Stylesheetsprachen, Skriptsprachen und Programmiersprachen. Das Grundgerüst der Homepage wird mit HTML aufgebaut, die Formatierung mittels CSS umgesetzt. Das Abfragen der aktuellen Gefahrenstufe erfolgt mittels PHP. Das interaktive Kartenfenster wird mit JavaScript unter Verwendung der Open Layers 3 Bibliothek programmiert.

¹ TU Graz, Institut für Geodäsie, Arbeitsgruppe für Geoinformation, Steyrergasse 30, A 8010 Graz, E-Mail: matthias-pilz@gmx.at, konrad.rautz@tugraz.at

2.2 Die aktuelle Lawinengefahr

Die aktuelle Lawinengefahr wird, meist täglich, durch die regionalen Lawinenwarndienste erhoben und durch den Lawinenlagebericht mitgeteilt. Dieser Bericht enthält neben einer regionalen Einschätzung der Gefahr in Form der Gefahrenstufe (siehe Abb. 1) Informationen über eine vorherrschende typische Situation, die Gefahrenbeurteilung, den Schneedeckenaufbau, das Wetter und die Tendenz. Der Lawinenlagebericht sollte bei jeder Skitourplanung herangezogen werden und steht über zahlreiche Kanäle kostenlos zur Verfügung.

Gefahrenstufe	Icon	Schneedeckenstabilität	Lawinen-Auslösewahrscheinlichkeit
5 - Sehr gross		Die Schneedecke ist allgemein schwach verfestigt und weitgehend instabil.	Spontan sind viele grosse, mehrfach auch sehr große Lawinen, auch in mässig steilem Gelände zu erwarten.
4 - Gross		Die Schneedecke ist an den meisten Steilhängen schwach verfestigt.	Lawinenauslösung ist bereits bei geringer Zusatzbelastung** an zahlreichen Steilhängen wahrscheinlich. Fallweise sind spontan viele mittlere, mehrfach auch grosse Lawinen zu erwarten.
3 - Erheblich		Die Schneedecke ist an vielen Steilhängen* nur mässig bis schwach verfestigt.	Lawinenauslösung ist bereits bei geringer Zusatzbelastung** vor allem an den angegebenen Steilhängen* möglich. Fallweise sind spontan einige mittlere, vereinzelt aber auch grosse Lawinen möglich.
2 - Mässig		Die Schneedecke ist an einigen Steilhängen* nur mässig verfestigt, ansonsten allgemein gut verfestigt.	Lawinenauslösung ist insbesondere bei grosser Zusatzbelastung**, vor allem an den angegebenen Steilhängen* möglich. Grosse spontane Lawinen sind nicht zu erwarten.
1 - Gering		Die Schneedecke ist allgemein gut verfestigt und stabil.	Lawinenauslösung ist allgemein nur bei grosser Zusatzbelastung** an vereinzellen Stellen im extremen Steigelände* möglich. Spontan sind nur Rutsche und kleine Lawinen möglich.

Abb. 1: Europäische Lawinengefahrenstufenskala (EAWS 2016)

2.3 Der Einfluss der Topografie auf die Lawinengefahr

Ein alter Spruch lautet: „Wer nichts von Schneedeckenaufbau und Lawinengefahr versteht, der muss zumindest ein Meister des Geländes sein!“

Der Topografie kommt eine wichtige Rolle bei der Entstehung von Lawinen zu (GEBETSBERGER 2010). Topografische Details haben einen grossen Einfluss auf die horizontale und vertikale Verteilung der Schneedecke. Doch selbst innerhalb eines einzelnen Hanges ist der Schneedeckenaufbau sehr unterschiedlich. Die Mächtigkeit ist nicht immer gleich und der Aufbau der einzelnen Schichten ist nicht immer homogen. Die aus einer lokalen Schneedeckenuntersuchung gewonnene Information darf demnach keinesfalls direkt auf den gesamten Hang übertragen werden. Problematisch dabei ist das Erkennen der ungünstigen Bereiche (Schwachstellen) innerhalb eines Hanges. Tendenziell kann aber davon ausgegangen werden, dass es günstige und weniger günstige Bereiche gibt. Neben den ohnehin bekannten ungünstigen Geländeformen wie Rinnen, Mulden und Hohlformen finden sich innerhalb dieser ungünstigen Formen im Detail noch besonders heikle Bereiche:

- Ränder von Rinnen und Mulden
- Geländeübergänge (flach zu steil oder steil zu flach)
- Übergänge von Schneemächtigkeiten (gering zu gross, aber auch gross zu gering)

An solchen Schwachstellen innerhalb eines Hanges kann ein Initialbruch deutlich leichter ausgelöst werden als an gut verfestigten Stellen, was dann wiederum die Bruchfortpflanzung aktiviert. Diese Stellen werden daher auch „superschwache Zonen“ oder „Hot-Spots“ genannt. Je größer die Lawinengefahr, umso mehr Hot-Spots sind vorhanden und damit steigt auch die Auslösewahrscheinlichkeit deutlich (siehe Abb. 2 und Abb. 3) (ZENKE 2012). Nach SCHWEIZER (2006) nimmt die Lawinenauslösewahrscheinlichkeit von Gefahrenstufe zu Gefahrenstufe exponentiell zu.



Abb. 2: Exemplarische Verteilung der Hot-Spots bei geringer Lawinengefahr (Gefahrenstufe 1) (eigene Darstellung, Bildgrundlage: eigenes Archiv)



Abb. 3: Exemplarische Verteilung der Hot-Spots bei großer Lawinengefahr (Gefahrenstufe 4) (eigene Darstellung, Bildgrundlage: eigenes Archiv)

Die Hot-Spot Theorie besagt auch, dass die Auslösewahrscheinlichkeit weitgehend unabhängig von der Hangsteilheit ist, sie wird vor allem durch den Schneedeckenaufbau und damit durch Exposition, Geländeform und Bewuchs beeinflusst. Nicht zu verwechseln ist die Auslösewahrscheinlichkeit mit der Wahrscheinlichkeit des Abgleitens eines Schneebretts, diese ist primär von der Hangsteilheit abhängig (SCHWEIZER 2006).

3 SNOWMAPS

Mit der Online-Plattform SNOWMAPS können Skitouren auf neuartige Weise geplant werden. Durch die gemeinsame Verwendung verschiedener Informationen kann der Nutzer daraus komplexes Wissen ableiten. Bei der Gestaltung der Website von SNOWMAPS wird besonders Wert darauf gelegt, sämtliche Funktionen jederzeit am Display zu haben. Dazu wird eine bildschirmfüllende Anzeige der Website verwendet (siehe Abb. 4). Dominierend ist hierbei das große Kartenfenster (A). Innerhalb des Kartenfensters finden sich Zoomleiste, Maßstabsbalken sowie Urheberrechtsinformation. Im oberen rechten Teil befindet sich das Analyse-Fenster (B), welches im Grundzustand das SNOWMAPS Logo zeigt. Hier werden nach der Tourenplanung das Höhenprofil sowie die Ergebnisse der Wegzeitberechnung angezeigt. Darunter befindet sich das Kartencenter (C), in welchem jederzeit sämtliche Karten oder Overlays ein- und ausgeblendet werden können sowie deren Transparenz geändert werden kann. Hier gibt es auch Links zum aktuellen Lawinenlagebericht sowie zu einem detaillierten Bergwetterbericht. Zudem finden sich kleine Informationsbuttons, mit deren Hilfe jederzeit ein einfaches Anzeigen der jeweiligen Kartenlegende möglich ist.

Des Weiteren gibt es ein geografisches Suchfenster, mit welchem nach Gipfel-, Orts- oder anderen geografischen Namen gesucht werden kann. Unterhalb der Karte ist die sogenannte Toolbar (D) angeordnet.

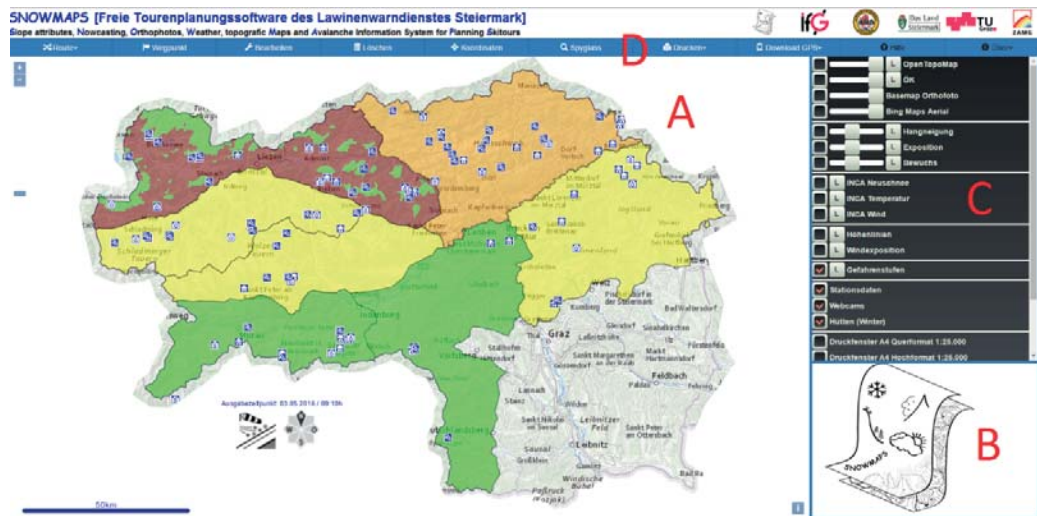


Abb. 4: Grundaufbau von SNOWMAPS (eigene Darstellung, Kartengrundlage: Basemap, Geobasisdaten: GIS-Steiermark, Lawinenwarnstufen: Lawinenwarndienst Steiermark)

3.1 Kartencenter

Im Kartencenter können einzelne Layer ein- und ausgeblendet sowie deren Transparenz verändert werden. Jederzeit kann auch eine entsprechende Legende abgerufen werden. Karten bzw. deren Informationen können somit beliebig kombiniert und daraus kann komplexes Wissen abgeleitet werden.

Hangneigungs-, Expositions- und Bewuchskarten werden unter dem Begriff ALS-derived Attribute Maps zusammengefasst. Im Rahmen der durchgeführten Experimente zeigt sich, dass durch Verwendung der Ausgangsdaten mit 10m Auflösung die besten Ergebnisse für den weiteren Einsatz erreicht werden können.

3.1.1 Topografische Basislayer

Die topografischen Karten bilden die Grundlage jeder Tourenplanung. Um dem Benutzer ein möglichst umfangreiches Bild der Situation zu geben, stehen verschiedene Basislayer zur Verfügung. Eine topografische Karte, in diesem Fall der Layer „Grau“ von Basemap, wird immer als Grundkarte angezeigt. Darüber hinaus können folgende Layer angezeigt werden:

- OpenStreetMap
- ÖK 25V
- Basemap Layer „Orthofoto“ mit Basemap „Overlay“
- Bing Maps „Satellite“

Die ÖK 25V stellt mit ihrer Rasterdarstellung immer noch die beste Grundlage zur alpinen Tourenplanung dar. Der hohe Detailreichtum an Informationen, besonders in Bezug auf die Geländebeschaffenheit, ist allen anderen Karten überlegen (WINKLER et al. 2008). Die amtlichen Ortsbezeichnungen der ÖK finden sich auch in den meisten Skitouren-Führern wieder.

