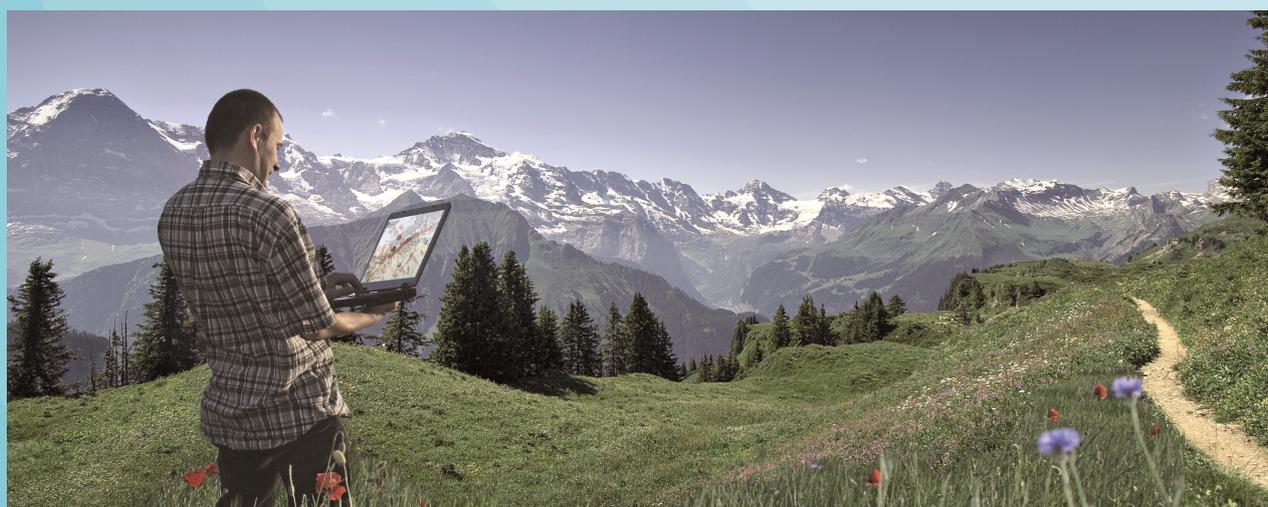


Publikationen der Deutschen
Gesellschaft für Photogrammetrie,
Fernerkundung und Geoinformation e.V.



Band 25

2016



Beiträge

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

7. – 9. Juni 2016
in Bern

***Dreiländertagung der SGPF, DGPF und OVG
Lösungen für eine Welt im Wandel***

ISSN 0942-2870

Thomas P. Kersten, Hrsg.

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



Band 25

2016

Beiträge

36. Wissenschaftlich-Technische
Jahrestagung der DGPF

7. – 9. Juni 2016
in Bern

*Dreiländertagung der SGPF, DGPF und OVG
Lösungen für eine Welt im Wandel*

ISSN 0942-2870

Thomas P. Kersten, Hrsg.

ISSN 0942-2870

Publikationen der
Deutschen Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
Band 25 421 S., Hamburg 2016
Hrsg.: Thomas P. Kersten

© Deutsche Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
München 2016

Zu beziehen durch:

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

Redaktion:

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

VORWORT

Die Schweizerische Gesellschaft für Photogrammetrie und Fernerkundung (SGPF) lädt anlässlich der 36. Wissenschaftlich-Technischen Jahrestagung der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation vom 7. bis 9. Juni 2016 nach Bern ein. Die Tagung findet in Kooperation mit dem durch die SOGI (Schweizerische Organisation für Geoinformation) veranstalteten GEOSummit statt.

Diese Einladung erfolgt auch im Namen der mitveranstaltenden Schwestergesellschaften aus Österreich und Deutschland, der Österreichischen Gesellschaft für Vermessung und Geoinformation sowie der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation, die nach 2001 in Konstanz, 2007 in Muttenz, 2010 in Wien und 2013 in Freiburg zum fünften Mal gemeinsam mit der SGPF diese Dreiländertagung veranstalten.

Wir haben für die Tagung als gemeinsames Motto gewählt:

DREILÄNDERTAGUNG D – A – CH:

LÖSUNGEN FÜR EINE WELT IM WANDEL

Die Bundesstadt Bern mit ihrer von der Aareschlaufe und dem Hausberg Gurten geprägten Topographie und der in das UNESCO-Welterbe aufgenommenen Altstadt bietet eine einladende Kulisse für die Dreiländertagung der drei Schwestergesellschaften. Mit der 1834 gegründeten kantonalen Universität blickt die Stadt Bern außerdem auf eine lange Tradition in Forschung und Lehre zurück. Mit dem BernEXPO-Kongresszentrum steht ein moderner Veranstaltungsort unweit der Berner Altstadt zur Verfügung.

In diesem Jahr wurden 69 Beiträge angemeldet, von denen 56 auf der Tagungs-CD enthalten sind. Insgesamt haben 223 Autoren an den Beiträgen mitgearbeitet. 11 von ihnen haben zu dieser Jahrestagung mehrfach publiziert.

Im Namen des Vorstandes der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF), der Österreichischen Gesellschaft für Vermessung und Geoinformation (OVG) und der Schweizerischen Gesellschaft für Photogrammetrie und Fernerkundung (SGPF) bedanken wir uns bei allen Autoren für die eingereichten Beiträge, die es uns erst ermöglichen, diesen Tagungsband herauszugeben.

Die lokale Organisation auf dem BernEXPO-Kongressgelände wird durch die Firma Hinte Messe GmbH in Zusammenarbeit mit dem Organisationskomitee GEOSummit und der SGPF durchgeführt.

Prof. Dr. Thomas H. Kolbe
Präsident DGPF



Dipl.-Ing. Julius Ernst
Präsident OVG



Dipl.-Ing. Francois Gervais
Präsident SGPF



INHALTSVERZEICHNIS

SENSOREN UND PLATTFORMEN - UAV

GRENZDÖRFFER, G.:	
Pflichtenheft zur Ausschreibung von UAS-Dienstleistungen.	1
CRAMER, M., SABOW, F. & FÖRG, P.:	
Direkte Georeferenzierung in UAS-Anwendungen - Eine Beispielanwendung der Landesvermessung Baden-Württemberg.	13
CLAUSEN, P., REHAK, M. & SKALOUD, J.:	
UAV Sensor Orientation with Pre-calibrated Redundant IMU/GNSS Observations: Preliminary Results.	26

FERNERKUNDUNG IN DER GEOLOGIE

ULRICH, C., BANNEHR, L., LAUSCH, A.:	
Ableitung von Eisen(II,III)oxid in Fließgewässern mittels Multispektraldaten.	34

3D STADTMODELLE I

BLAHÁ, M., VOGEL, C., RICHARD, A., WEGNER, J. D., POCK, T. & SCHINDLER, K.:	
Towards Integrated 3D Reconstruction and Semantic Interpretation of Urban Scenes.	44
WILLENBORG, B., SINDRAM, M. & KOLBE, T. H.:	
Semantic 3D City Models Serving as Information Hub for 3D Field Based Simulations.	54
ARTUSO, R., BÉDERT, N. & GANDOR, F.:	
Challenges for the Swiss Federal Office of Topography swisstopo for the Production of Digital Images and Altimetrical Models using the ADS80 Sensor for the Glacier Monitoring in the Swiss Alps.	66

SENSOREN UND PLATTFORMEN

JACOBSEN, K.:	
Ersatz des SRTM Oberflächenmodells durch AW3D30?	80
KNEER, C., JENAL, A., WEBER, I. & BONGARTZ, J.:	
Ein adaptives und kompaktes Fernerkundungssystem für UL-Fluggeräte – Konzept und Anwendungen.	89
SPRECKELS, V., SCHLIENKAMP, A., GREIWE, A. & EBERHARDT, L.:	
Eignung von ALS, Aero- und UAS-Photogrammetrie zur Früherkennung und Erfassung von Tagesbrüchen.	97
CAVEGN, S., NEBIKER, S. & HAALA, N.:	
Ein systematischer Vergleich zwischen direkter und bildbasierter Georeferenzierung von Mobile Mapping-Stereosequenzen in einem anspruchsvollen Stadtgebiet.	113

AUSWERTUNG VON FERNERKUNDUNGSDATEN I

BAHR, T., HOLZER, N. & LASICA, R.: Automated Point Cloud Analysis for 3D Topographic Change Detection.	124
MANDLBURGER, G., OTEPKA, J., BRIESE, C., MÜCKE, W., SUMMER, G., PFEIFER, N., BALTRUSCH, S., DORN, C. & BROCKMANN, H.: Automatische Ableitung von Strukturlinien aus 3D-Punktwolken.	131
BART, R.: Von der Punktwolke zum Terrainmodell.	143

BILDANALYSE UND BILDVERSTEHEN

BRODOVSKII, M., KOROVIN, D., CHIZHOVA, M., BRUNN, A. & STILLA, U.: Recognizing Separate Structural Elements of Churches Using Neural Networks.	153
SEPPKE, B., DRESCHLER-FISCHER, L. & WILMS, C.: Robuste Superpixel-basierte Segmentierung von spektralen Ferner- kundungsaufnahmen zur Überwachung der Stadtentwicklung anhand von Straßennetzen.	165
HANEL, A., KLÖDEN, H., HOEGNER, L. & STILLA, U.: Personendetektion in Grauwertbildern eines Stereo-Fahrzeugkamera- systems in Nachtsituationen mit variierenden Beleuchtungsverhältnissen. ..	177
KASSER, M.: Utilisation de la corrélation diachronique d'images pour la métrologie de ponts.	188

GEOINFORMATIK

WÜEST, R., ZWICK, M. & NEBIKER, S.: Geospatial Augmented Reality als interaktives Erlebnis im Museum am Beispiel der Swissarena AR App.	197
AGOUB, A., KUNDE, F. & KADA, M.: Potential of Graph Databases in Representing and Enriching Standardized Geodata.	208
CHRISTEN, M., BLASER, S., LOESCH, B., NEBIKER, S.: Prozedurale Modellierung von historischen 3D-Stadtmodellen und deren Aufbereitung für 3D-Visualisierungen in einer Smartphone-Applikation am Beispiel der römischen Stadt Augusta Raurica.	217

RADARFERNERKUNDUNG UND FLUGZEUGLASERSCANNING

- RICHTER, K., WESTFELD, P., MAAS, H.-G. & WEIß, R.:
Korrektur der Signaldämpfung und Analyse zur Bestimmbarkeit der
Gewässertrübung in Laserbathymetrie-Daten. 230
- HOLLAUS, M., PFEIFER, N., MANDLBURGER, G. & INSINNA, P.:
Baumartenklassifikation mittels multi-temporaler Full-Waveform
Laserscanning-Daten. 240
- MORS DORF, F., BRUGGISSER, M., GUILLEN ESCRIBA, C., KÜKENBRINK, D., LEITERER, R.,
SCHNEIDER, F., TORABZADEH, H. & SCHAEPMAN, M.:
Physically-based Multi-temporal LiDAR Traits for Species Discrimination
in a Temperate Mixed Forest. 251

OPTISCHE 3D-MESSTECHNIK I

- REZNICEK, J., HASTEDT, H., EKKELE, T., LUHMANN, T. & JEPPIING, C.:
Analysen zur Datumsfestlegung in photogrammetrischen Projekt
großer Volumina. 259
- KERSTEN, T., PRZYBILLA, H.-J., LINDSTAEDT, M., TSCHIRSCHWITZ, F. &
MISGAISKI-HASS, M.:
Genauigkeitsuntersuchungen handgeführter Scannersysteme. 271

SENSOREN UND PLATTFORMEN - KAMERASYSTEME

- THOMAS, N., CREMONESE, G., ZIETHE, R., GERBER, M., BRÄNDLI, M., ERISMANN, M.,
GAMBICORTI, L., GERBER, T., GHOSE, K., GRUBER, M., GUBLER, P., MISCHLER, H.,
JOST, J., PIAZZA, D., POMMEROL, A., RIEDER, M., ROLOFF, V., SERVONET, A.,
TROTSMANN, W., UTHAICHAROENPONG, T., ZIMMERMANN, C., VERNANI, D.,
JOHNSON, M., PELÒ, E., WEIGEL, T., VIERTL, J., DE ROUX, N., LOCHMATTER, P.,
SUTTER, G., CASCIELLO, A., HAUSNER, T., FICAI VELTRONI, I., DA DEPPO, V.,
ORLEANSKI, P., NOWOSIELSKI, W., ZAWISTOWSKI, T., SZALAI, S., SODOR, B.,
TROZNAI, G., BANASKIEWICZ, M., BRIDGES, J.T., BYRNE, S., DEBEI, S.,
EL-MAARRY, M.R., HAUBER, E., HANSEN, C.J., HENSON, R., IVANOV, A.,
KESTAY, L., KIRK, R., KUZMIN, R., MANGOLD, N., MARINANGELI, L.,
MARKIEWICZ, W.J., MASSIRONI, M., McEWEN, A.S., OKUBO, C.,
TORNABENE, L.L., WAJER, P. & WRAY, J.J.:
The Colour and Stereo Surface Imaging System (CaSSIS) for
ESA's ExoMars Trace Gas Orbiter. 288
- KREMER, J. & GRIMM, P.:
Möglichkeiten und Einschränkungen des kombinierten Einsatzes digitaler
Senkrecht- und Schrägluftbildsensoren in einer Großformatkamera. 298

AUSWERTUNG VON FERNERKUNDUNGSDATEN II

BÜHLER, Y., STOFFEL, A., ADAMS, M., BÖSCH, R. & GINZLER, C.: UAS Photogrammetry on Homogenous Snow Cover.	306
LANDTWING, S. & GÖTZ, T.: Unbeschwertes und sicheres Skivergnügen dank Geodaten.	593
HERMLE, D. & RUNKEL, I.: Geoprodukte im Aufwind - Die GeoApp.UAS.	317

OPTISCHE 3D-MESSTECHNIK II

HOLLENSTEIN, D. & NEBIKER, S.: Bild-basierte 3D-Rekonstruktion texturarmer Kleinwiederkäuerknochen – erste Leistungsuntersuchungen.	327
WAHBEH, W. & NEBIKER, S.: Kombination unterschiedlicher Photogrammetrischer Ansätze zur Rekonstruktion zerstörter Kulturdenkmäler.	338
WEINMANN, M., MALLET, C. & BRÉDIF, M.: Segmentation and Localization of Individual Trees from MMS Point Cloud Data Acquired in Urban Areas.	351
GALLIANI, S., LASINGER, K. & SCHINDLER, K.: Gipuma: Massively Parallel Multi-view Stereo Reconstruction.	361

3D STADTMODELLE II

WAGNER, D., CASPER, E. & COORS, V.: CityGML Semantik unter der Lupe – welche Konsequenzen ergeben sich für die Validierung von 3D Stadtmodellen?.	370
KUTZNER, T. & KOLBE, T. H.: Extending Semantic 3D City Models by Supply and Disposal Networks for Analysing the Urban Supply Situation.	382
STEIN, A. & BACHMANN, M.: Innovative, quantitative 3D-Planungswerkzeuge für sichere Städte der Zukunft.	395

AUSWERTUNG VON FERNERKUNDUNGSDATEN III

ZIELEWSKA-BÜTTNER, K., ADLER, P., PETERSEN, M. & BRAUNISCH, V.: Parameters Influencing Forest Gap Detection Using Canopy Height Models Derived From Stereo Aerial Imagery.	405
IMMITZER, M., VUOLO, F., EINZMANN, K., NG, W. T., BÖCK, S. & ATZBERGER, C.: Verwendung von multi-spektralen Sentinel-2 Daten für die Baumarten- klassifikation und Vergleich mit anderen Satellitensensoren.	417

AUS- UND WEITERBILDUNG

RIENOW, A., GRAW, V., HEINEMANN, S., SCHULTZ, J., SELG, F. & MENZ, G.: Mathematikunterricht aus dem All – Interdisziplinäre Lernwerkzeuge für den Einsatz von Erdbeobachtung im Schulunterricht.	428
MEYER, T. & BRUNN, A.: Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen durch "Blended Learning".	436
CANNELLE, B.: Les drones au service de l'enseignement.	449

POSTER

EDER, K., HOEGNER, L., RIEGER, W., KAISER, M. & STILLA, U.: Photogrammetric Recording and Modelling of a Valley Floor using RPAS-mounted RGB and TIR Cameras.	461
GÄNSRICH, A., BAYER, S. & BUCHER, T.: Entwicklung einer zeitoptimierten automatischen Extraktion von Wasserflächen aus hochauflösenden Luftbilddaten (True Ortho Mosaik, Digitales Oberflächenmodell).	471
RÖDER, M., HEUSINGER, A. & BRUNN, A.: Konzeption und geometrische Bewertung der flächenhaften Berechnung von Digitalen Oberflächenmodellen aus Luftbildern.	477
STEINER, G. & CANNELLE, B.: La Photogrammétrie Peut-Elle Remplacer La Topométrie Traditionnelle Pour Du Levé De Bâtiments ?.	487
YONG, X. & SÖRGEL, U.: Recognition of Road Network in Aerial Images using CRF and Marked Point Process.	496

BEITRÄGE KARL-KRAUS-NACHWUCHSPREIS

BAUMANN, A. G. & BALTSAVIAS, E.:	
Einzelbaumdetektion in bewaldetem Gebiet auf Basis von luftgestützten LiDAR-Daten.	506
COENEN, M. & ROTTENSTEINER, F.:	
Klassifikation von Stereobildern aus Mobile Mapping Daten mittels Conditional Random Fields.	516
CROMMELINCK, S.:	
Simulating an Autonomously Operating Low-Cost Static Terrestrial LiDAR for Multitemporal Maize Crop Height Measurements.	524
JUNG, M.:	
Cloud basiertes Rendering von komplexen 3D-Stadtmodellen.	531
KAHMEN, O. & LUHMANN, T.:	
Entwicklung einer großen Invardraht-Maßverkörperung zur Anwendung in der Industriephotoграмmetrie.	538
SCHORCHT, M.:	
Entwicklung und Anwendung einer Methode zur Erstellung von Wechsel- bilanzen der Flächennutzung auf Grundlage des ATKIS Basis-DLM.	545
SUDMANN, M., TIEDE, D., AUGSTEN, N., BARALDI, A., BELGIU, M. & LANG, S.:	
Array-Datenbanken für semantische inhaltsbasierte Suche und Analyse in Satellitenbildarchiven.	555
XAVER, A., DORIGO, W. & WAGNER, W.:	
Automated Quality Control Procedures for the International Soil Moisture Network.	565
XIAO, M. & CRAMER, M.:	
Superweitwinkel- und Fisheye-Kameras in photogrammetrischen Anwendungen.	574
MANANDHAR, A., Hoegner, L. & Stilla, U.:	
Detection of Star-shaped Objects in Aerial Images: Use Case Oil Palm Tree Detection.	582
AUTORENVERZEICHNIS	603
PROGRAMMKOMITEE	609

Pflichtenheft zur Ausschreibung von UAS-Dienstleistungen

GÖRRES GRENZDÖRFFER¹

Zusammenfassung: Der Bedarf und die Nachfrage nach UAS-Befliegungsdienstleistungen durch Behörden und Firmen, z.B. zur 3D-Datenerfassung und Kartierung von beliebigen Objekten nehmen rasant zu. Gleichzeitig entstehen viele neue Firmen, die photogrammetrische Standardprodukte anbieten wollen, ohne über entsprechende photogrammetrische Grundkenntnisse zu verfügen. Da es ein völlig neues Feld ist, fehlt es – im Unterschied zum klassischen Bildflug – an verbindlichen Standards und Richtlinien, um diese Dienstleistungen fachgerecht nachfragen und anbieten zu können. Der Beitrag enthält ein detailliertes Pflichtenheft zur Ausschreibung von UAS-Befliegungsdienstleistungen, das Behörden und Unternehmen mit geringen individuellen Anpassungen direkt verwenden können. Der Beitrag geht allerdings aus Platzgründen nicht auf die dafür notwendigen photogrammetrischen Grundlagen ein.