3.1.2 Thematische Basislayer

Den thematischen Basislayer bildet die Gefahrenstufenkarte. Dabei werden die aktuellen Gefahrenstufen grafisch aufbereitet und der Grundkarte überlagert. Eine ganzjährig gültige und

höhenunabhängige Gefahrenstufe wird dabei durch eine statische Farbgebung der entsprechenden Gefahrenstufenfarbe visualisiert. Herrscht eine höhenabhängige Gefahrenstufe, so wird diese auch in ihrer tatsächlichen geografischen Ausdehnung dargestellt. Eine vom Tagesgang betroffene Gefahrenstufe wird durch einen dynamischen Farbübergang visualisiert. Ergänzt werden diese Informationen mit meteorologischen Daten wie Windgeschwindigkeit, Windrichtung, Böen, Temperatur und Neuschneezuwachs. Webcams, meteorologische Stationsdaten sowie im Winter geöffnete Hütten werden als POI's überlagert. Durch einen Klick auf den jeweiligen Punkt werden entsprechende Informationen angezeigt.

3.1.3 Hangneigungskarte

Die Farbgebung und Klassifizierung wird basierend auf Überlegungen bezüglich der Verwechslungsgefahr mit den Farben der Gefahrenstufen gewählt. Diese Art der Darstellung ermöglicht auf den ersten Blick eine Zweiteilung des Geländes in Bereiche mit Hangsteilheit unter 30° und über 30° . Graue Flächen repräsentieren dabei Gelände unter 30° , farbige Flächen Gelände über 30° (siehe Abb. 5). Damit wird die Hangneigungskarte der Erkenntnis gerecht, wonach die potenzielle Lawinengefahr ab 30° massiv ansteigt (KRONTHALER 2011). Vor allem für unerfahrenere Tourengeher bewirkt die Vermeidung der „bunten Hänge“ eine deutliche Risikoreduktion. Experten können in Kombination mit anderen Informationen, beispielsweise der Geländeinformation über Höhenschichtenlinien, eine Analyse über möglicherweise in einem Hang vor sich gehende bzw. gegangene Prozesse durchführen. Auch wenn die Hangneigung vor allem bei der praktischen Durchführung einer Tour nach aktuellem Wissen eine geringere Rolle spielt als bisher angenommen, so ist ihre Wichtigkeit in der Tourenplanung unumstritten.

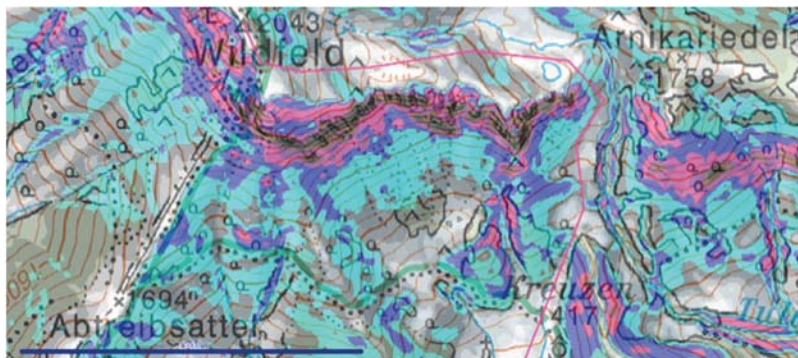


Abb. 5: Geplante Route mit eingeblendeter Hangneigung (eigene Darstellung, Kartengrundlage: Basemap, Geobasisdaten: GIS-Steiermark)

Des Weiteren wird SNOWMAPS damit auch den Ansprüchen der auf der klassischen Reduktionsmethode basierenden Strategien gerecht, in welchen die Hangneigung den wichtigsten Verzichtsfaktor bildet (WÜRTL 2011). In diesem Fall ist eine Klassifizierung der Hänge in 5° -Schritten unumgänglich.

3.1.4 Bewuchskarte

Die Bewuchskarte dient in erster Linie der Erkennung der in einem Gebiet herrschenden Vegetation. Diese wiederum beeinflusst einerseits die Windeinwirkung innerhalb eines Hanges sowie andererseits die Eignung desselben für die Befahrung durch Skifahrer.

3.2 Toolbar

In der Toolbar sind alle dem Benutzer zur Verfügung stehenden Interaktionsfunktionen für das Zeichnen einer Haupt- sowie einer Alternativroute zusammengefasst. Weiters ist die Eingabe von Wegpunkten möglich. Routen und Wegpunkte können mit der Funktion „Bearbeiten“ geändert, ergänzt oder reduziert werden. Eingezeichnete Elemente sind in gängigen Formaten auf ein GPS-Gerät übertragbar und der gesamte Kartenausschnitt kann ausgedruckt werden.

4 Fazit & Ausblick

In dem Beitrag wurde ein neuartiges System zur Entscheidungsfindung (Planung) für Skitourengeher vorgestellt. Die gemeinsame Bereitstellung dafür relevanter Daten wie ALS-derived Attribute Maps, meteorologischer Daten sowie topografischer Basisdaten stellt eine Neuerung auf diesem Sektor dar und soll zu einer entscheidenden Erhöhung der Sicherheit beitragen. Bisher konnten diese Informationen nur aus unterschiedlichen Kanälen zusammengetragen werden. Im Vordergrund steht dabei die Bereitstellung hochqualitativer Informationen, auf Basis welcher selbstständige Entscheidungen getroffen werden können. SNOWMAPS ist ab dem Winter 2016/17 unter www.snowmaps-steiermark.com verfügbar.

Hinweis: Um eine leichtere Lesbarkeit zu gewährleisten, wird im Beitrag die Sprachform des generischen Maskulinums verwendet. Es wird betont, dass die Verwendung dieser Form geschlechtsneutral zu verstehen ist.

5 Literaturverzeichnis

- EAWS, 2016: European Avalanche Warning Services, Gefahrenstufen, http://www.avalanches.org/eaws/en/includes/basics/basicsImages/scale_de.png, (04/2016)
- GEBETSBERGER, P., 2010: Winter3, 1. Auflage, Wien.
- KRONTHALER, G., 2011: Systematische Lawinenkunde. Ausbildungshandbuch der Tiroler Lawinenkommissionen, 2. Auflage, Amt der Tiroler Landesregierung (Hrsg.), Innsbruck.
- SCHWEIZER, J., 2006: Hangneigung. Zeitschrift bergundsteigen, Deutscher Alpenverein, Schweizer Alpen-Club SAC, Alpenverein Südtirol, Österreichischer Alpenverein (Hrsg.), Ausgabe 4/06, Innsbruck.
- WINKLER, K., BREHM, H. & HALTMEIER, J., 2008: Bergsport Winter – Technik, Taktik, Sicherheit, 2. Auflage, Bern.
- WÜRTL, W., 2011: Gefahrenzeichen im Gelände. Ausbildungshandbuch der Tiroler Lawinenkommissionen, 2. Auflage, Amt der Tiroler Landesregierung (Hrsg.), Innsbruck.
- ZENKE, B. & KRONTHALER, G., 2012: Strukturiertes Vorgehen in der Beurteilungstätigkeit örtlicher Lawinenkommissionen in Bayern. Conference Proceedings INTERPRAEVENT, Entwurf, Grenoble

Untersuchung und Modellierung des Rolling-Shutter-Effekts für photogrammetrische Einzel- und Mehrbildauswertungen

ROBIN ROFALLSKI¹ & THOMAS LUHMANN¹

Zusammenfassung: CMOS-Sensoren nutzen ein elektronisches Verschlussverfahren, welches das Bild sequentiell aufnimmt und ausliest. Dieses als „Rolling Shutter“ bezeichnete Prinzip führt in kinematischen Anwendungen zu charakteristischen Bildverzerrungen, die sich qualitätsmindernd auf photogrammetrische Messbilder auswirken.

Der vorliegende Beitrag behandelt das Thema des Rolling-Shutter-Effekts und führt eine Erweiterung der Kollinearitätsgleichungen mit zwei äußeren Orientierungen und geschwindigkeitskonstanter Interpolation ein. Das Modell kommt in Anwendungen der Einzel- und Mehrbildphotogrammetrie zum Einsatz, wird anhand eines Versuches im Nahbereich untersucht und die Ergebnisse mit Hilfe eines unabhängigen Messsystems verifiziert. Die Ergebnisse zeigen Genauigkeitssteigerungen von > Faktor 5 gegenüber klassischen Berechnungsverfahren.

1 Einleitung

CMOS-Sensoren sind in digitalen Spiegelreflexkameras mittlerweile als Standard etabliert und drängen zunehmend CCD-Sensoren aus dem Markt. Neben vielen Vorteilen hinsichtlich der Bildqualität, Komprimierbarkeit und Geschwindigkeit besitzt die CMOS-Technologie einen bauartbedingten Nachteil bei der elektronischen Verschlusssteuerung. Bei diesem als *Rolling Shutter* bezeichneten Verfahren werden die Sensorzeilen sequentiell belichtet und anschließend ausgelesen. Für kinematische Anwendungen, wie etwa der UAV-Photogrammetrie, resultieren dadurch Bildverzerrungen, welche als *Rolling-Shutter-Effekt* bezeichnet werden. Abhängig von Geschwindigkeit, Bildmaßstab und Auslesedauer des Sensors entstehen nichtlineare Bildverzerrungen, die von photogrammetrischen Berechnungsverfahren bislang nicht berücksichtigt werden.

In der Disziplin des Computer Vision ist die Thematik seit etwa zehn Jahren Forschungsgegenstand und es existieren Modelle, welche auf der projektiven Geometrie basieren und vorrangig für die Generierung einer möglichst dichten Punktwolke vorgesehen sind (siehe bspw. MEINGAST et al. 2005 oder AIT-AIDER et al. 2006). Obgleich einige Beiträge zum Thema veröffentlicht sind, ist eine Lücke im Bereich der Photogrammetrie mit signalisierten Punkten vorhanden. Diese soll mit Fokus auf Vergleichbarkeit und Genauigkeit durch die Validierung mit einem unabhängigen Messsystem geschlossen werden.

Der vorliegende Beitrag behandelt die Ergebnisse einer Masterarbeit, die zum Abschluss des Studiengangs *Geodäsie und Geoinformatik* an der Jade Hochschule Oldenburg vorgelegt wurde. Ziel war eine funktionale Modellierung der vorliegenden Problematik, Umsetzung in photogrammetrischen Berechnungsverfahren und deren Verifikation anhand eines geeigneten Versuchsaufbaus.

¹ Jade Hochschule Oldenburg, Institut für Angewandte Photogrammetrie und Geoinformatik, Ofener Straße 16/19, D-26121 Oldenburg, E-Mail: [robin.rofallski,thomas.luhmann]@jade-hs.de

2 Funktionale Modellierung

Klassisch sehen die Kollinearitätsgleichungen (Gl. 1) für jedes Bild genau eine äußere Orientierung mit sechs Parametern für Translation und Rotation vor (LUHMANN 2010). Dieses Modell wird als *Global-Shutter-Modell* (GS) bezeichnet, da es die Annahme trifft, dass die Bildaufnahme global – also zu genau einem Zeitpunkt – stattfindet.

$$\begin{aligned} x' &= x'_0 + z' \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta x' \\ y' &= y'_0 + z' \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta y' \end{aligned} \quad (1)$$