1 Einleitung

In den letzten zwei bis drei Jahren ist eine Vielzahl kleinerer Firmen entstanden, die mit UAS-Aufnahmen ihr Geld verdienen wollen. Auch einzelne Vermessungs- und Ingenieurbüros haben sich ein oder mehrere UAS angeschafft, um ihr Portfolio zu verbreitern und an dem rasanten technologischen Fortschritt teilhaben zu können (GRENZDÖRFFER 2016). Das Spektrum der angebotenen Dienstleistungen unter den zurzeit gültigen rechtlichen Rahmenbedingungen (BMVI 2014) ist groß und reicht von Fotos und Videos für Werbezwecke und Eventbegleitung, über die Kartierung von Wildschäden und die Vermessung von Bauwerken bis hin zu Spezialbefliegungen zur Inspektion von Solar- und Windkraftanlagen. Es gibt keine genaue Übersicht oder ein Branchenverzeichnis für die Anbieter solcher Dienstleistungen. Deshalb ist es schwierig eine genaue Marktübersicht zu geben. Viele Dienstleister, die z.B. im „Drohnenforum“ registriert sind, bieten auch „Vermessungsflüge“ an, aber nicht alle verfügen über die dafür notwendige Technik und das dazugehörige Knowhow.

Bei der Beauftragung für einen UAS-Bildflug sind mehrere Rahmenbedingungen vom Auftraggeber zu spezifizieren, um einerseits ein optimales Ergebnis zu bekommen und es gleichzeitig dem Auftragnehmer zu ermöglichen ein passgenaues und angemessenes Angebot zu erstellen. Das trifft im Besonderen auf Behörden zu, die solche Dienstleistungen in der Regel ausschreiben müssen, und sich schon aus diesem Grund intensiver mit dem Thema beschäftigen müssen. Gleichwohl gibt es keine verbindlichen Vorgaben, so dass die beauftragten Firmen entweder teilweise unnötig hohe Anforderungen (= Kosten) zu erfüllen haben oder aber auch Produkte abliefern dürfen, die für die nachfolgenden Auswertungen nur bedingt geeignet sind.

Die nachfolgenden Ausführungen konzentrieren sich auf die Ausschreibung von UAS-Bildflügen mit dem Ziel der Erstellung photogrammetrischer Standardprodukte. Im Einzelnen handelt es sich dabei um digitale „true“ Orthophotos, Digitale Oberflächenmodelle und eine 3D-Punktwolke. D.h. die Richtlinie ist für alle Aufträge gedacht, die einen gewissen Vermessungsanspruch haben bzw.

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

für ein längerfristiges Monitoring gedacht sind. In diesem Falle ist eine genaue Georeferenzierung zwingend, um beispielsweise (Massen)Veränderungen detailliert erfassen zu können.

Das Pflichtenheft ist nicht zur Anwendung zu bringen, wenn es sich um die Ausschreibung von reinen Werbevideos oder Inspektionsflügen handelt, die im Anschluss nicht weiter geometrisch ausgewertet werden sollen. Wenn es bei einem Projekt nicht auf eine hohe absolute Lagegenauigkeit in cm-Bereich ankommt, kann man durchaus auf eine passpunktlose Aerotriangulation übergehen, und die genäherten GPS-Daten der Aufnahmen verwenden. Schließlich sind die Einmessung von Passpunkten und deren manuelle Bildmessung aufwendig und teuer. In diesem Falle liegen die erzielbaren Genauigkeiten bei ca. 1 - 3 m in absoluter Lage und Höhe. Der Begriff „absolute“ Lagegenauigkeit bedeutet in diesem Zusammenhang, dass die genannten Abweichungen an Kontrollpunkten gemessen werden können. Der Bildverband selber ist in sich homogen und es treten keine offensichtlichen Fehler an den Schnittkanten benachbarter Bildern auf. Nur muss man sich für diesen Fall auch im Klaren sein, dass eine multitemporale Auswertung schwierig wird, da Bilder unterschiedlicher Zeitpunkte nicht mehr übereinander passen und somit Veränderungen nur unzureichend erfasst werden können.

Zwei Sachen vorweg:

- Die Photogrammetrie misst nur die direkt luftbildsichtbare Oberfläche, d.h. alle von oben zu sehenden Bildbereiche die in mindestens zwei Bildern zu erkennen sind. D.h. in abgeschatteten, von oben nicht einsehbaren Bereichen gibt es keine 3D-Informationen. Durch eine geeignete Bildflugplanung kann der Anteil der sichttoten Bereiche verringert, aber nicht verhindert werden. Bei bewachsenen Flächen oder im Wald wird auch nur dessen Oberfläche gemessen und nicht die Höhe der darunterliegenden Oberfläche. Im Nachgang, d.h. im Rahmen der Analyse der 3D-Punktswolke kann man versuchen, die so genannten „nicht Bodenpunkte“ von den Bodenpunkten mit Hilfe automatischer statistischer Verfahren zu trennen. Anschließend wird zwischen den verbleibenden Bodenpunkten interpoliert, um ein digitales Geländemodell abzuleiten. Das funktioniert für viele Oberflächen auch ganz zufriedenstellend, aber nicht für alle.
- Grundsätzlich ist anzumerken, dass ein UAS-Bildflug in der freien Natur stattfindet und nicht jede Witterung einen Flug zulässt. Das bedeutet, es ist immer ein gewisses Zeitfenster (mind. 2 Wochen) für eine Befliegung einzuplanen. Falls zusätzliche Erschwernisse vorliegen, z.B. Befliegungen nur bei bestimmten Tidewasserständen oder zu definierten Uhrzeiten oder bei definierten Licht- bzw. Wetterverhältnissen ist ein entsprechend größeres Zeitfenster einzuplanen. Ausschlusskriterien für einen Flug sind dabei starker Wind (> 4 Bft), Nebel, Regen, Schneefall und anderes schlechtes Wetter. Vorzugsweise sollte der Flug entweder bei Sonnenschein oder gleichmäßig bedecktem Himmel stattfinden. Bei unterschiedlich bewölktem Himmel kann es sein, dass sich z.B. ein Wolkenschatten durch die Bilder schiebt. Diese sind auch im fertigen Orthophoto zu sehen, was rein optisch nicht so schön ist und eine spätere Nutzung unter fernerkundlichen Aspekten (z.B. Klassifikation) erschwert, bzw. ganz verhindert. Für Messaufgaben hingegen, z.B. zur Volumenbestimmung, spielen die Schatten nur eine untergeordnete Rolle. Das bedeutet, wenn es auf schöne Bilder ankommt, die anschließend auch inhaltlich klassifiziert werden sollen, muss ein durchgehend wolkenfreier Himmel gegeben sein und der Flug muss bei wechselnden Wolken abgebrochen werden. Wenn es mehr auf die Geometrie ankommt, sind Wolkenschatten nur ein untergeordnetes Problem und es ergeben sich deutlich größere Flugfenster.

2 Problemstellung

Das umfangreiche Pflichtenheft dient zudem dazu nur potentiell geeignete Firmen für diese Aufgaben zu interessieren, die über das entsprechende Knowhow verfügen und sich mit den entsprechenden Fachbegriffen und der notwendigen Hard- und Software auskennen. Durch das standardisierte Verfahren ergibt sich für die beteiligten Firmen zudem der Vorteil an Ausschreibungsverfahren kostengünstiger und effektiver teilnehmen zu können, da ein Teil der Unterlagen mehrfach genutzt werden können. Das Pflichtenheft beinhaltet folgende Punkte:

- Allgemeines und Rechtliches
- Spezifikation der Dienstleistung (UAS-Befliegung)
- Spezifikation der photogrammetrischen Standardprodukte (digitales „true“ Orthophoto, Digitales Oberflächenmodell und 3D-Punktwolke)
- Spezifikationen für die Kontrolle der Genauigkeit, Datenabgabe und Lieferung
- transparentes Auswahlverfahren für den wirtschaftlichsten Anbieter

Die nachfolgende tabellarische Darstellung fasst die oben genannten Punkte für eine Beauftragung einer UAS-Befliegung für vermessungstechnische Zwecke zusammen. Grundlage für die Tabelle ist teilweise der ADV-Leitfaden zur Ausschreibung einer digitalen Luftbildbefliegung als Grundlage zur Herstellung von ATKIS®-DOP und stereoskopischen Auswertung, ADV 2015. Viele Punkte sind für die Ausschreibung von UAS-Flügen modifiziert und ggf. ergänzt worden. Zudem sind an mehreren Stellen Hinweise eingearbeitet, die bei einer Ausschreibung entfernt werden müssen. Die rot markierten Informationen sind individuell für jeden Flug zu definieren.

2.1 Allgemeines und rechtliches

Zweck der Befliegung	<i>Hinweis: Beschreibung des Zwecks, damit der AN die Ziele der Befliegung versteht und ggf. seine Vorstellungen mit einbringen kann.</i>
Gebietsbeschreibung und Einschränkungen	<p>(1) Dem Auftragnehmer wird das Untersuchungsgebiet in Form einer Karte und digital als Shape-, KML-Datei zur Verfügung gestellt</p> <p>(2) Es bestehen folgende räumliche Einschränkungen im Hinblick auf</p> <p>a) Erreichbarkeit (<i>mit Auto, zu Fuß, Schiff</i>)</p> <p>b) Begehbarkeit (z.B. Naturschutzgebiet, steile Böschung, Schilf ...)</p> <p>(3) Über Flugbeschränkungen im Bildfluggebiet hat sich der Auftragnehmer selbst zu informieren und ggf. eine entsprechende Erlaubnis einzuholen.</p> <p>(4) Der Auftraggeber stellt die für eine etwaige Einzelerlaubnis notwendigen Nachweise (Eigentümer bzw. Nutzungsnachweis) zur Verfügung.</p> <p>(5) Die Verantwortung für eine fristgerechte Auftragsausführung liegt beim Auftragnehmer. Damit sind z.B. zeitliche und räumliche Flugbeschränkungen, wetterbedingte oder sonstige technische Risiken einzuplanen und somit Bestandteil des Angebotes.</p>
Nutzungsrechte Eigentum	<i>/ Hinweis: Mit der nachfolgend beschriebenen Rechteübertragung auf den Auftraggeber gehen auch Pflichten bei der Veröffentlichung der</i>

Daten einher. Das betrifft insbesondere den Datenschutz, dieser ist nämlich durch den Rechteinhaber bei der etwaigen Veröffentlichung der Bilder zu gewährleisten! Diese Verpflichtung ist unabhängig von der datenschutzrechtlichen Verpflichtung des UAS-Betreibers, die er zur Erlangung einer Aufstiegserlaubnis unterschreibt.