Aufgrund der zeilenweisen Belichtung gilt diese Annahme im Rolling-Shutter-Modell (RS) nicht mehr, weshalb eine Erweiterung des zentralperspektivischen Lochkameramodells vorgenommen wird. Für eine hypothesenfreie Beschreibung des Modells müsste für jede Sensorzeile eine äußere Orientierung geschätzt werden. Da dies in praktischen Anwendungen zu einem hochgradig unterbestimmten Gleichungssystem führt, wird das Modell auf zwei äußere Orientierungen an den Rändern beschränkt und dazwischen unter Annahme einer konstanten Geschwindigkeit interpoliert. Für die Abbildung des Rolling-Shutter-Effekts in den Kollinearitätsgleichungen werden der Verschiebungsvektor $\vec{V} = [V_x \ V_y \ V_z]^T$ für die Translation und zwei unabhängige Einheitsquaternionen q_1, q_2 für die Parametrisierung der Rotation eingeführt. Gesteuert wird die Interpolation in Abhängigkeit von der y' -Komponente der Bildkoordinate, indem der normierte Interpolationsfaktor τ unter Berücksichtigung der vertikalen Sensorgröße gebildet wird zu:

$$\tau = \frac{y'_{max} - y'_i}{y'_{max} - y'_{min}} \quad (2)$$

Die Interpolation geschieht für die Translation linear mit einer Steigungsfunktion und für die Rotation mit Hilfe von *Spherical Linear Interpolation* (SLERP), einem Verfahren zur Interpolation von Quaternionen unter konstanter Winkelgeschwindigkeit (SHOEMAKE 1985). Aus den interpolierten Quaternionen ergibt sich anschließend die interpolierte orthonormale Rotationsmatrix \mathbf{R}_{int} . Zusammengefasst und formalisiert in einer Erweiterung der Kollinearitätsgleichungen resultiert für die Anpassung an das RS-Modell folgender im Rahmen dieser Arbeit entwickelter funktionaler Zusammenhang:

$$\begin{aligned} x' &= x'_0 + z' \frac{r_{int}(\tau)_{11}(X - (X_0 + \tau V_x)) + r_{int}(\tau)_{21}(Y - (Y_0 + \tau V_y)) + r_{int}(\tau)_{31}(Z - (Z_0 + \tau V_z))}{r_{int}(\tau)_{13}(X - (X_0 + \tau V_x)) + r_{int}(\tau)_{23}(Y - (Y_0 + \tau V_y)) + r_{int}(\tau)_{33}(Z - (Z_0 + \tau V_z))} + \Delta x' \\ y' &= y'_0 + z' \frac{r_{int}(\tau)_{12}(X - (X_0 + \tau V_x)) + r_{int}(\tau)_{22}(Y - (Y_0 + \tau V_y)) + r_{int}(\tau)_{32}(Z - (Z_0 + \tau V_z))}{r_{int}(\tau)_{13}(X - (X_0 + \tau V_x)) + r_{int}(\tau)_{23}(Y - (Y_0 + \tau V_y)) + r_{int}(\tau)_{33}(Z - (Z_0 + \tau V_z))} + \Delta y' \end{aligned} \quad (3)$$

Offensichtlich geht τ in jeden Parameter der äußeren Orientierungen ein und steuert damit kontinuierlich die Veränderung ebenjener. Es ergeben sich damit 14 unbekannte Parameter mit zwei Bedingungsgleichungen für die Bestimmung der äußeren Orientierungen. Abb. 1 verdeutlicht das gegebene Modell anhand der beiden äußeren Orientierungen in farbiger Darstellung und zwei exemplarischen interpolierten Orientierungen in grauer Darstellung. Zu beachten ist hierbei in Grün die ausgelesene Zeile, welche mit dem Wert von τ bezogen auf die Bildhöhe korrespondiert.

Das funktionale Modell wird in den photogrammetrischen Berechnungsverfahren der Bestimmung eines räumlichen Rückwärtsschnittes und der Bündelausgleichung implementiert, wodurch die äußeren Orientierungen und Objektpunkte unter Annahme einer bewegten Kamera bestimmt werden können. Dazu wird eine eigene Bildkoordinatensimulation zu Test- und Simulationszwecken entwickelt, welche aufgrund der nichtlinearen Bewegungsform mittels Newton-Verfahren aufgelöst wird.

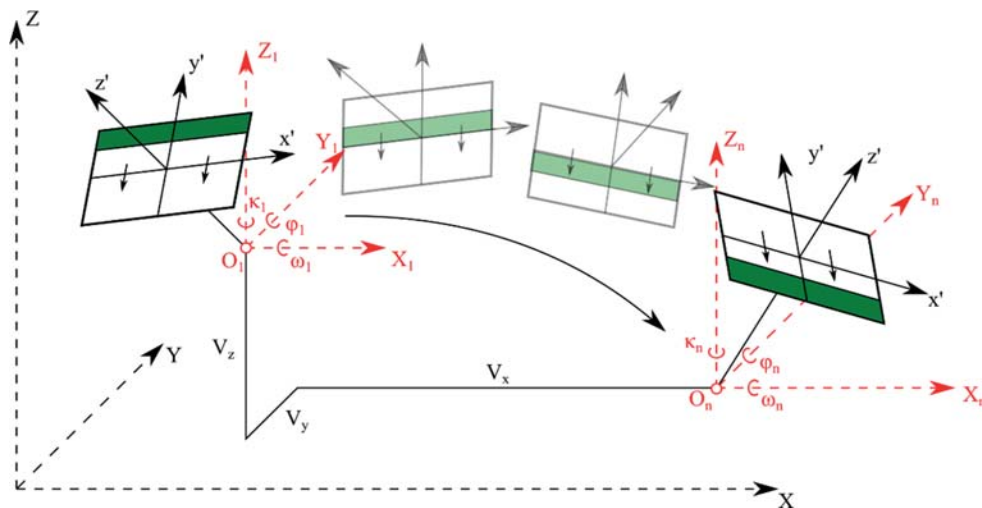


Abb. 1: Darstellung der beiden geschätzten und exemplarischen interpolierten äußeren Orientierungen

3 Verifikation des Modells

Zur Verifikation des dargestellten Modells wird ein Versuch mit einem rotierenden Messobjekt durchgeführt, welches von einer Nikon D4 im Videomodus (1920 x 1080 Px) beobachtet wird. Die Kamera besitzt einen CMOS-Sensor mit Rolling Shutter und ist damit zur Überprüfung des Modells geeignet. Die Bilddaten werden zum einen mittels Methoden der Einzelbilddatenauswertung und zum anderen mit einer Bündelausgleichung zur Berechnung von Objektpunkten ausgewertet. Vergleichend dazu werden die Daten ohne Berücksichtigung des Rolling-Shutter-Effekts mit Hilfe der Kollinearitätsgleichungen (GS-Modell) ausgewertet und die Ergebnisse gegenübergestellt.

3.1 Versuchsaufbau

Der Versuchsaufbau besteht im Wesentlichen aus:

- Nikon D4 mit 28 mm-Objektiv (vorab kalibriert)
- AICON MoveInspect HF4 für die Referenzmessung (vorab kalibriert)
- Bewegtes Messobjekt der Dimension 600 x 400 x 120 mm mit retroreflektierenden Targets und vier Hochpunkten (vorab eingemessen)
- Referenzrahmen zur geometrischen Transformation der Messsysteme (vorab eingemessen)

Vorab werden einige Simulationen durchgeführt, um eine bestmögliche Versuchsgeometrie zu gewährleisten. Die Ergebnisse zeigen, dass das Modell erst mit mindestens einem Hochpunkt lösbar wird. Weiterhin wird der Kamerastandort in einem Meter Entfernung von schräg unten auf das

Messobjekt blickend gewählt. Eine Abschätzung über die zu erwartende Bildwanderung durch Rolling-Shutter-Effekt wird auf Grundlage des Verfahrens von MEINGAST et al. (2005) durchgeführt. Hierbei wird die Auslesedauer jeder Bildzeile mit Hilfe einer LED, welche mit fester Frequenz ausgelöst wird, zu $14 \mu\text{s}/\text{Zeile}$ (entsprechend $15,5 \text{ ms}/\text{Bild}$) bestimmt. Die Belichtungsdauer wird auf $1/1000 \text{ s}$ reduziert, um den Rolling-Shutter-Effekt nicht durch Bewegungsunschärfe zu überlagern und die Rotationsgeschwindigkeit mit 30 RPM (entsprechend $1,1 \text{ m/s}$ am Rand) angepasst. Mit diesen Versuchsparametern ist eine Bildwanderung durch den Rolling-Shutter-Effekt von ca. 15 Px zu erwarten.

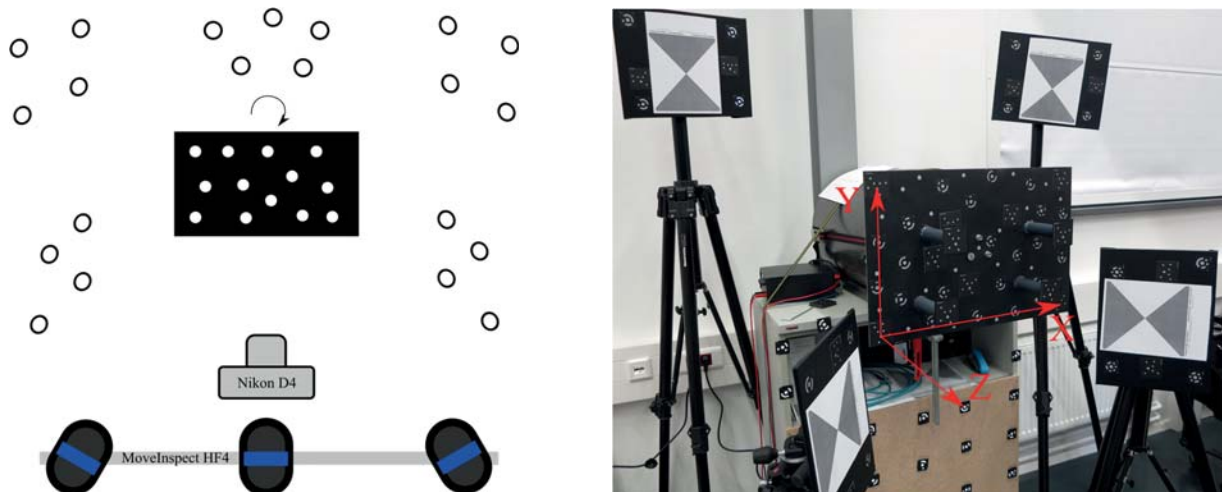


Abb. 2: Schematischer Versuchsaufbau (links) und Realisierung des Aufbaus mit Referenzrahmen und Messobjekt (rechts)

Zur unabhängigen Überprüfung der Ergebnisse aus Kameradaten wird das AICON MoveInspect HF4 als weiteres optisches 3D-Messsystem eingesetzt. Dies beobachtet die Rotation des Messobjekts sowie den statischen Referenzrahmen simultan und liefert Vergleichswerte mit einer Genauigkeit von $0,1 \text{ mm}$ im gewählten Versuchsvolumen (AICON 3D SYSTEMS GMBH 2016).

Abb. 2 zeigt den Versuchsaufbau, der von beiden Systemen über einen Zeitraum von 40 Sekunden beobachtet wird. Die Kamera nimmt im Videomodus mit einer Bildwiederholrate von 24 Hz auf (entspricht ca. 960 Bildern im Versuchszeitraum), während das MoveInspect mit 100 Hz etwa 4000 Messepochen aufzeichnet. Währenddessen vollführt das Messobjekt etwa 20 Umdrehungen, dessen Bewegung den Rolling-Shutter-Effekt auf dem Bildsensor bewirkt.

3.2 Einzelbildauswertung

Zur Realisierung eines gemeinsamen Punktes aus Kamera- und MoveInspect-Daten werden Rückwärtsschnitte aus dem Referenzrahmen und dem Messobjekt berechnet. Letzterer wird jeweils mit dem Global- und dem hier entwickelten Rolling-Shutter-Modell gebildet. Mittels 6DoF-Berechnung wird der Ursprung des Messobjekts (Lokatorpunkt) anschließend in das Referenzsystem transformiert. Die Parameter einer Transformation der vorab bestimmten Objektkoordinaten auf die MoveInspect-Daten liefern die Ursprungscoordinate auch aus diesem Messsystem.

Da die Kamera im Videomodus kein Triggersignal aussendet, kann das MoveInspect nicht ohne zusätzliche Komponenten synchronisiert werden. Daher wird mit dem Ansatz von LÖSLER & NITSCHKE (2010) aus den Lokatorpunkten des MoveInspects eine räumliche Ellipse im Gauß-Helmert-Modell geschätzt und diese als Referenz zugrunde gelegt. Der Abstand des Lokatorpunkts von der Ellipse ist die Zielgröße. Die benötigten Näherungswerte resultieren aus dem Global-Shutter-Modell, sodass für die Einzelbildauswertung im gegebenen Versuch keine Näherung für die Bewegungsparameter nötig ist.