(1) Der Auftragnehmer räumt dem Auftraggeber das ausschließliche, örtlich unbeschränkte, übertragbare, dauerhafte, unwiderrufliche, umfassende und unkündbare Nutzungsrecht ein, über das der Auftraggeber uneingeschränkt weiter verfügen kann, an den im Zusammenhang mit der Ausführung des Auftrages aufgezeichneten digitalen Daten sowie an allen aufgrund des Bildfluges erstellten Daten.

(2) Das Nutzungsrecht umfasst insbesondere das Recht, die o.g. Daten - digital oder analog dauerhaft oder temporär zu speichern, zu laden, anzuzeigen und ablaufen zu lassen;

- abzuändern, zu bearbeiten oder auf anderem Weg umzugestalten;

- auf einem beliebigen bekannten Medium oder in einer anderen Weise zu speichern, zu vervielfältigen, auszustellen, zu veröffentlichen, zu verbreiten oder öffentlich wiederzugeben;

- in Datenbanken, Datennetzen und Online-Diensten einzusetzen, einschließlich des Rechts, die Daten den Nutzern der vorgenannten Datenbanken, Netzen und Online-Dienste zur Recherche und zum Abruf mittels vom Auftraggeber gewählter Tools bzw. zum Herunterladen zu Verfügung zu stellen;

- durch Dritte zu nutzen oder für den Auftraggeber betreiben zu lassen.

(3) Der Auftragnehmer ist nicht berechtigt, die Daten selbst zu nutzen oder eine Nutzung durch Dritte zu ermöglichen.

(4) Soweit der Auftragnehmer sich eines Nachauftragnehmers bedient, hat er diesem gegenüber vertraglich sicherzustellen, dass die Nutzungsrechte mit dem oben beschriebenen Umfang auf den Auftraggeber übergehen.

(5) Das Nutzungsrecht erstreckt sich auch auf Teilergebnisse.

(Optional) (6) Es dürfen während des Fluges nur Aufnahmen für diesen Auftrag gemacht werden.

Alternative für eine mehrfache Nachnutzung der Befliegungsergebnisse

(1) Der Auftraggeber räumt dem Auftragnehmer das örtlich unbeschränkte, übertragbare, dauerhafte, unwiderrufliche, umfassende und unkündbare Nutzungsrecht ein, über das der Auftragnehmer uneingeschränkt weiter verfügen kann, an den im Zusammenhang mit der Ausführung des Auftrages aufgezeichneten digitalen Daten sowie an allen aufgrund des Bildfluges erstellten Daten.

(2) Der Auftraggeber verzichtet auf das ausschließliche Nutzungs- und Verwertungsrecht.

Hinweis: Die Landesvermessungsämter haben die Erfahrung gemacht, dass ein Verzicht des ausschließlichen Nutzungs- und Verwertungsrechts zu einer Preisreduktion bis zu 25% führen kann. Allerdings trifft das wohl nur auf solche landesweiten Befliegungen zu, die z.B. an Google o.ä. auch verkauft werden können. Bei UAS-Befliegungen wird sich seltener ein Zweitverwerter finden lassen.

- Bieter / Bietergemeinschaft** (1) Bieter / Bietergemeinschaft sind in § 6 ff. VOL/A (Teilnahme am Wettbewerb) geregelt.
(2) Die Fachkunde, Leistungsfähigkeit und Zuverlässigkeit ist von jedem im Angebot aufgeführten Unternehmen einzeln nachzuweisen. Werden besondere Nachweise gefordert, sind diese vom jeweiligen Ausführenden zu erbringen.
- Nachauftragnehmer (Unterauftragnehmer / Subunternehmer)** (1) Ohne Genehmigung des Auftraggebers dürfen über den Inhalt des erteilten Auftrages keine Subverträge abgeschlossen werden.
- Ansprechpartner** (1) Bei Angebotsabgabe ist ein Ansprechpartner für die Ausschreibung und für die Durchführung des Projektes zu benennen.
- Mitwirkung des Bieters** 1) Zum Schutz des Auftraggebers:
Enthalten die technischen Bedingungen nach Auffassung des Bieters Unklarheiten, so hat der Bieter vor Angebotsabgabe schriftlich darauf hinzuweisen, auch wenn er den Hinweis schon vorher in anderer Form gegeben hat. (Quelle: § 4 Abs. 3 VOL/B)
- Haftpflicht** (1) Der Auftragnehmer verpflichtet sich, den Auftraggeber von allen Haftpflichtansprüchen freizustellen, die gegen diesen im Zusammenhang mit dem übernommenen Auftrag von Dritten erhoben werden.
(2) Der Auftragnehmer hat auf Verlangen nachzuweisen, dass er hinsichtlich aller Haftpflichtansprüche, die sich aus der Ausführung des übernommenen Auftrages ergeben können, eine Haftpflichtversicherung in ausreichender Höhe abgeschlossen hat und laufend unterhält. Der Auftraggeber ist berechtigt, rückständige Prämien anstelle des Auftragnehmers an den Versicherungsunternehmer zu zahlen und die Beiträge von der dem Auftragnehmer zustehenden Vergütung oder einer von ihm hinterlegten Sicherheit einzubehalten.
- Nachweis der Eig-nung (§ 7 Abs. 1 VOL/A)** (1) Informationen über einzusetzende UAS und Kameras, technische Parameter
(2) Angaben zum GPS-System oder Tachymeter, das für die Positionsbestimmung der Passpunkte verwendet wird, Typ und Genauigkeit
(3) Angaben zum Zugriff auf die Ressourcen zu (1) und (2)
(4) Nachweis der Allgemeinerlaubnis für das Befliegungsgebiet (Bundesland)

(5) Angaben zum auftragsbezogenen technischen Workflow incl. Besonderheiten und Maßnahmen zur Qualitätskontrolle (Verwendete Software, Passpunktmarken, Qualitätsprotokoll der GPS-Messung etc.)

(6) **Option 1** Referenz-Ergebnisse aus einem vergleichbaren Auftrag (UAS, Flugzeitpunkt, Bodenauflösung, Kamera)

Option 2 Referenzen über vergleichbare Aufträge der letzten 2 Jahre
Anlage 1: Verbindliche Angaben des Bieters über Angaben nach VOL sowie Kontaktdaten

Inhalte des Angebots

(1) oben genannte Nachweise und Informationen zur verwendeten Technik

(2) graphische Darstellung der Bildflugplanung und der ungefähren Lage der Passpunkte

(3) Zeitplanung (Ausführungs- / Liefertermine),

(4) Aufschlüsselung des Angebotspreises nach folgenden Kriterien:
Befliegungskosten

Kosten für die Messung der Pass- und Kontrollpunkte

Kosten für die photogrammetrische Datenauswertung

Fahrt- und ggf. Übernachtungskosten

Kosten für Datenlieferung, Präsentation und Projektmanagement

2.2 Spezifikation der UAS-Befliegung

Digitale Kamera

Hinweis: Die Vorgaben zur Kamera sind abhängig vom Zweck der Befliegung und der dafür notwendigen Sensoren. Bei RGB-Kameras steht die Bildqualität im Vordergrund. Diese wird im Wesentlichen durch die Größe der Sensorelemente und die Güte der Optik bestimmt. Gleichzeitig bedeutet eine höhere Bildqualität in der Regel eine schwerere Kamera und somit ein größeres und teureres UAS. Es können z.B. folgende Vorgaben gemacht werden:

(1) Zulässig sind ausschließlich digitale CCD-Kameras, die über einen Zentral- oder Schlitzverschluss verfügen. Kameras mit einem digitalen Verschluss sind nicht zulässig.

(2) Der Öffnungswinkel des Objektivs soll nicht größer sein als 80°.

(3) Die Auslösezeit soll so gewählt werden, dass Bewegungsunschärfen < 0.5 Pixel liegen, d.h. nicht länger als 1/500 s.

Option bei Videoflügen für Inspektions-, Dokumentations- oder Werbezwecke

(1) Zulässig sind ausschließlich digitale Videokameras, die über eine Stabilisierung (Gimbal) verfügen und Full HD Video (1080p) oder in 4K aufzeichnen können.

- (2) Neben dem UAS-Piloten ist ein zweiter Operateur als Kamerabedienter vorzusehen. (Diese Vorgabe ist in Abhängigkeit des Bundeslandes (z.B. BaWü) ggf. auch Pflicht bei Videoflügen)
- Option für ergänzende Schrägaufnahmen zur Dokumentation** (1) Das Untersuchungsgebiet ist mit Schrägaufnahmen unter einem Winkel von **30 – 60°** überblicksmäßig aus einer Höhe **von ?? m** zu erfassen.
(2) Die Aufnahmepositionen sind
Option 1 in der angehängten Shape-Datei vorgegeben
Option 2 noch mit dem Auftraggeber zu vereinbaren
Option 3 nach freiem Ermessen so zu wählen, dass die Sonne nicht in das Bild einfällt und zu Überstrahlungen führt
- Bodenauflösung** (1) Es ist eine originäre Bodenauflösung von mind. **? cm** zu garantieren.
(2) Die Bodenauflösung gilt für den tiefsten Geländepunkt des Bildfluggebietes.
- Längs- / Querüberdeckung** (1) Es ist eine Querüberdeckung von mindestens **60 %** und eine Längsüberdeckung von mindestens **80 %** einzuhalten.
(2) Die geforderten Überdeckungsverhältnisse sind am höchsten Geländepunkt des Bildfluggebietes zu erreichen.
Hinweis: Bei stark bewegtem Gelände kann aus der Bedingung der Bodenauflösung in Kombination mit der Längs- und Querüberdeckung das Erfordernis nach mehreren Flugniveaus resultieren. Dieses sollte in der Bildflugplanung gesondert angezeigt werden.
- Gebietsabdeckung** (1) Das Befliegungsgebiet ist so aufzunehmen, dass eine lückenlose stereoskopische Erfassung möglich ist.
(2) Die Flugstreifen sind über das Befliegungsgebiet hinaus um je eine zusätzliche Basislänge am Anfang und am Ende zu erweitern.
- Koordinatenreferenzsystem** Abbildung: **UTM32, bezogen auf ETRS89, Höhenreferenzsystem DHHN 92**
- Ausführungsfrist** (1) Die Ausführungsfrist bezieht sich nur auf die Durchführung der Befliegung.
(2) Es wird ein Befliegungszeitraum festgesetzt (**vom xx bis xx**).
(3) Die individuelle Ausführungsfreigabe eines jeden Bildflugs durch den Auftraggeber zu Beginn der Ausführungsfrist ist vom Auftragnehmer zu beachten.
(4) Der Auftraggeber behält sich eine Verlängerung der Frist vor. Ein Rechtsanspruch darauf besteht nicht.
(5) Der Befliegungsauftrag erlischt ohne Setzen einer Nachfrist mit Ablauf des letzten Tages des gesetzten Befliegungszeitraums beziehungsweise der eingeräumten Verlängerung.
(6) Der Auftragnehmer hat das Recht, ausgeführte Befliegungen innerhalb der Ausführungsfrist nachzubessern oder zu wiederholen. Der Auftraggeber ist von dieser Absicht unverzüglich zu unterrichten.