Tab. 1: Ergebnisse der Einzelbildauswertung

	RMS (XYZ)	Max. Abweichung	Nicht konvergierte Bilder
Global Shutter	1,302 mm	6,072 mm	0,0 %
Rolling Shutter	0,246 mm	0,920 mm	2,7 %

Tabelle 1 unterstreicht dies mit zugehörigen statistischen Kennziffern. Die Ergebnisse sind um Faktor 5 genauer, als die Lösung im GS-Modell (s. auch Abb. 3). Einschränkend ist darauf hinzuweisen, dass etwa 3 % der Bilder zu keiner Lösung konvergieren. Dies resultiert vermutlich aus geometriebedingten Problemen, da das Messobjekt zum Zeitpunkt der jeweiligen Bildaufnahme nahezu horizontal ausgerichtet ist, was im gegebenen Modell scheinbar zu den Problemen führt.

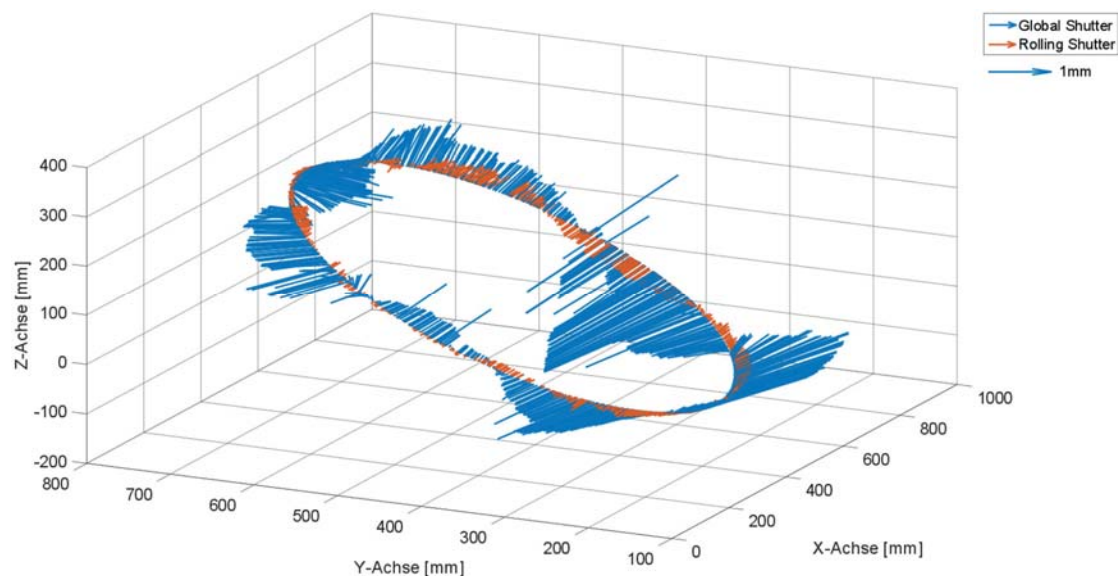


Abb. 3: Abweichungen des Lokatorpunkts von der Ellipse im Global- und Rolling-Shutter-Modell

3.3 Mehrbildauswertung

Die konvergent zum Messobjekt ausgerichtete Kameraposition sorgt für einen Rundumverband um das Messobjekt, der mit einer Bündelausgleichung ausgewertet werden kann. Die entstehenden 3D-Koordinaten des Messobjekts können anschließend mit den aus der Einmessung bekannten Punkten verglichen und die Restklaffen einer räumlichen Ähnlichkeitstransformation untersucht werden.

Tab. 2: Ergebnisse der Mehrbildauswertung

	RMS (XYZ)	Max. Abweichung
Global Shutter	0,864 mm	1,910 mm
Rolling Shutter	0,101 mm	0,183 mm

Die dargestellten Restklaffen in Abb. 4 zeigen deutlich größere Abweichungen im Global-Shutter-Modell und implizieren ebenfalls einen systematischen rotatorischen Effekt in den Hochpunkten, der sich gegenläufig zur Rotationsbewegung äußert. Die Ergebnisse des RS-Modells zeigen hingegen homogen verteilte Restklaffen, die sich vorrangig in Aufnahme­richtung (Z) widerspiegeln, betragsmäßig aber deutlich kleiner sind (siehe auch Tabelle 2). Damit sorgt auch in der Mehrbildauswertung das RS-Modell für genauere Ergebnisse von mehr als Faktor 8 gegenüber dem GS-Modell.

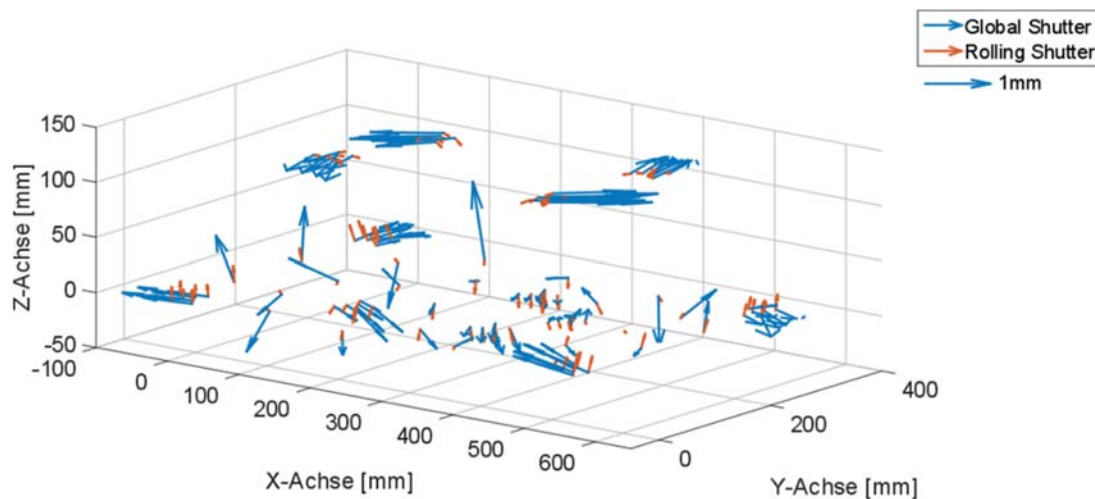


Abb. 4: Restklaffen der räumlichen Ähnlichkeitstransformation

4 Fazit und Ausblick

Die Ergebnisse zeigen deutliche Genauigkeitssteigerungen in beiden Auswertemethoden von mehr als Faktor 5. Das RS-Modell ist damit deutlich zutreffender, als das Global-Shutter-Modell und vermag im beschriebenen Versuch Genauigkeiten zu produzieren, die nahe der zu erwartenden Messgenauigkeit liegen. Die Restabweichungen können u. a. durch die Nichtberücksichtigung des Rolling-Shutter-Effekts bei der Markenmessung und der unvollkommenen Trajektorienbeschreibung durch eine Ellipse entstanden sein, was ggf. zu untersuchen ist.

Mit dem entwickelten Modell wird es möglich, die Daten, welche zuvor nur unter hohen Qualitätseinbußen genutzt werden konnten, mit ihren Verzerrungen zu berücksichtigen und Bildverbände mit hoher Genauigkeit auszuwerten. Insbesondere im Falle der UAV-Photogrammetrie kommen gewichts- und kostenbedingt häufig Kameras mit Rolling Shutter zum Einsatz, für die mit diesem Modell ein signifikanter Qualitätsgewinn geschaffen werden könnte. Auch ist hiermit die Nutzung von günstigen CMOS-Kameras mit Bildwiederholraten von mehr als 200 Hz, wie etwa der *GoPro Hero4 Black*, als kostengünstige Alternative zu High-Speed-Kameras künftig denkbar.

5 Literaturverzeichnis

- AICON 3D SYSTEMS GMBH, 2016: MoveInspect HF – Technische Daten. <http://aicon3d.de/produkte/moveinspect-technology/moveinspect-hf/technische-daten.html>, letzter Zugriff am 18.10.2016.
- AIT-AIDER, O., ANDREFF, N., LAVEST, J.M. & MARTINET, P., 2006: Exploiting Rolling Shutter-Distortions for Simultaneous Object Pose and Velocity Computation Using a Single View. International Conference on Computer Vision Systems, IEEE (Hrsg.), 35-41.
- LÖSLER, M. & NITSCHKE, M., 2010: Bestimmung der Parameter einer Regressionsellipse in allgemeiner Raumlage. Allgemeine Vermessungs-Nachrichten **3/2010**, 113-117.
- LUHMANN, T., 2010: Nahbereichsphotogrammetrie – Grundlagen, Methoden, Anwendungen. 3. Auflage, Wichmann VDE, Berlin/Offenbach.
- MEINGAST, M., GEYER, C. & SASTRY, S., 2005: Geometric models of rolling-shutter cameras. Computing Research Repository, arXiv preprint cs/0503076.
- SHOEMAKE, K., 1985: Animating Rotation with Quaternion Curves. SIGGRAPH '85: Proceedings of the 12th annual conference on Computer graphics and interactive techniques, 245-254.

Semantische Analyse von ALS- und MLS-Daten in urbanen Gebieten mittels Conditional Random Fields

MORITZ STEINSIEK¹, PRZEMYSŁAW POLEWSKI¹, WEI YAO¹ & PETER KRZYSTEK¹

Zusammenfassung: In dieser Arbeit wird ein Verfahren zur semantischen Analyse von ALS- und MLS-Daten in urbanen Gebieten entwickelt, das auf einem Conditional Random Field (CRF) zur kontextbasierten Klassifikation von 3D-Punktwolken aufbaut. Das CRF nutzt einen Random Forest (RF) Klassifikator für die Knotenpotentiale und eine Variante des kontrast-sensitiven Potts-Modells für die Kantenpotentiale. Die Grundlage der Klassifikation bilden verschiedene geometrische und radiometrische Merkmale, die mittels zweier lokaler Punktnachbarschaften abgeleitet werden. Die Genauigkeit des Verfahrens wird anhand von jeweils einem ALS- und einem MLS-Benchmark-Datensatz mit Referenzdaten überprüft. Zusätzlich erfolgt ein Vergleich der Ergebnisse der kontextbasierten Klassifikation mit einer individuellen Klassifikation mittels RF. Bei Anwendung dieses Verfahrens können Gesamtgenauigkeiten von 80,5% für den ALS-Datensatz und 97,7% für den MLS-Datensatz erzielt werden. Von der Nutzung kontextbasierter Klassifikationsverfahren profitieren im Vergleich zu individuellen Verfahren vor allem unterrepräsentierte Objektklassen deutlich, also solche Klassen mit einer geringen Anzahl von Punkten in den Daten.

1 Einleitung

Die automatische Interpretation von 3D-Punktwolken ist ein hochaktuelles Thema und gewinnt im Bereich der Fernerkundung, Photogrammetrie, Computer Vision und der Robotik zunehmend an Bedeutung. Gründe dafür sind unter anderem die in den letzten Jahren vollzogene Entwicklung von mobilen Laserscanning Systemen (MLS), die Weiterentwicklung des luftgestützten Laserscannings (ALS) und die steigende Verbreitung von *Unmanned Aerial Systems* zur Datenakquisition.

MLS werden im Allgemeinen für kleinräumige Gebiete, beispielsweise einzelne Straßenzüge, eingesetzt und liefern genaue und dichte 3D-Daten der Objektoberflächen. Die Punktdichte ist allerdings von der Entfernung des Scanners zu den einzelnen Objekten abhängig, was zu einer inhomogenen 3D-Punktwolke führt. Das hauptsächliche Einsatzgebiet für ALS ist im Gegensatz dazu die Aufnahme von großflächigen Umgebungen, wie zum Beispiel ganze Teile von Stadt- oder Waldgebieten mit weitaus geringeren aber meist homogenen Punktdichten.

Besonders im urbanen Raum finden sich viele Anwendungsbereiche, deren Grundlage 3D-Punktwolken darstellen. Darunter fallen etwa die Generierung von 3D-Stadtmodellen zur Planung von Infrastruktur und der Stadtentwicklung oder die Nutzung zur Fahrzeugnavigation. Ein grundlegender Schritt für die meisten der genannten Anwendungen ist die Durchführung einer Klassifikation der 3D-Punktwolke. Hierbei wird jedem einzelnen 3D-Punkt eine semantische Objektklasse (z.B. Auto, Gebäude, Vegetation oder Straße) zugeordnet.

¹ Hochschule München, Fakultät für Geoinformation, Karlstr. 6, D-80333 München,
E-Mail: moritz.steinsiek@gmx.de, [polewski, yao, krzystek]@hm.edu

Häufig konzentrieren sich die Untersuchungen auf einzelne Objektklassen wie Gebäude (ROTTENSTEINER et al. 2014), Vegetation (WURM et al. 2014) oder Straßenlaternen (YONGTAO et al. 2014). NIEMEYER et al. (2014) und WEINMANN et al. (2015) nutzen *Conditional Random Fields* (CRF) zur punktweisen Klassifikation von mehreren Objekten in urbanen ALS-beziehungswise MLS-Daten. Bei der Anwendung von CRF zeigt sich eine deutliche Verbesserung der Klassifikationsergebnisse im Vergleich zu herkömmlichen Verfahren, wie beispielsweise *Random Forest* (RF) oder *Support Vector Machines* (SVM). Durch das komplexere Modell der CRF erhöht sich jedoch der Berechnungsaufwand. YANG & FÖRSTNER (2011) und WEGNER et al. (2011) hingegen nutzen zur Klassifikation urbaner Gebiete CRF auf Basis von Rasterdaten statt 3D-Punktwolken.

Die häufigsten Schwierigkeiten bei der Klassifikation von Punktwolken ergeben sich vor allem aufgrund der großen Menge an Daten, Variationen in der Punktdichte, unvollständigen Strukturen durch Verdeckungen oder wegen der Unterschiede zwischen Objekten innerhalb derselben Klasse. Auf Grund von sehr komplexen Strukturen und vielen unterschiedlichen benachbarten Objektklassen, sind 3D-Szenen in urbanen Gebieten besonders herausfordernd. Das Ziel dieser Arbeit besteht darin, ein Verfahren zur semantischen Analyse von 3D-Punktwolken in urbanen Gebieten zu entwickeln, das sowohl für ALS- als auch MLS-Daten anwendbar ist. Die Analyse umfasst dabei alle nötigen Schritte, um jedem 3D-Punkt eine Objektklasse zuzuweisen: (i) die Auswahl einer passenden lokalen Punktnachbarschaft für die (ii) Extraktion von markanten Merkmalen und (iii) die Durchführung der Klassifikation mit einem geeigneten Verfahren.

2 Methodik

Das Ziel der Klassifikation von Punktwolken besteht darin, jedem 3D-Punkt P_i ein Objektklassenlabel zuzuweisen. In vielen Ansätzen werden individuelle Klassifikationsverfahren wie SVM (CORTESA & VAPNIK 1995) oder RF (BREIMAN 2001) eingesetzt. Bei diesen Verfahren wird jeder 3D-Punkt unabhängig von seiner Nachbarschaft klassifiziert. In der letzten Zeit wurden jedoch vor allem kontextbasierte Klassifikationsverfahren vorgestellt, bei denen auch die Objektklassen der Punkte in der Nachbarschaft in den Zuweisungsprozess der Labels miteinbezogen werden. Dies hat zur Folge, dass alle Punkte gleichzeitig klassifiziert werden müssen. In der vorliegenden Arbeit werden eine kontextberücksichtigende Klassifikation basierend auf CRF und eine individuelle Klassifikation mittels RF eingesetzt (Abb. 1).

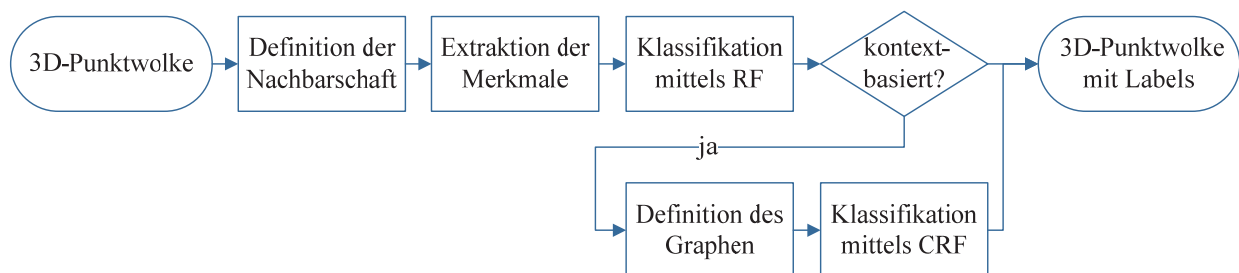


Abb. 1: Ablaufdiagramm mit den einzelnen Schritten zur Klassifikation der 3D-Punktwolke

2.1 Definition der lokalen Punktnachbarschaft

Ausgehend von einer Punktwolke mit N Punkten $P_i = (X_i, Y_i, Z_i)^T \in \mathbb{R}^3$ mit $i \in \{1, \dots, N\}$, besteht der erste Schritt darin, eine geeignete lokale Punktnachbarschaft festzulegen. Die Definition der Nachbarschaft dient zur Beschreibung der lokalen 3D-Struktur um einen gegebenen Punkt P_i und bildet damit die Grundlage für die Extraktion von geometrischen Merkmalen. In dieser Arbeit werden zwei Nachbarschaftsdefinitionen betrachtet: Sphärische Nachbarschaften N_s mit fixiertem Radius r_s und eine optimale sphärische Nachbarschaft N_{opt} mit variierendem Radius, basierend auf den $k \in \mathbb{N}$ nächsten Nachbarn und der Eigenentropie (WEINMANN et al. 2014).

Der Wert k für einen Punkt P_i hängt unter anderem von der lokalen 3D-Struktur und der lokalen Punktdichte um P_i ab. Infolgedessen basieren einige vorgeschlagene Verfahren zur automatischen Bestimmung von individuellen Werten für k auf den lokalen geometrischen Eigenschaften der Punktwolke. Zur Ableitung dieser Eigenschaften wird die 3D-Kovarianzmatrix $C \in \mathbb{R}^{3 \times 3}$ für einen gegebenen Punkt und seine k nächsten Nachbarn berechnet. Da es sich bei C um eine symmetrische, positiv-semidefinite Matrix handelt, existieren ihre drei Eigenwerte. Die Eigenwerte sind positiv und korrespondieren zu einem orthogonalen System von Eigenvektoren. Die drei Eigenwerte $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, mit $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ repräsentieren die Ausdehnung eines dreidimensionalen Kovarianzellipsoids entlang seiner Hauptachsen und eignen sich damit zur Beschreibung der lokalen 3D-Struktur. Die *Shannon Entropie* der normalisierten Eigenwerte e_1, e_2, e_3 dient als Grundlage für die Berechnung der Eigenentropie $E_{\lambda,i}$.

$$E_{\lambda,i} = -e_1 \ln(e_{1,i}) - e_2 \ln(e_{2,i}) - e_3 \ln(e_{3,i}) \quad (1)$$

Die Eigenentropie ist ein Maß für die Ordnung beziehungsweise die Unordnung der 3D-Punkte innerhalb der lokalen Nachbarschaft. Die Idee dieses Ansatzes besteht darin, den optimalen Wert von k zu finden, der die Unordnung der 3D-Punkte innerhalb der Nachbarschaft und damit die *Shannon Entropie* minimiert. Dafür wird der Wert k für jeden Punkt in einem festgelegten Intervall von k_{min} bis k_{max} mit einer Schrittweite von Δk variiert. Der Wert k , der zur minimalen *Shannon Entropie* korrespondiert, wird als optimaler Wert angenommen.

2.2 Extraktion von Merkmalen

Nach der erfolgten Definition der lokalen Punktnachbarschaft für jeden Punkt P_i , können anschließend Merkmale abgeleitet werden, welche die lokale 3D-Struktur beschreiben. In dieser Arbeit werden die Merkmale adaptiert, die in WEINMANN et al. (2015) vorgeschlagen wurden. Dabei handelt es sich um eine Kombination von 2D- und 3D-Merkmalen, mit der bereits gute Ergebnisse bei der Klassifikation von 3D-Punktwolken erzielt werden konnten.

Die 3D-Merkmale werden für jeden Punkt und seine durch die Nachbarschaft festgelegten Nachbarn abgeleitet. Basierend auf den drei Eigenwerten der 3D-Kovarianzmatrix, werden insgesamt acht geometrische Merkmale bestimmt:

$$L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_3} \quad L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_3} \quad (2)$$

$$\begin{aligned}
S_\lambda &= \frac{\lambda_3}{\lambda_{13}} & E_\lambda &= \sqrt[3]{\lambda_1 \lambda_2 \lambda_3} \\
A_\lambda &= \frac{\lambda_1 - \lambda_3}{\lambda_1} & E_\lambda &= - \sum_{i=1}^3 \lambda_i \ln(\lambda_i) \\
\Sigma_\lambda &= \lambda_1 + \lambda_2 + \lambda_3 & C_\lambda &= \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}
\end{aligned}$$

Die Linearität L_λ , Planarität P_λ und die Sphärizität S_λ liefern Informationen darüber, ob es sich um eine lineare 1D-Struktur, eine planare 2D-Struktur oder eine volumetrische 3D-Struktur handelt. Die weiteren eigenwertbasierten Merkmale sind die Omnivarianz O_λ , Anisotropie A_λ , Eigenentropie E_λ , die Summe der Eigenwerte Σ_λ und die Änderung der Krümmung C_λ . Als weiteres Merkmal ist die Vertikalität $V = 1 - |z_{\vec{n}}|$ zu nennen, die sich mit Hilfe der dritten Komponente $z_{\vec{n}}$ des Normalenvektors \vec{n} berechnen lässt. Zusätzlich werden der Radius r_{knn} der Kugel, die durch die k nächsten Nachbarn gebildet wird, und die resultierende lokale Punktdichte D_{3D} bestimmt.

$$D_{3D} = \frac{k + 1}{3/4 \pi r_{knn}^3} \quad (3)$$

Weitere markante Merkmale können mit Hilfe der Höhe abgeleitet werden. Eine wichtige Eigenschaft stellt die normalisierte Höhe H_{norm} dar, welche der über einem digitalen Geländemodell entspricht. Für die ALS-Daten wird die normalisierte Höhe des Punktes P_i als Merkmal verwendet, da die Höhe in größeren Gebieten stärker variiert. Im Falle von MLS-Daten mit kleinräumiger Ausdehnung wird dagegen die absolute Höhe H_{abs} verwendet. Weiter dienen die Standardabweichung der Höhe σ_H und die maximale Höhendifferenz ΔH innerhalb der lokalen Nachbarschaft als Merkmale. Beim ALS-Datensatz wird darüber hinaus für jeden Punkt der Intensitätswert I_i genutzt.