- (7) Soweit der Auftragnehmer die Befliegung unverschuldet nicht termingerech zu Ende bringt, kann der Auftraggeber die mängelfreie Teilbefliegung abnehmen und vergüten, wenn die Daten verwertbar sind. Ein Rechtsanspruch darauf besteht nicht.
- Pass- und Kontrollpunkte**
- (1) Als Messmethode zur Passpunktbestimmung soll folgendes Gerät eingesetzt werden: **RTK-GPS, Tachymeter**
- (2) Es sind mindestens **6** Passpunkte zu messen. Pro 100 Bilder sind 2 zusätzliche Punkte einzumessen.
- (3) Die Messung und Genauigkeit der Pass- und Kontrollpunkte ist durch eine geeignete Passpunktconfiguration und Methodik zu gewährleisten und zu dokumentieren.
- (4) Es sind mindestens **3** unabhängige Kontrollpunkte zu messen. Pro 100 Bilder ist **ein** zusätzlicher Punkt einzumessen.
- Wetterbedingungen**
- (1) Die Bildqualität darf nicht durch Nebel, **Wolkenschatten**, großflächigen Rauch oder Dunst beeinträchtigt sein.
- (2) Schnee, Raureif, Überflutung, Eis
- a) Das Befliegungsgebiet muss grundsätzlich frei von großflächigen Überflutungen, Schnee, Raureif und Eis sein.
- b) Ausnahmen sind nur nach vorhergehender Absprache und der Zustimmung des Auftraggebers zugelassen.
- Verfahren bei Unterbrechung eines Flugstreifens**
- (1) Wird ein Flugstreifen unterbrochen, dann ist dieser vollständig erneut zu befliegen, so dass ein eindeutiger, widerspruchsfreier Datensatz erzeugt wird.

2.3 Spezifikation der Ergebnisse

- Orthophoto**
- (1) Die Bodenauflösung des Orthophotomosaiks soll der **tatsächlichen** Bodenauflösung entsprechen. (*Hinweis: geraden Wert vorgeben*)
- (2) Das Orthophotomosaik soll homogene radiometrische Eigenschaften aufweisen, d.h. die Grenzen benachbarter Luftbilder sollen farblich angeglichen sein.
- Digitales Oberflächenmodell**
- (1) Die Bodenauflösung des digitalen Oberflächenmodells soll dem **3fachen Wert der tatsächlichen** Bodenauflösung entsprechen. (*Hinweis: geraden Wert vorgeben*)
- Digitales Geländemodell**
- (1) Die Bodenauflösung des digitalen Geländemodells soll der dem **3fachen Wert der tatsächlichen** Bodenauflösung entsprechen. (*Hinweis: geraden Wert vorgeben*)
- (2) Die Höhengenaugigkeit hängt ab von der Gitterweite, dem Bewuchs, dem Erfassungszeitpunkt, etc. Beim Bewuchs können, abhängig vom Befliegungszeitpunkt und der Art des Bewuchses, größere Abweichungen auftreten.

3D-Punktwolke

(1) Die georeferenzierte 3D-Punktwolke ist vor der Klassifizierung derart zu plausibilisieren, dass offensichtliche Fehlmessungen, Punkte die im Bezug zum Bearbeitungsgebiet deutlich zu hoch oder zu tief liegen, aus den Daten eliminiert werden. Die Höhenwerte dieser Filterung sind zwischen AG und AN abzustimmen.

(2) Die Punktwolkendichte soll um den Faktor **3** kleiner sein als die Bodenauflösung.

2.4 Spezifikationen der Genauigkeitskontrolle, Datenabgabe und Lieferung

Kontrolle der Zielgenauigkeit

Hinweis: Die nachfolgend geforderten Genauigkeiten sind nur in Bildverbänden realisierbar, in denen eine günstige Verteilung von Passpunkten möglich war und in denen keine größeren Wasser- oder Forstflächen enthalten sind, die ein genaues Matching beeinträchtigen. Das bedeutet, ausreichende Anzahl von Passpunkten ist vorhanden und die Blockränder sind gut abgebildet.

(1) Überprüfung anhand von Passpunkten:

a) Mittlere Residuen an den Passpunkten \leq 2-fache Bodenauflösung für Lage und Höhe für den Fall, dass dieser Wert größer ist als die Messgenauigkeit (z.B. GPS, Tachymeter) der Passpunkte.

b) Falls die Genauigkeit der Messmethode der Passpunkte (z.B. GPS, Tachymeter) geringer ist als die 2-fache Bodenauflösung, ist die 2-fache Messgenauigkeit der jeweiligen Messmethode anzusetzen.

(2) Überprüfung anhand von Kontrollpunkten:

a) Die mittleren Abweichungen in Lage und Höhe an unabhängigen Kontrollpunkten dürfen das Dreifache der Bodenauflösung oder das dreifache der Messgenauigkeit der zugrundeliegenden Kontrollpunktmessgenauigkeit (z.B. GPS, Tachymeter) nicht überschreiten.

Qualitätskontrolle beim Auftragnehmer

(1) Der Auftragnehmer liefert nur Daten,

a) die vor Abgabe auf offensichtliche Fehler überprüft wurden und

b) die für den genannten Zweck uneingeschränkt verwendbar sind.

(2) Ergibt sich während der Abnahmekontrollen des Auftraggebers, dass derartige, offensichtliche Fehler bereits bei Abgabe vorgelegen haben, hat der Auftraggeber das Recht, Ersatz für die entstandenen Aufwendungen zu verlangen.

Protokoll

Dem Auftraggeber ist das Ergebnisprotokoll der verwendeten Software zu übergeben. Das Protokoll muss mindestens folgende Punkte beinhalten:

Eingesetzte Kamera und innere Orientierung, Anzahl prozessierter Bilder, Verteilung der Passpunkte und deren Residuen an Passpunkten und den Kontrollpunkten, Mehrfachüberlappung, ...

- Hinweis: Optional können vom AN die Äußeren Orientierungen der einzelnen Bilder übergeben werden. Das ist allerdings nur dann sinnvoll, wenn mit den Bilddaten eine weitere photogrammetrische Prozessierung, z.B. Stereoauswertung geplant ist.*
- Datenträger**
- (1) Die Lieferung erfolgt auf **externer Festplatte / USB-Stick mit USB 3.0 Schnittstelle**. (*optional weiterer Kopien*)
 - (2) Datenträger gehen in das **Eigentum des Auftraggebers über / verbleiben im Eigentum des Auftragnehmers**.
 - (3) Die Festplatte, Verzeichnisse und Dateien sind ohne einschränkende Rechtevergabe zu sichern.
- Datenlieferung / Kachelung**
- (1) Es ist ein Gesamtergebnis inklusive aller Dokumentationen und Metadaten zu liefern.
 - (2) Teillieferungen erfordern die Absprache mit dem Auftraggeber.
 - (3) Die Daten sind in **x m x x m** großen Kacheln auf volle Meter im oben genannten Bezugssystem aufzubereiten. (*Hinweis: Wenn das Gebiet kleiner ist, kann auch auf eine Kachelung verzichtet werden.*)
 - (4) Das Orthophotomosaik ist im ***.tif / *.jpg**-Format zu liefern
 - (5) Das digitale Oberflächenmodell und das digitale Geländemodell sind im ***.tif** Format zu liefern
 - (6) Die 3D-Punktwolke ist im ***.las Format 1.2** zu liefern.

2.5 Bewertungsschema zur Ermittlung des wirtschaftlichsten Angebots

Im Anschluss an die Ausschreibung ist das wirtschaftlichste Angebot zu ermitteln. Welches Angebot am Ende den Zuschlag erhält, wird in einem zweistufigen Verfahren geklärt:

- a) Ist die Firma überhaupt geeignet (§§ 2 und 7 VOL)? Wenn diese Frage nicht positiv beantwortet werden kann, wird das Angebot ausgeschlossen. Zur Klärung der Eignung ist die berufliche Zuverlässigkeit, die finanzielle und wirtschaftliche sowie die technische Leistungsfähigkeit und Fachkunde zu beurteilen.
- b) Welche der verbleibenden Firmen am besten geeignet ist, wird nach im Verfahren bekanntgegebenen Kriterien bewertet.

Im Gegensatz zu dem Bewertungsverfahren zur Ausschreibung amtlicher digitaler Bildflüge (ADV 2015) ist bei der Ausschreibung digitaler UAS-Bildflüge eine größere Vielfalt an eingesetzter Technik zu erwarten, die sich zudem schnell weiterentwickelt. Das bedeutet, dass sich die nachfolgend genannten Bewertungskriterien noch in der Praxis bewähren müssen, ggf. modifiziert und erweitert werden müssen. Eine einfache Umrechnung in ein Punktesystem ist derzeit ebenfalls noch nicht möglich, da Praxiserfahrungen mit entsprechenden Ausschreibungen fehlen.

2.5.1 Bewertungskriterien

1. Informationen zur Auftrags- und Projektabwicklung

Diese lässt sich nur ungefähr über die bisherigen Referenzen abschätzen.

2. Informationen über UAS und digitales Aufnahmesystem

- Sicherheitsfeatures des UAS (Redundanz im Autopiloten, Rotoren etc., Notfallprozeduren, Automatisches Starten und Landen, Fallschirm, Notlandefunktion, Coming Home Funktion, ...). Beim Vergleich zwischen Flächenflüglern und Multicoptern ist immer zu berücksichtigen, dass im Falle einer fatalen Funktionsstörung (Motor- oder Servoausfall) Flächenflügler nicht einfach abstürzen, sondern in der Regel ziemlich kontrolliert gelandet werden können.
- Qualität der Kamera (Festbrennweite, keine bewegten Teile in der Optik (Bildstabilisierung oder Sensorreinigungssystem), Anzahl Pixel und Pixelgröße, RAW-Format ...)
- Kamerastabilisierung

3. Bildflugplanung

Die Vorlage einer detaillierten Bildflugplanung wird bereits zur Angebotsabgabe von den Anbietern gefordert, um die Umsetzung der technischen Vorgaben gemäß Leistungsbeschreibung zu beurteilen. Hieraus kann ein Eindruck über das Engagement und die fachliche Eignung gewonnen werden. Folgende Aspekte werden für eine Prüfung empfohlen:

Gebietsabdeckung

- Flughöhe entsprechend Bodenauflösung am niedrigsten Geländepunkt
- Längs- und Querüberdeckung am höchsten Geländepunkt
- Bodenauflösung
- Struktur und Format

4. Weitere Informationen aus dem Angebotstext

Das Angebot soll die Ausführung der geforderten Leistungen und das zugehörige Qualitätsmanagement umfassen. Hierbei sollten wesentliche Aussagen zu folgenden Punkten betrachtet und beurteilt werden:

- Erforderlicher Flugaufwand (Eine kürzere Befliegung, die entweder über eine Kamera mit einer größeren Auflösung oder einem schnelleren UAS realisiert werden kann, ist grundsätzlich besser, da sich die Licht- und Wetterverhältnisse in einem kurzen Zeitraum weniger ändern als in einem längeren.)
- Anfahrt und Logistik
- Zugriffsmöglichkeit auf die technische Ausrüstung (UAS, Piloten, RTK-GPS für Referenzmessung). (Es wird erwartet, dass ein Anbieter mit mehr als nur einem verfügbaren Komplettsystem (Flugzeug, GNSS/INS, Kamera, Besatzung) flexibler auf unerwartete und den Befliegungserfolg gefährdende Ereignisse reagieren kann. Die zu bewertenden Systeme müssen sich im Direktzugriff des Anbieters befinden. Die alleinige Möglichkeit der weiteren Hinzuziehung von Komplettsystemen über potenzielle Partner wird nicht bewertet.)
- Zugriffsmöglichkeit auf Postprocessing- und photogrammetrische Auswertesysteme
- Aussagen zu etwaigen Flugbeschränkungen
- Aussagen zu topographischen Besonderheiten und Einschränkungen

- Qualitätsmanagement incl. Maßnahmen zur Qualitätssicherung sowie die Art und Form der dabei entstehenden Ergebnisse
- Aussagen zum Lieferzeitraum. Der Bieter hat anzugeben, in welchem Zeitraum eine vollständige Lieferung des kompletten Datenpakets nach Abschluss der Flugleistung garantiert ist. Je schneller die Lieferung zugesagt wird, desto besser.

5. Bonus für regionalen Anbieter

Bei der Bewertung der Ausschreibungsunterlagen ist eine Bevorzugung regionaler Anbieter anzuraten, da diese im Allgemeinen flexibler auf das aktuelle Befliegungswetter reagieren können und somit auch in der Lage sind kleinere Wolkenlücken etc. auszunutzen. Schließlich basiert die Entscheidung die Flugkampagne zu beginnen immer auf einer Wetterprognose, die bekannter Maßen nicht immer so eintreten muss.

3 Fazit & Ausblick

Das vorgestellte Pflichtenheft für die Ausschreibung von UAS-Befliegungen ist in der Zukunft weiter zu entwickeln und an den technischen und regulatorischen Fortschritt anzupassen. Durch eine hoffentlich große Verbreitung der Ausschreibungsrichtlinie für öffentliche Ausschreibungen von UAS-Befliegungen werden auch die notwendigen Praxiserfahrungen gesammelt, um das Bewertungsschema quantifizieren zu können und dadurch auch die Transparenz für potentielle Auftragnehmer zu erhöhen.

4 Literaturverzeichnis

- AdV, 2015: Leitfaden zur Ausschreibung einer digitalen Luftbildbefliegung als Grundlage zur Herstellung von ATKIS®-DOP und stereoskopischen Auswertung. 30 S., <http://www.adv-online.de/AdV-Produkte/Standards-und-Produktblaetter/Standards-der-Geotopographie/>.
- BMVI, 2014: Kurzinformation über die Nutzung von unbemannten Luftfahrtsystemen. 17 S., https://www.bmvi.de/SharedDocs/DE/Publikationen/LF/unbemannte-luftfahrtsysteme.pdf?__blob=publicationFile).
- GRENZDÖRFFER, G., 2016: Unbemannte Flugroboter – aktuelle technologische Entwicklungen und die wichtigsten rechtlichen Rahmenbedingungen. 148. DVW-Seminar UAV 2016 – Vermessung mit unbemannten Flugsystemen, 18. und 19. Februar 2016, Bonn, Schriftenreihe des DVW, Band 82, Wißner-Verlag, Augsburg, 31-51.

Direkte Georeferenzierung in UAS-Anwendungen – Eine Beispielanwendung der Landesvermessung Baden-Württemberg

MICHAEL CRAMER¹, FELIX SABOW¹ & PETRA FÖRG²

Zusammenfassung: Auch in der Landesvermessung spielen technisch und wirtschaftlich optimierte Datenerfassungsmethoden und Auswerteprozesse eine entscheidende Rolle. Daher wird in dieser Studie ein Verfahren der direkten Georeferenzierung, ohne Verwendung von Passpunkten, in einer UAS-Anwendung untersucht. Das Gebiet einer Straßenbaumaßnahme wurde im März 2015 in einer Blockbefliegung mit einer Bodenpixelgröße (GSD) von 3 cm in einer mittleren Flughöhe von etwa 125 m mit dem MAVinci Sirius Pro System im Auftrag des Landesamts für Geoinformation und Landentwicklung Baden-Württemberg (LGL BW) erfasst. Die Auswertung der photogrammetrischen Daten inkl. der zur Verfügung gestellten GNSS-Trajektorie erfolgte mit zwei unabhängigen Softwarepaketen. Referenzdaten in Form von mit GNSS eingemessenen natürlichen Punkten bzw. kinematisch erfassten GNSS-Straßenprofilen erlauben eine Abschätzung der erzielten 3D-Objektpunktgenauigkeit aus der photogrammetrischen Auswertung. Neben der Einzelpunktgenauigkeit wird auch die Genauigkeit der 3D-Punktwolke analysiert. Die Einzelpunktgenauigkeit dieser dichten 3D-Punktwolke soll laut Anforderungen der Landesvermessung 1 dm betragen.

1 Einführung

Die Landesvermessung hat die Aufgabe, Veränderungen in der Landschaft zeitnah zu dokumentieren und so die amtlichen Geobasisdaten zu aktualisieren. Luftbilder sind für die Landesvermessung schon seit jeher die typische Datenquelle, aus denen diese Geobasisdaten abgeleitet werden können. Um eine hohe Aktualität der Geobasisinformationen gewährleisten zu können, müssen diese ständig fortgeführt werden. Daher besteht großes Interesse, die kostenintensive Erfassung der Ausgangsdaten, unter anderem durch eine Effizienzsteigerung bezüglich der Datenerfassung und Datenverarbeitung, zu optimieren (GÜTLINGER 2015). Die flexible Datenerfassung kleinräumiger Gebiete mit einem UAS stellt somit eine gute Alternative oder Erweiterung zu klassischen Bildflügen dar, welche der Aktualisierung von Geobasisinformationen dienen.

Das Landesamt für Geoinformation und Landentwicklung Baden-Württemberg (LGL BW) hat seit 2012 mehrfach UAS für Aufgaben der Landesvermessung eingesetzt. Im Jahr 2012 wurde eine umfassende Kampagne zur Evaluierung der UAS-gestützten Bilderfassung im Rahmen eines Flurneuerordnungsverfahrens gemeinsam mit dem Institut für Photogrammetrie (ifp) und dem Institut für Flugmechanik und Flugregelung (ifr) der Universität Stuttgart durchgeführt (CRAMER et al. 2013). Dieses sogenannte Hessigheim-Projekt (benannt nach der benachbart liegenden

¹ Institut für Photogrammetrie, Universität Stuttgart, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, E-Mail: michael.cramer@ifp.uni-stuttgart.de, felixsabow@outlook.com

² Landesamt für Geoinformation und Landentwicklung Baden-Württemberg (LGL), Referat 53, Fernerkundung, Postfach 4065, 76025 Karlsruhe, E-Mail: petra.foerg@lgl.bwl.de

Ortschaft) diene zugleich als Pilotprojekt für alle Landesvermessungsbehörden in Deutschland zur Abschätzung des Potenzials von UAS im Rahmen der Landesvermessung. In der Studie konnte nachgewiesen werden, dass bei guter Blockgeometrie und dichter Passpunktverteilung 3D-Objektpunktgenauigkeiten (RMS aus Kontrollpunktdifferenzen) im Bereich von 2 – 3 cm in der Lage bzw. ca. 5 cm in der Höhe erzielt werden können. Damit können 3D-Einzelpunkte subpixelgenau bestimmt werden. Die korrespondierende Bodenpixelgröße betrug 5 – 8 cm.

Im nachfolgenden Jahr 2013 wurden diese Untersuchungen in einer neuen Kooperation zwischen dem LGL BW, ifp und dem kommerziellen UAS-Anbieter senseFly fortgesetzt. Zur Dokumentation einer Hangrutschung im Bereich Talheim-Mössingen, Schwäbische Alb, sollte mit dem senseFly eBee System eine 3D-Oberflächenpunktswolke photogrammetrisch bestimmt werden, um durch Vergleich mit dem vorliegenden Laserscanner-DGM die aktuelle Veränderung des Geländes zeitnah zu dokumentieren. Aufgrund der Unwegsamkeit des Geländes war keine alternative terrestrische Datenerfassung möglich. Der Flug wurde im November 2013 unter sehr anspruchsvollen Befliegungsbedingungen (starke Windböen, große Helligkeitsunterschiede, niedriger Sonnenstand) durchgeführt. Der jahreszeitlich späte Befliegungstermin ergab sich aus der Anforderung, in der vegetationsfreien Periode zu fliegen, um möglichst ungehinderte Sicht auf das Gelände zu haben. Die Hangrutschung liegt zu großen Teilen in dicht bewaldetem Gebiet. Im Unterschied zu dem Hessigheim-Projekt lag der Schwerpunkt in dieser Untersuchung auf der Analyse der Genauigkeiten der abgeleiteten Oberflächenmodelle, nicht auf der Einzelpunktbestimmung.