Zur Ableitung der 2D-Merkmale werden zunächst alle 3D-Punkte auf die XY-Ebene projiziert. Beruhend auf dieser 2D-Darstellung der Daten, wird der Radius $r_{knn,2D}$ der kreisförmigen Nachbarschaft, die sich aus P_i und seinen k Nachbarn ergibt, bestimmt. Mit $r_{knn,2D}$ kann anschließend die lokale Punktdichte D_{2D} berechnet werden.

$$D_{2D} = \frac{k + 1}{\pi r_{knn,2D}^2} \quad (4)$$

Überdies wird die 2D-Kovarianzmatrix C_{2D} berechnet und daraus die beiden Eigenwerte $\lambda_{1,2D}$ und $\lambda_{2,2D}$ abgeleitet. Auf Basis der beiden Eigenwerte werden die Summe $\Sigma_{\lambda,2D}$ und das Verhältnis $R_{\lambda,2D} = \lambda_{2,2D} / \lambda_{1,2D}$ gebildet und als weitere Merkmale genutzt.

Weiterhin werden 2D-Merkmale aus einem Akkumulationsbild extrahiert, indem die Projektionsebene in diskrete, quadratische Bins mit einer Seitenlänge von s unterteilt wird. Für jeden Punkt werden aus dem ihm zugewiesenen Bin folgende Merkmale abgeleitet: Die Anzahl der Punkte N_{Acc} in dem Bin, die maximale Höhendifferenz ΔH_{Acc} und die Standardabweichung der Höhe $\sigma_{H,Acc}$.

2.3 Conditional Random Field

Die CRF gehören zu den probabilistischen graphischen Modellen und erlauben die Modellierung von Abhängigkeiten zwischen benachbarten Punkten. Die CRF sind diskriminative Modelle und modellieren die bedingte Klassenwahrscheinlichkeit $P(Y|x)$ direkt (KUMAR & HEBERT 2006):

$$P(Y|x) = \frac{1}{Z(x)} \exp \left(\prod_{i \in n} \phi(x, y_i) \prod_{i \in n} \prod_{j \in N_i} \psi(x, y_i, y_j) \right) \quad (5)$$

$Z(x)$ entspricht einer Normalisierungskonstanten und wird auch als Partitionsfunktion bezeichnet. Da diese nicht von den Klassenlabels abhängt, kann sie in Bezug auf die Klassifikation vernachlässigt werden. Die Gleichung besteht somit aus zwei Funktionen: Dem Assoziationspotential $\phi(x, y_i)$ und dem Interaktionspotential $\psi(x, y_i, y_j)$. Das CRF wird durch einen ungerichteten Graphen $G = (n, e)$ mit einer Menge von Knoten n und einer Menge von Kanten e repräsentiert. Die Knoten $n_i \in n$ entsprechen den 3D-Punkten der Punktwolke; die Kanten $e_{ij} \in e$ verbinden zwei benachbarte Knoten n_i und n_j und modellieren deren Abhängigkeiten. Innerhalb einer 3D-Punktwolke ist die Definition des Graphen in Bezug auf die lokale Nachbarschaft komplexer als für Bilder. Die 3D-Punkte innerhalb einer Punktwolke haben eine unregelmäßige Anordnung im dreidimensionalen Raum. In Bildern können hier beispielsweise die vier Nachbarn anhand des Bildrasters ausgewählt werden, was bei Punktwolken nicht möglich ist. Häufig basiert die Definition der lokalen Punktnachbarschaft auf den k nächsten Nachbarn. Dementsprechend erfolgt die Definition der Kanten mittels den in Abschnitt 2.1 beschriebenen Nachbarschaften, die bereits zur Extraktion der Merkmale genutzt werden. Angesichts der hohen Anzahl an Punkten in den Punktwolken, wird die Größe der lokalen Nachbarschaft innerhalb des Graphen aus Gründen der Rechenkapazitäten durch einen Schwellwert k_{thr} eingeschränkt. Der Vektor Y enthält die Klassenlabels y_i für jeden 3D-Punkt. x entspricht den Zufallsvariablen der Merkmale. Das Ziel der Klassifikation besteht darin, die optimale Konfiguration Y zu finden, für die $P(Y|x)$ maximal wird.

2.3.1 Definition der Potentiale

Das Assoziationspotential $\phi(x, y_i)$ verbindet die Daten x und die entsprechenden Klassenlabels y_i . Für dieses Potential kann das Ergebnis jedes beliebigen diskriminativen Klassifikators, beispielsweise RF, verwendet werden. Der Klassifikator muss in der Lage sein, eine Verteilung über die möglichen Werte der Klassenlabels zu liefern. Diese Möglichkeit ergibt sich daraus, dass $\phi(x, y_i)$ für einen Knoten n_i von den gesamten Daten x abhängen kann (KUMAR & HEBERT 2006). Infolgedessen wird für jeden Knoten ein Merkmalsvektor $f_i(x)$ auf Basis der in Abschnitt 2.2 beschriebenen Merkmale erstellt. Die einzelnen Komponenten des Merkmalsvektors sind folglich Funktionen der lokalen Nachbarschaft des Knotens. Das Assoziationspotential kann als die a posteriori Wahrscheinlichkeit eines lokalen diskriminativen Klassifikators beruhend auf $f_i(x)$ definiert werden (KUMAR & HEBERT 2006): $\phi(x, y_i) = \log(p(y_i|f_i(x)))$.

Für das Assoziationspotential $\phi(x, y_i)$ wird ein RF verwendet. Bei RF wird ein Ensemble von N_T unabhängigen Entscheidungsbäumen erstellt, die, basierend auf dem *Bagging* Prinzip, mit unterschiedlichen Trainingsmengen trainiert werden. Die anschließende Klassifikation wird mittels einer Mehrheitsentscheidung getroffen, bei der jeder Baum für eine Klasse y^l stimmt. RF

zählt zu den nicht-probabilistischen Klassifikationsverfahren, die nicht direkt auf Wahrscheinlichkeiten beruhen. Mittels der Anzahl N_l der Bäume, die für eine Klasse y^l gestimmt haben, kann allerdings im Anschluss die a posteriori Wahrscheinlichkeit abgeleitet werden:

$$p(y_i = y^l | f_i(x)) = \frac{N_l}{N_T} \quad (5)$$

Das Interaktionspotential $\psi(x, y_i, y_j)$ beschreibt den lokalen Kontext und verbindet die Klassenlabels (y_i, y_j) zweier benachbarter Knoten mit den Daten x . Für das Interaktionspotential kann ebenfalls das Ergebnis jedes beliebigen diskriminativen Klassifikators verwendet werden. In dieser Arbeit wird für dieses Potential eine Variante des häufig verwendeten kontrast-sensitiven Potts-Modells (BOYKOV & JOLLY 2001) angewendet (WEINMANN et al. 2015):

$$\log(\psi(x, y_i, y_j)) = \delta_{y_i y_j} \cdot w_1 \cdot \frac{N_a}{N_{k_i}} \left[w_2 + (1 - w_2) \cdot \exp\left(-\frac{d_{ij}(x)^2}{2\sigma^2}\right) \right] \quad (6)$$

$\delta_{y_i y_j}$ entspricht dem Kronecker-Delta, welches den Wert eins, falls $y_i = y_j$ und ansonsten null annimmt. $d_{ij}(x)$ ist die Euklidische Distanz zwischen den beiden Merkmalsvektoren $f_i(x)$ des Knoten i und $f_j(x)$ des Knoten j , die durch eine gemeinsame Kante e_{ij} verbunden sind. Der Wert N_a ist die durchschnittliche Anzahl der Kanten, mit denen ein Knoten im Graphen verbunden ist und N_{k_i} die Anzahl der Knoten in der lokalen Nachbarschaft des Knoten i . Der Parameter σ ist die durchschnittliche Distanz zwischen den Merkmalsvektoren benachbarter Punkte. Bei den beiden Parametern w_1 und w_2 handelt es sich um Gewichtsparameter. Das erste Gewicht bestimmt, welchen Einfluss das Interaktionspotential auf die Klassifikation hat und kann beliebige nichtnegative, reelle Werte annehmen. Das zweite Gewicht $w_2 \in [0; 1]$ beeinflusst den Grad der Glättung abhängig von den Daten x .

2.3.2 Training und Inferenz

Zunächst müssen die entsprechenden Parameter und Gewichte der unterschiedlichen Klassifikatoren in einem Trainingsprozess bestimmt werden. Dazu werden Trainingsmengen beziehungsweise Referenzdaten der 3D-Punkte mit den entsprechenden Objektklassenlabels benötigt. Für den verwendeten RF müssen im Trainingsschritt die N_T Entscheidungsbäume angelehrt werden. Dazu wird eine Trainingsmenge T mit N_{tr} zufällig ausgewählten 3D-Punkten pro Klasse erzeugt. Durch den Ausgleich der Klassen wird vermieden, dass Entscheidungsbäume entstehen, die mit nur sehr wenigen oder keinen Beispielen für eine unterrepräsentierte Klasse trainiert werden.

Für die Bestimmung der beiden Gewichte w_1 und w_2 kann beispielsweise eine Methode wie die Kreuzvalidierung verwendet werden. Die beiden Parameter werden in dieser Arbeit mittels einer Gittersuche unter Verwendung des gesamten Trainingsdatensatzes ausgewählt. Bei der Klassifikation mittels CRF wird die optimale Belegung der Labels Y bestimmt, für welche $P(Y|x)$ maximal wird. Für Graphen mit Zyklen ist die exakte Inferenz nicht in praktischer Zeit lösbar, weshalb im Allgemeinen nur approximative Lösungen verwendet werden. In dieser Arbeit wird die *Quadratic Pseudo-Boolean Optimization* (BOROS & HAMMER 2002) genutzt.

3 Experimente und Ergebnisse

3.1 Datensätze

Die Genauigkeit der Methode wird anhand von jeweils einem ALS- und einem MLS-Benchmark-Datensatz mit Referenzdaten überprüft.

Der ALS-Datensatz (NIEMEYER et al. 2014) wird im Rahmen des *3D Semantic Labeling Contest* der ISPRS zur Verfügung gestellt.² Dieser wurde über dem Stadtgebiet von Vaihingen in Deutschland aufgenommen. Die Aufnahme erfolgte im August 2008 mit einem *Leica ALS50* System in einer gemittelten Flughöhe von 500 m über dem Boden und einem Sichtfeld von 45°. Der durchschnittliche Überlappungsbereich der Spuren beträgt 30%, die Punktdichte liegt bei ungefähr 8 Punkte/m². Zusätzlich wurden Intensitäten und Echoinformationen aufgezeichnet. Der Datensatz wurde in einen Trainingsdatensatz mit ungefähr 750k Punkten und einen davon unabhängigen Testdatensatz mit ungefähr 412k Punkten aufgeteilt. Der Trainingsdatensatz ist vollständig mit den entsprechenden Referenzlabels versehen. Das Ziel des *3D Semantic Labeling Contest* ist es, jedem 3D-Punkt im Testdatensatz eine von neun Klassen zuzuordnen: *Hochspannungsleitung* (0,07%), *niedrige Vegetation* (23,9%), *undurchlässige Oberfläche* (25,7%), *Auto* (0,61%), *Zaun/Hecke* (1,60%), *Dach* (20,2%), *Fassade* (3,61%), *Strauch* (6,31%) und *Baum* (17,9%).

Bei dem MLS-Datensatz handelt es sich um das *Oakland 3D Point Cloud Dataset* (MUNOZ et al. 2009) mit den fünf Objektklassen *Hochspannungsleitung* (0,3%), *Stange/Mast* (0,6%), *Fassade* (8,4%), *Boden* (70,5%) und *Vegetation* (20,2%).³ Der Datensatz wurde in der Umgebung des Campus der *Carnegie Mellon University* (CMU) in Oakland, USA aufgenommen. Die Daten wurden mit Hilfe des *Mobile Mapping Systems NavlabII* der CMU aufgenommen. Nach der Aufnahme erfolgte eine Zuordnung der entsprechenden Referenzlabels jedes 3D-Punktes in einem semi-automatischen Prozess. Die Daten liegen aufgeteilt in einem Trainingsdatensatz mit ungefähr 37k Punkten, einem Validationsdatensatz mit 91k Punkten und mehreren Testdatensätzen mit insgesamt ungefähr 1,3M Punkten, vor.

3.2 Experimente

Die Experimente werden mit den beiden in Abschnitt 2.1 beschriebenen lokalen Nachbarschaften durchgeführt. Für die optimale Nachbarschaft N_{opt} wird das Intervall für k durch $k_{min} = 10$ und $k_{max} = 100$ mit einer Schrittweite von $\Delta k = 1$ definiert. Für die sphärische Nachbarschaft N_s werden für den ALS-Datensatz die Radien $r_s = (0,5m; 1,0m; 2,0m)$ und für den MLS-Datensatz $r_s = (0,4m; 0,6m; 0,8m)$ untersucht.

Die Extraktion der Merkmale (vgl. Abschnitt 2.2) erfolgt basierend auf den k nächsten Nachbarn, die sich aus der entsprechenden Nachbarschaftsdefinition ergeben. Die Seitenlängen s der quadratischen Bins für die Merkmale N_{Acc} , ΔH_{Acc} und $\sigma_{H,Acc}$ wurde empirisch bestimmt und im Falle der ALS-Daten auf $s = 1,25m$ und für die MLS-Daten auf $s = 0,25m$ gesetzt.

²Der Datensatz kann unter <http://www2.isprs.org/commissions/comm3/wg4/data-request-form2.html> angefordert werden (Letzter Aufruf: 06.01.2017).

³Der Datensatz ist frei verfügbar unter http://www.cs.cmu.edu/~vmr/datasets/oakland_3d/cvpr09/doc/ (Letzter Aufruf: 06.01.2017)

Für alle Experimente wird für den RF eine gleichbleibende Trainingsmenge T mit jeweils $N_{tr} = 1000$ zufällig ausgewählten Beispielen pro Klasse erzeugt. Bei einer Erhöhung des Wertes, konnten keine signifikant besseren Klassifikationsergebnisse erzielt werden. Daraus ergibt sich für die ALS-Daten mit neun Klassen eine Trainingsmenge T mit $|T| = 9000$ und für die MLS-Daten mit fünf Klassen eine Trainingsmenge mit $|T| = 5000$. Die Anzahl der Entscheidungsbäume wurde empirisch bestimmt und auf $N_T = 100$ festgelegt.

Die Bestimmung der beiden Gewichtsparameter w_1 und w_2 des Potts-Modells erfolgt mittels einer Gittersuche auf dem gesamten Trainingsdatensatz. Hierzu werden die beiden Parameter jeweils in einem Intervall variiert und das CRF mit allen sich daraus ergebenden Kombinationen trainiert. Anhand der Trainingsgenauigkeiten wird daraufhin die beste Kombination ausgewählt. Dabei ergibt sich für den ALS-Datensatz $w_1 = 0,6$ und $w_2 = 0$ und für den MLS-Datensatz $w_1 = 5$ und $w_2 = 0,5$. Der Schwellwert für die maximale Nachbarschaftsgröße innerhalb des Graphen wird auf $k_{thr} = 25$ gesetzt.

3.3 Ergebnisse und Diskussion

Die Tab. 1 zeigt jeweils ein Ergebnis der Klassifikation mit RF und CRF für den ALS-Datensatz mit der Gesamtgenauigkeit (O_A), Cohens Kappa Koeffizient κ und dem $F1$ -Score für jede Klasse. Der $F1$ -Score ist definiert als der harmonische Mittelwert aus Korrektheit und Vollständigkeit. Für beide Klassifikatoren wurde die Nachbarschaft N_{opt} genutzt.

Tab. 1: Gesamtgenauigkeit, Cohens-Kappa und F1-Score der einzelnen Klassen für den ALS-Datensatz.

	O_A	κ	Hoch.	n. Veg.	Un. Ober.	Auto	Zaun	Dach	Fass.	Strauch	Baum
RF	71,0	0,71	14,3	65,8	86,1	24,9	19,8	84,8	43,9	40,8	69,5
CRF	80,5	0,75	69,8	73,8	91,5	58,2	29,9	91,6	54,7	47,8	80,2

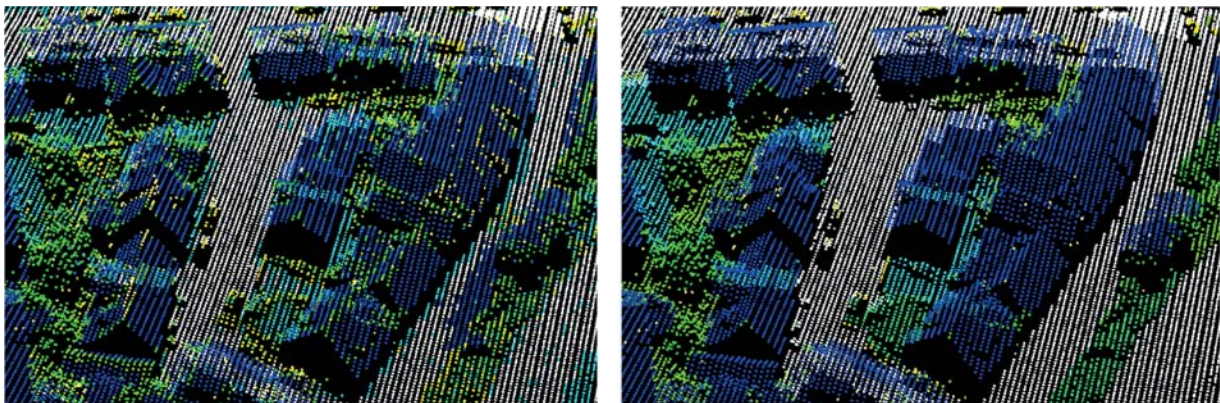


Abb. 2: Vergleich der Ergebnisse von RF (links) und CRF (rechts) anhand eines Ausschnitts der ALS-Daten.

Der Unterschied der beiden Klassifikationsverfahren in der Gesamtgenauigkeit beträgt 9 Prozentpunkte. Die Differenzen im $F1$ -Score der einzelnen Klassen liegen jedoch zwischen 5 und 55 Prozentpunkten. Besonders die unterrepräsentierten Klassen *Hochspannungsleitung* und *Auto*, die lediglich einen Anteil von 0,07% und 0,6% bezogen auf die Gesamtpunktzahl aufweisen, zeigen Zunahmen von 55 beziehungsweise 33 Prozentpunkten im $F1$ -Score. Bei beiden Verfahren ist auffällig, dass vor allem die Klassen *niedrige Vegetation*, *Strauch* und

Zaun/Hecke schlecht voneinander getrennt werden können. In Abb. 2 werden die beiden Klassifikationsergebnisse gegenüber gestellt. Dabei ist besonders bei der Klasse *Dach* (blau) der Einfluss der kontextbasierten Klassifikation sichtbar.

In Tab. 2 sind die Ergebnisse der Klassifikation mit RF und CRF für den MLS-Datensatz dargestellt. Die besten Resultate konnten mit einem RF mit der Nachbarschaft N_{opt} und mit einem CRF mit einer sphärischen Nachbarschaft mit fixiertem Radius $r_s = 0,8\text{m}$ erzielt werden.

Tab. 2: Gesamtgenauigkeit, Cohens-Kappa und F1-Score der einzelnen Klassen für den MLS-Datensatz.

	O_A	κ	Hoch.	Stange	Fass.	Boden	Veg.
RF	92,0	0,83	12,1	32,6	72,9	98,8	87,9
CRF	97,7	0,95	33,0	70,0	94,3	99,5	96,1

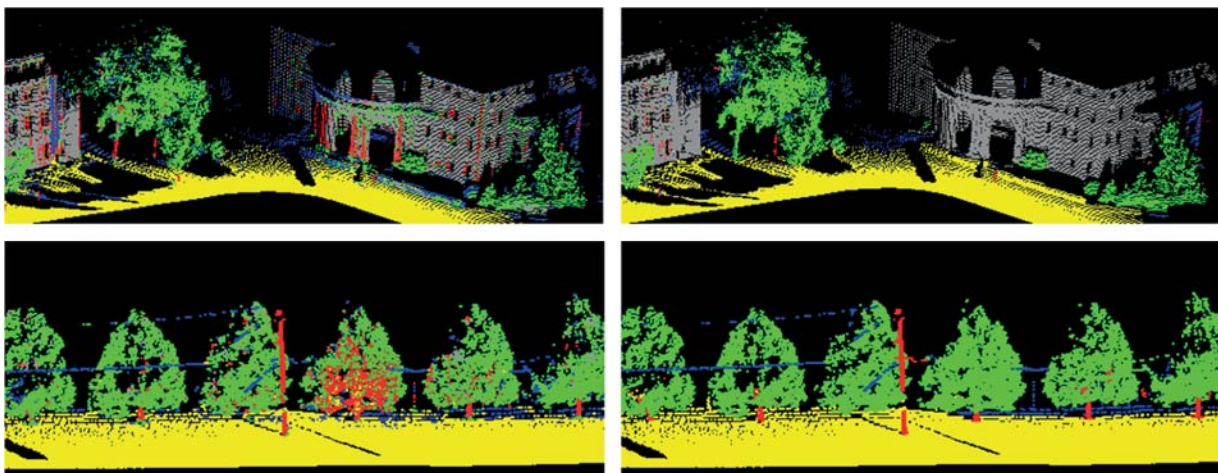


Abb. 3: Vergleich der Ergebnisse von RF (links) und CRF (rechts) anhand eines Ausschnitts der MLS-Daten.

Die hohen Gesamtgenauigkeiten von 92% und 97,7% werden durch die sehr gut klassifizierte Objektklasse *Boden* stark beeinflusst, die 70,5% der Punkte im Testdatensatz ausmacht. Bei beiden Verfahren werden die Klassen *Hochspannungsleitung* und *Stange/Mast* mit nur einer geringen Anzahl an Punkten weniger gut erkannt. Im Vergleich mit RF, können mit Hilfe des CRF vor allem für solche Klassen höhere Werte für den *F1-Score* erreicht werden. Insgesamt ist das CRF-Klassifikationsergebnis bei den MLS-Daten, wie bei den ALS-Daten, glatter als bei RF. Ein Vergleich anhand zweier Ausschnitte der klassifizierten Testdaten ist in Abb. 3 dargestellt. Besonders sichtbar ist dieser Glättungseffekt bei den Klassen *Fassade* (grau) und *Vegetation* (grün).

Zum Vergleich des vorgestellten Verfahrens eignen sich für die ALS-Daten die auf der Seite des *ISPRS Semantic Labeling Contest*⁴ eingereichten Ergebnisse. RAMIYA et al. (2016) verwenden für die Klassifikation des Datensatzes eine 3D-Segmentierung und extrahieren aus den Segmenten spektrale und geometrische Merkmale. Hierfür wurde neben der 3D-Punktwolke auch ein *True Orthophoto* verwendet. Ein weiteres Ergebnis wurde von CVIRN et al. (2016)

⁴ Die eingereichten Ergebnisse des ISPRS Semantic Labeling Contest sind unter <http://www2.isprs.org/vaihingen-3d-semantic-labeling.html> abrufbar.

eingereicht. Zur Klassifikation wird *Multiclass Machine Learning* mit einer *one-vs-one* Strategie, verbunden mit einem genetischen Algorithmus zur Merkmalsauswahl, angewendet. Als Merkmale werden radiometrische, strukturelle und geometrische Attribute genutzt. Das Verfahren von RAMIYA et al. (2016) resultiert in eine Gesamtgenauigkeit von 76,2%. Allerdings wurde die Objektklasse *Fassade* nicht klassifiziert, was einen negativen Einfluss auf das Ergebnis hat. Mit dem in dieser Arbeit verwendeten Verfahren kann, im Vergleich zu RAMIYA et al. (2016), bei den einzelnen Objektklassen ein etwas besserer *F1-Score* erzielt werden. Die Unterschiede liegen zwischen 0,3 (*Auto*) und 15,4 Prozentpunkten (*Hochspannungsleitung*). CVIRN et al. (2016) erzielten eine Gesamtgenauigkeit von 80,8% für den Datensatz. Auch bei diesem Ansatz werden die unterrepräsentierten Klassen *Hochspannungsleitung*, *Auto* und *Zaun/Hecke*, gemessen am *F1-Score*, schlechter detektiert. Die Klassen *niedrige Vegetation* und *Dach* werden besser erkannt.