Die flächenhaften Vergleiche der photogrammetrischen Punktwolken mit den Referenzdaten aus der Laserbefliegung zeigen gute Übereinstimmungen im zentralen, von Passpunkten kontrollierten Bereich. Die Abweichungen liegen bei ca. -10 cm – +20 cm, wobei gewisse Anteile auch durch die Vegetation beeinflusst sind. Der Vergleich erfolgte zwischen dem Geländemodell aus Laserscanning und dem photogrammetrisch abgeleiteten Oberflächenmodell. Weitere Analysen anhand von DOM-Vertikaldifferenzen an 19 signalisierten Referenzpunkten lieferten je nach Orientierung bzw. DSM-Verfahren Genauigkeiten (RMS) von 5 – 10 cm, entsprechend 0.6 – 1.2 pix bei einer mittleren GSD von 8 cm. Diese Höhengenaugigkeit im Bereich von etwa 1 pix GSD korrespondiert gut mit den von HAALA (2013) genannten Genauigkeiten photogrammetrischer Punktwolken, die im Rahmen des EuroSDR-Tests für großformatige photogrammetrische Kamerasysteme ermittelt wurden. Das vom LGL BW angestrebte Genauigkeitspotenzial der photogrammetrischen DOM-Bestimmung liegt ebenfalls bei etwa 10 cm.

Die Ergebnisse der UAS-Projekte Hessigheim bzw. Talheim zeigen, dass die Qualität der 3D-Einzelpunkt- bzw. Oberflächenmodellbestimmung durchaus mit der Genauigkeit der klassischen photogrammetrischen Kamerasysteme bzw. Laserscanner verglichen werden kann. Das gilt zumindest für derartige, eher kleinräumige Befliegungsprojekte mit sehr dichter Bildüberdeckung und bei guter bis sehr guter Passpunktverteilung, was einen gewissen Mehraufwand für die Signalisierung bzw. Einmessung dieser Punkte mit sich bringt. Dieses steht in einem Gegensatz zu der photogrammetrischen Praxis, bei der großflächige Auswertungen mit minimaler Passpunktanzahl realisierbar sind. Die präzise GNSS-Bahnbestimmung liefert genaue Messungen der Projektionszentrumskoordinaten, die als gewichtete Beobachtungen in die Auswertung eingehen und damit die Passpunkte vom Gelände „in die Luft“ verlagern. Im Idealfall sind so komplett passpunktfreie Auswertungen möglich, sofern keine zusätzlichen Offset- und Drift-

Effekte zu korrigieren sind und es sich um einen gut verknüpfenden Block mit mehreren sich überlappenden Flugstreifen handelt (ACKERMANN 1997).

In wie weit dieses effiziente Anwendungsszenario nun auch mit kommerziellen UAS realisierbar ist, soll im Rahmen dieses Beitrags anhand eines durchgeführten Beispielprojekts analysiert werden. Dazu hat das LGL BW eine Firma beauftragt, das Gebiet einer Straßenbaumaßnahme im März 2015 mit einem sog. RTK (real-time-kinematic) UAS zu befliegen. Ein solches UAS verfügt neben der bildgebenden Kamera über weitere Sensorik für eine präzise GNSS-Trajektorienbestimmung, die ähnlich wie im bemannten Bildflug eine (quasi) passpunktfreie Auswertung der Bilddaten erlauben soll.

Dieses soll im Folgenden vorgestellt werden. Im nachfolgenden Abschnitt wird RTK UAS diskutiert mit Schwerpunkt auf dem MAVinci Sirius Pro System, welches für diese Anwendung gewählt wurde. Anschließend wird das vom LGL BW beauftragte Befliegungsprojekt (Abschnitt 3) erläutert. Die photogrammetrische Auswertung und Präsentation der Ergebnisse erfolgt in Abschnitt 4. Abschnitt 5 diskutiert dann vor allem auch die seitens der Landesvermessung angestrebte Verwendung UAS-gestützter 3D-Punktwolken zur Fortführung des bestehenden Geländemodells. Der Beitrag schließt mit einer Zusammenfassung.

2 RTK UAS

Unter RTK versteht man im Allgemeinen die kinematische, differenzielle Positionierung in Echtzeit unter Verwendung der Phasenmessung mit Auflösung der Integer-Phasenmehrdeutigkeiten. Für die Echtzeitanwendung ist zwingend eine permanente Verbindung zu einer GNSS-Basisstation notwendig, die am Boden auf einem bekannten Punkt aufgestellt wird und permanente Korrekturdaten an das Flugsystem sendet. Auf diesen Datenlink kann verzichtet werden, wenn die Auswertung der GNSS-Daten im Post-Processing erfolgt, was neben der Echtzeit-Filterung auch die nachträgliche Glättung der Daten und damit Steigerung der Genauigkeit erlaubt. Die präzisen Positionen stehen dem Träger dann allerdings nicht in Echtzeit zur Verfügung. Die ausgewerteten Messdaten sind identisch, die Auswertung findet lediglich zeitversetzt, nach der Landung statt. Dieses Verfahren wird mit post-processed kinematic Auswertung (PPK) bezeichnet, was für photogrammetrische Anwendungen oftmals ausreichend ist. Da in beiden Fällen die Qualität der Bahnbestimmung mit zunehmender Basislänge abnimmt, ist die Referenzstation im Idealfall im Testgebiet platziert, oder es werden – sofern möglich – Korrekturdienste eines GNSS-Permanentreferenzstationsnetzwerks genutzt.

Standard UAS-Flugplattformen sind für die Echtzeitnavigation im Regelfall mit einfachen GNSS-Empfängern ausgestattet, die nur auf einer Frequenz arbeiten. Dieses Standard-GNSS erlaubt unter Auswertung der Codephase Echtzeitpositionsgenauigkeiten im Bereich von wenigen Metern, was für die Navigation des Trägers ausreichend ist. Für die präzise Messung der Projektionszentrumsordinate des zu orientierenden Sensorsystems sind jedoch Positionsmessungen von 10 cm oder besser notwendig, je nach angestrebtem Genauigkeitsbereich bzw. Anwendung. Dieses Genauigkeitspotenzial wird durch die Verwendung von Mehrfrequenz-GNSS-Empfängern realisiert, ähnlich wie sie auch im Rahmen der klassischen geodätischen Basislinienmessung bzw. für die photogrammetrische GNSS-gestützte Aerotriangulation zum Einsatz kommen. Die dort verwendeten Systeme können aufgrund von Gewicht und Größe aber

nicht in UAS mit einem maximalen Gesamtabfluggewicht von unter 5 kg eingesetzt werden. Erst mit der Miniaturisierung erschließen sich diese Anwendungsszenarien. BÄUMKER et al. (2013) berichten über ein Zwei-Frequenz UA-System auf Basis des TOPCON B110-Empfängers sowie einer Zwei-Frequenz-Antenne Maxtena M1227 HCT. Das eigentliche GNSS-Empfänger-Board hat Abmessungen von 40x55x10 mm und ein Gewicht <20 g. Das Gesamtgewicht inkl. Antenne liegt bei ca. 100 g. Vergleichbare Systeme finden sich mittlerweile in diversen kommerziellen UAS-Lösungen. Zur Unterscheidung von der regulären UAS-Basisausstattung wird oftmals das Kürzel RTK zur Kennzeichnung real-time-kinematic GNSS-Funktionalität ergänzt. Beispiele dafür sind die UAS-Flächenflieger senseFly ebee RTK bzw. C-Astral Bramor rTK. Andere Hersteller ergänzen alternative Kürzel, unter denen sich die präzise GNSS-Bahnbestimmung verbirgt, wie z.B. Trimble UX5 HP (high precision) oder das hier eingesetzte MAVinci Sirius Pro (professional) System. Auch Rotorflügler wie z.B. der Hexakopter Aibotix X6 Version 2 sind mit RTK-GNSS ausgestattet. Die Hersteller versprechen für diese Systeme durchgängig GNSS-Positionsgenauigkeiten von bis zu wenigen Zentimetern.

Bei dem MAVinci Sirius Pro System handelt es sich um fixed-wing UAS. Starrflügler oder Flächenflieger haben gegenüber Rotorflüglern den Vorteil, dass sie deutlich höhere Flugzeiten erreichen und somit größere Flächen in einem Flug abdecken können was Anwendungen der Landesvermessung vermutlich entgegenkommt. Das Fluggerät selber besteht größtenteils aus Elapor, einem leichten aber gleichzeitig stabilen Hartschaum der Firma Multiplex. Die Spannweite beträgt 1.63 m, die Länge 1.22 m. Das Abfluggewicht liegt inklusive der möglichen 550 g Nutzlast bei 2.7 kg. Bildflüge mit dem UAS sind bis zu einer Windstärke von 50 km/h (Böen bis 65 km/h) durchführbar und es kann bei Außentemperaturen von -20° C bis +45° C, sowie bei Regen eingesetzt werden. Die Fluggeschwindigkeit liegt bei etwa 65 km/h (alle Angaben laut Hersteller). Mittlerweile stattet MAVinci die Systeme mit der Kamera X-M1 von Fujifilm in Kombination mit einem 18 mm f/2.0 Festbrennweitenobjektiv aus. Die kompakte spiegellose Systemkamera verfügt über einen 16 Megapixel APS-C Bildsensor mit den Abmessungen 23.6 x 15.6 mm² und wiegt 450 g inklusive Objektiv. Die im Rahmen des hier untersuchten Projekts erfassten Daten wurden allerdings noch mit einer Panasonic GX-1 und einem 14 mm f/2.5 Pancake Objektiv (entsprechend 35 mm Brennweite im Kleinbildformat) aufgenommen. Die Panasonic GX-1 verfügt ebenfalls über 16 Megapixel, die aber auf einem Micro Four Thirds Sensor angeordnet sind. Mit einer Fläche von 17.3 x 13.0 mm² ist dieser etwa 40 % kleiner als der Sensor der aktuellen XM-1, was sich vor allem auch auf die radiometrische Qualität der Bilder (schlechteres Signal-Rausch-Verhältnis aufgrund kleinerer Pixelgröße) auswirken kann.

Das Sirius Pro System verfügt über ein RTK-fähiges GNSS-Board der Firma Topcon. Laut PRZYBILLA et al. (2015) handelt es sich dabei um das B110 Empfänger-Board, die entsprechenden Produktinformationen von MAVinci spezifizieren einen 100 Hz Topcon L1/L2-Zwei-Frequenzempfänger der Genauigkeitsklasse 2 cm.