Für den MLS-Datensatz werden im Rahmen eines Vergleichs die Ergebnisse von MUNOZ et al. (2009) herangezogen. Die kontextbasierte Klassifikation der Punktwolke geschieht mittels *Functional Max-Margin Markov Networks*. Der Vergleich anhand des *F1-Scores* zeigt lediglich größere Unterschiede bei den Klassen *Hochspannungsleitung* und *Stange/Mast*. Während mit dem in dieser Arbeit verwendeten CRF die Klasse *Stange/Mast* einen um 30,6 Prozentpunkte höheren *F1-Score* aufweist, ist dieser für die Klasse *Hochspannungsleitung* um 31,3 Prozentpunkte geringer.

4 Fazit und Ausblick

Das in dieser Arbeit verwendete Verfahren liefert sowohl für den ALS-, als auch für den MLS-Datensatz gute Ergebnisse. Die Anpassungen bezüglich der zwei unterschiedlichen Aufnahmemethoden beschränken sich lediglich auf die Optimierung der beiden Gewichtsparameter. Durch das Einbeziehen von Kontextinformationen kann das Klassifikationsergebnis hinsichtlich der individuellen Klassifikation in allen Belangen verbessert werden. Besonders unterrepräsentierte Klassen mit wenigen Punkten im Datensatz profitieren von der kontextbasierten Klassifikation.

Eine weiterführende Entwicklung des vorgestellten Verfahrens könnte bei der Verbesserung der Klassifikationsgenauigkeit ansetzen, indem weitere Merkmale und Nachbarschaften auf ihre Eignung untersucht würden. Zusätzlich könnte die Übertragbarkeit des trainierten Klassifikators auf unterschiedliche Gebiete und Sensoren überprüft werden. Abschließend könnten die Ergebnisse genutzt werden, um einzelne Objekte wie zum Beispiel Bäume zu extrahieren.

5 Literaturverzeichnis

- BOROS, E. & HAMMER, P. L., 2002: Pseudo-boolean optimization. *Discrete Applied Mathematics* **123**(1-3), 155-225.
- BOYKOV, Y. Y. & JOLLY, M.-P., 2001: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. *Proceedings of the Eighth IEEE International Conference on Computer Vision* **1**, 105-112.
- BREIMAN, L., 2001: Random Forests. *Machine Learning* **45**(1), 5-32.

- CORTESA, C. & VAPNIK, V., 1995: Support-Vector Networks. *Machine Learning* **20**(3), 273-297.
- CVIRN, R.; HORVAT, D. & MONGUS, D., 2016. http://ftp.ipi.uni-hannover.de/ISPRS_WGIII_website/ISPRSIII_4_Test_results/papers/UM_DescriptionOfAlgorithms.pdf, letzter Zugriff 20.01.2017.
- KUMAR, S. & HEBERT, M., 2006: Discriminative Random Fields. *International Journal of Computer Vision* **68**(2), 179-201.
- MUNOZ, D.; BAGNELL, J. A.; VANDAPEL, N. & HEBERT, M., 2009: Contextual Classification with Functional Max-Margin Markov Networks. *IEEE Conference on Computer Vision and Pattern Recognition 2009*, 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.
- RAMIYA, A. M.; NIDAMANURI, R. R. & KRISHNAN, R., 2016: 3D semantic labelling of colored LiDAR point cloud using supervoxels based segmentation approach/color based region growing segmentation approach. http://ftp.ipi.uni-hannover.de/ISPRS_WGIII_website/ISPRSIII_4_Test_results/papers/IIST_3D_ext.pdf, letzter Zugriff 20.01.2017.
- ROTTENSTEINER, F.; SOHN, G.; JUNG, J.; GERKE, M.; BAILLARD, C.; BENITEZ, S. & BREITKOPF, U., 2014: Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing* **93**, 256-271.
- WEGNER, J. D.; HÄNSCH, R.; THIELE, A. & SOERGEL, U., 2011: Building Detection From One Orthophoto and High-Resolution InSAR Data Using Conditional Random Fields. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**(1), 83-91.
- WEINMANN, M.; JUTZI, B. & MALLET, C., 2014: Semantic 3D scene interpretation: a framework combining optimal neighborhood size selection with relevant features. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **II-3**, 181-188.
- WEINMANN, M.; SCHMIDT, A. & MALLET, C., 2015: Contextual classification of point cloud data by exploiting individual 3D neighborhoods. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **II-3/W4**, 271-278.
- WURM, K. M.; KRETZSCHMAR, H.; KÜMMERLE, R.; STACHNISS, C. & BURGARD, W., 2014: Identifying vegetation from laser data in structured outdoor environments. *Robotics and Autonomous Systems* **62**(5), 675-684.
- YANG, M. Y. & FÖRSTNER, W., 2011: A hierarchical conditional random field model for labeling and classifying images of man-made scenes. *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 169-203.
- YONGTAO, Y.; JONATHAN, L.; HAIYAN, G. & CHENG, W., 2015: Automated Extraction of Urban Road Facilities Using Mobile Laser Scanning Data. *IEEE Transactions on intelligent transportation systems* **16**(4), 2167-2181.

AUTORENVERZEICHNIS

A

ARENS, MICHAEL 262, 302

B

BAHR, THOMAS 223

BALTRUSCH, SVEN 90, 280

BAUMHAUER ROLAND 363

BECKER, SUSANNE 147

BOERNER, RICHARD 371

BORGMANN, BJÖRN 262

BRAUN, ANDREAS 110

BRAUNECK, JENS 45

BRUNN, ANDREAS 276

BRUNN, ANSGAR 434

BURGGRAAFF, PETER 203

C

CHIZHOVA, MARIA 434

CHRISTEN, MARTIN 67

CONJO, ANGELO 383

D

DEGGIM, SIMON 454

DINKEL, AGNES 428

DORNER, WOLFGANG 403

DOSTÁLEK, LIBOR 403

DRAUSCHKE, MARTIN 122

E

EDER, KONRAD 464

F

FEHR, MARKUS 67

FELDHAUS, HERIBERT 45

FISCHER, JÜRGEN 482

FRITSCH, DIETER 147

G

GEHRUNG, JOACHIM 302

GRENZDÖRFFER, GÖRRES 1

GÜLCH, EBERHARD 422

GURIANOV, ANDREI 434

H

HAALA, NORBERT 147

HANEL, ALEXANDER 160

HARLANDER, MAX 22

HAUNERT, JAN-HENRIK 138

HEBEL, MARCUS 262, 302

HEIPKE, CHRISTIAN 245

HINGER, JÜRGEN 22

HIJAZI, IHAB.....	288
HINZ, STEFAN.....	110
HOBERG, THORSTEN.....	245
HOEGNER, LUDWIG.....	371, 428, 464
HOHN, RÜDIGER.....	22
HOLZER, NICOLAI.....	223
HORNÍČKOVÁ, KATEŘINA.....	403
HOSSAIN, IMRAN MD.....	342
HUANG, HAI.....	122
HUTH, JULIA.....	245
I	
IWASZCZUK, DOROTA.....	129
J	
JACOBSEN, KARSTEN.....	13
JÜRGENS, CARSTEN.....	216
JUTZI, BORIS.....	179
K	
KELLER, SINA.....	110
KEMPER, GERHARD.....	170
KERSTEN, THOMAS.....	35, 454
KNIEPS, ELMAR.....	203
KNOPP, LISA.....	464
KOCH, TOBIAS.....	129
KÖHN, ALEXANDER.....	472
KOLBE, THOMAS H.....	288, 316
KOROVIN, DMITRII.....	434
KREBS, BENNET.....	332
KRUSE, CHRISTIAN.....	245
KRZYSZEK, PETER.....	521
KUHN, ANDREAS.....	122
KUTZNER, TATJANA.....	288
L	
LANGE, EVA.....	363
LINDSTAEDT, MAREN.....	35
LORENZ, ECKEHARD.....	482
LUHMANN, THOMAS.....	350, 514
M	
MAIWALD, FERDINAND.....	55, 444
MALLET, CLÉMENT.....	179
MAYER, HELMUT.....	122
MEHRING, JANNA.....	332
METTIG, NORA.....	482
MEYER, KERSTIN.....	492
MEYER, M. FABIAN.....	216
MICHELINI, MARIO.....	122
N	
NEUBERGER, HAJNALKA.....	245
NGUATEM, WILLIAM.....	122

O

OEHRLEIN, JOHANNES	138
ORTWEIN, ANNETTE	332

P

PAČLÍKOVÁ, KLÁRA	403
PETERS, JULIANE	90
PFEFFER, INGO	216
PFEIL, ISABELLA MARIA	498
PILZ, MATTHIAS	508
PÖCHTRAGER, MARKUS	414
POLEWSKI, PRZEMYSŁAW	521
PREUSKER, RENÉ	482
PREUSZ, MICHAL	403

Q**R**

RAUTZ, KONRAD	508
REIMER, CHRISTOPH	498
REINHARDT, WOLFGANG	342
REULKE, RALF	22, 192, 276, 280, 383, 391, 482
RIENOW, ANDREAS	332
ROFALLSKI, ROBIN	514
ROSSMANN, HANNES	22
ROTTENSTEINER, FRANZ	245
RUEß, DOMINIK	192
RUNCEANU, LAVINIA	147

S

SCHMITZ, MATTHIAS	122
SCHNEIDER, DANILO	55, 444
SCHREYER, QUIRIN	422
SCHRÖDER, UWE	1
SCHULTHEIß, JÖRN	203
SCHULTZ, JOHANNES	332
SMITH, PHILIP	223
SÖRGEL, UWE	233
SPIESS, CHRISTOPHER	391
STEINSIEK, MORITZ	521
STEPHANI, MANFRED	428
STILLA, UWE	97, 129, 160, 262, 302, 371, 434, 464

T

TAUTZ, JÜRGEN	422
TEMPEL, MICHAEL	203
TSCHIRSCHWITZ, FELIX	454

U

ULLMANN, TOBIAS	363
-----------------------	-----

V

VAN DIJK, THOMAS C.	138
VIETZE, THERESA	444
VILLAMIL LOPEZ, CARLOS	97
VOLKMANN, ARMIN.....	76

W

WAGNER, WOLFGANG	498
WEICHEL, HORST.....	276
WEINFURTNER, ANNE.....	403
WEINMANN, MARTIN.....	110, 179
WILLIGES, CHRISTIAN.....	22

X**Y**

YANG, CHIA-HSIANG	233
YAO, WEI	521
YAO, ZHIHANG	316

Z

ZIEMS, MARCEL	245
---------------------	-----

PROGRAMMKOMITEE

Abmayr, Thomas

Baltrusch, Sven

Bareth, Georg

Brunn, Ansgar

Grenzdörffer, Görres

Gülch, Eberhard

Haala, Norbert

Hauert, Jan-Henrik

Höfle, Bernhard

Jacobsen, Karsten

Kaufmann, Viktor

Kersten, Thomas

Kleinschmit, Birgit

Petzold, Bettina

Schneider, Danilo

Sörgel, Uwe

Stilla, Uwe

Weidner, Uwe

Weinmann, Martin

Wetzel, Hans-Ulrich

Wolff, Kirsten