3 LGL BW Projekt Neutrassierung Landstraße L 583

Die Landstraße L 583 im Neckar-Odenwald-Kreis verbindet die Ortschaften Großeicholzheim und Seckach. Der bisherige Streckenverlauf hatte unübersichtliche Kuppen und Wannen, enge Kurven und im Bereich der Randstreifen keinen festen Untergrund. In den Jahren 2011/2012 wurde

deshalb die Straße über eine Länge von ca. 2 km vollkommen neu gebaut und im Bereich einer scharfen Kurve über eine Länge von ca. 700 m seitlich verlegt. Begleitend wurden für ca. 5 km die Entwässerungsanlagen hergestellt oder neu gestaltet. Der neue Straßenverlauf und die Geländeänderungen durch die Baumaßnahmen erforderten die Aktualisierung der Geobasisdaten des LGL, insbesondere des DGM. Zur zeitnahen Dokumentation hat das LGL BW den geänderten Abschnitt der Landstraße im März 2015 durch eine klassische UAS-Blockbefliegung erfassen lassen. Um den Verlauf der Straße entsprechend zu berücksichtigen, wurde die Befliegung auf zwei Bildflüge aufgeteilt, mit einem Überlappungsbereich, der in beiden Flügen erfasst wurde. Bei einer mittleren Flughöhe von 125 m über Grund (resultierende Bodenpixelgröße 3 cm) wurden im ersten Flug in 26 Flugstreifen 313 Bilder aufgenommen (Vorwärtsüberdeckung >85%, lange Bildseite in Flugrichtung, Querüberdeckung ca. 70%). Nach der Änderung der Hauptflugrichtung um 45° wurden anschließend weitere 515 Aufnahmen in 17 Flugstreifen gemacht. Das erfasste Gebiet hat eine Größe von etwa 0,9 km² und variiert in der vertikalen Ausdehnung um ungefähr 75 m.

Planung und Durchführung und auch die photogrammetrische Auswertung der Befliegung wurden komplett von der durch das LGL BW beauftragten Firma erledigt. Für die Auswertung konnte diese Firma lediglich auf zwei bekannte Punkte zurückgreifen, die als Standorte für die



Abb. 1: Übersicht über das erfasste Gebiet mit Verteilung der Referenzpunkte (rot: Standpunkte für die GNSS-Referenzstationen, gelb: unabhängige Kontrollpunkte (natürliche Punkte))

terrestrischen GNSS-Referenzstationen (je Block eine Station) genutzt wurden und im Befliegungsgebiet liegen. Weitere bekannte Objektpunkte standen der Firma für die Auswertung nicht zur Verfügung, sodass sich die 3D-Lagerung des Blocks bzw. Georeferenzierung der Bilder ausschließlich auf die direkt gemessenen GNSS-Projektionszentrumskordinaten und die Verknüpfung der Bilder im Bildverband stützt.

Als Referenzwerte für die hier durchgeführten Genauigkeitsuntersuchungen dienen die vom LGL BW zur Verfügung gestellten Daten. Hierzu zählen neben dem landesweiten DGM aus der Laserscannerbefliegung sowohl mit statischem GNSS eingemessene natürliche Punkte (13

natürliche Punkte), wie auch kinematisch erfasste GNSS-Straßenprofile (im Wesentlichen entlang der L 583 und Nebenstraßen). Für diese Trajektorie wurde ein GNSS-Empfänger auf dem Dach eines Fahrzeugs befestigt, das die entsprechenden Straßen abfuhr; die Trajektorie wurde somit mittels kinematischem GNSS erfasst. Die Höhe der Antenne vom Boden wurde mit dem Meterstab gemessen und als konstanter Offset zur Reduktion der Vertikalkoordinaten auf das Straßenniveau angebracht. Diese Korrektur ist nicht streng, da das Fahrzeug durch die Fahrbewegungen federt. Daher ist die Höhengenaugigkeit der Trajektorie auf nur ca. 10 – 15 cm festzusetzen. Die statisch bestimmten Referenzpunkte werden mit einer Genauigkeit von 2 – 3 cm angenommen. Bei diesen Kontrollpunkten handelt es sich um Straßenmarkierungen, Kanaldeckel und speziell angebrachte Farbmarkierungen, die in den Bildern klar zu identifizieren sind. Die Abb. 1 zeigt das Untersuchungsgebiet (abgeleitetes Orthophoto aus der UAS-Befliegung) mit den zur Verfügung stehenden (statischen) Referenzpunkten und den beiden Referenzstationsstandpunkten.

4 Photogrammetrische Auswertung

Wie bereits erwähnt ist das Projekt vom LGL BW komplett extern vergeben worden und die beauftragte Firma hat die aus der Befliegung abzuleitenden Endprodukte (Orthophoto (GSD 3 cm), dichte Punktwolke und Oberflächenmodell (Rasterweite 6 cm)) geliefert. Die Auswertung der Daten wurde, wie von der UAS-Herstellerfirma MAVinci empfohlen, mit dem Programmpaket agisoft PhotoScan Pro durchgeführt. MAVinci bietet hierfür eine spezielle Schnittstelle (Plugin) an, um die gemessenen präzisen GNSS-Projektionszentren entsprechend ihrer Genauigkeit zur Lagerung des Blocks zu verwenden. Es standen der Firma wie in Abschnitt 3 genannt nur zwei bekannte Geländepunkte (Referenzstationskoordinaten) für diese externe Auswertung zur Verfügung. Damit beruhen (lt. Sprachgebrauch von MAVinci) die Ergebnisse auf einer passpunktfreien Auswertung. Die beiden im Befliegungsgebiet stehenden und in den Bildern sichtbaren GNSS-Referenzstationen werden aber zumindest als Kontrollpunkte bzw. zur a priori Offsetkorrektur (z.B. Offset Antenne-Kameraprojektionszentrum) genutzt. Aus der Praxis der GNSS-gestützten Aerotriangulation ist bekannt, dass oftmals konstante Offsets pro Bildverband auftauchen – wobei die genauen Ursachen dafür zunächst keine Rolle spielen –, die aber mit nur einem Kontroll- bzw. Passpunkt vollständig aufgedeckt und korrigiert werden können.

Zu Vergleichszwecken wurde eine alternative Auswertung der Daten mit der Software Trimble UASMaster durchgeführt. Die direkt gemessenen GNSS-Projektionszentrumskordinaten flossen dazu mit einer Gewichtung von 3 cm in die Bündelausgleichung ein. Ansonsten erfolgte die Auswertung auch passpunktfrei. Die Genauigkeiten beider Auswertung ergeben sich nachfolgend aus dem Vergleich mit den vorliegenden Referenzdaten.

4.1 Einzelpunktgenauigkeit

Zur Beurteilung der absoluten Genauigkeit der 3D-Punktbestimmung wurden die 13 Kontrollpunkte sowohl in PhotoScan, als auch in UASMaster manuell gemessen, um so die Abweichungen zu den Referenzwerten bestimmen zu können. Die resultierenden Kontrollpunktdifferenzen sind in der Abb. 2 für die 13 Punkte angegeben. Sie zeigen die absolute Genauigkeit der Einzelpunktbestimmung. Die Differenzen sind getrennt für die Auswertungen mit PhotoScan (Quadratsymbollinien) bzw. UASMaster (Dreiecksymbollinien) gegeben. Die

Abbildung zeigt die hohe Übereinstimmung der Ergebnisse. Die Differenzen in horizontaler Richtung sind praktisch mittelwertfrei, in der Vertikalkoordinate lässt sich für beide Auswertungen ein mittlerer Offset von ca. 7.5 cm ablesen. Die Streuungen liegen bei 2.2 cm, 2.3 cm, 3.3 cm für PhotoScan und 3.6 cm, 3.7 cm, 2.6 cm für UASMaster, jeweils angegeben für Ost-, Nord- und Vertikalkoordinate. Die PhotoScan Genauigkeit ist geringfügig besser in den horizontalen Komponenten, wobei UASMaster in der Höhenkomponente leicht besser abschneidet. Die resultierenden RMS-Werte in der Vertikalen sind aber praktisch wieder gleich und liegen bei 8.2 cm für PhotoScan bzw. 7.9 cm für UASMaster. Bezogen auf die Bodenpixelgröße von etwa 3 cm wird in der Lage eine Genauigkeit von etwa 1 pix erreicht, in der Höhe liegt der RMS-Wert bei etwa 2.5 – 3 pix.

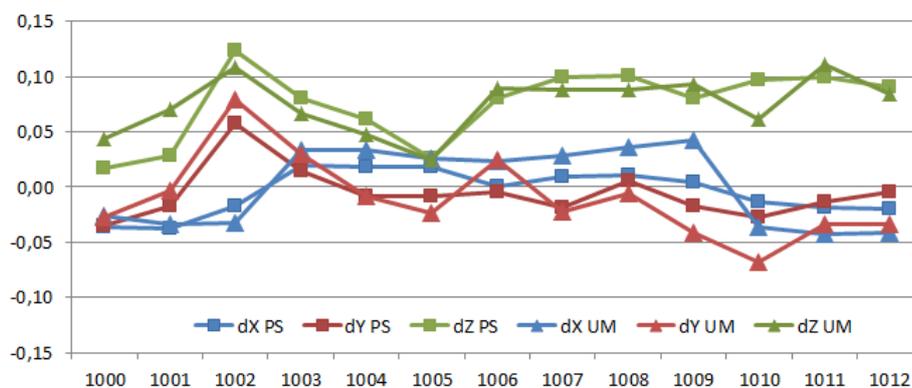


Abb. 2: Kontrollpunktdifferenzen [m] aufgetragen über der Punktnummer (PS: PhotoScan, UM: UASMaster)

4.2 Flächendeckende DHM-Genauigkeit

Die bisher gezeigten Resultate erlauben die Beurteilung der punktuellen Genauigkeit der 3D-Punktbestimmung an den 13 vorliegenden Kontrollpunkten. Da es im Rahmen dieses Projekts eher auf die flächenhafte (Höhen-)Genauigkeit ankommt, werden Vergleiche an den kinematisch erfassten Geländepunkten bzw. mit dem Laserscanner-DGM vorgenommen. Von den 405 kinematisch erfassten Geländepunkten liegen 379 im durch die Bildflüge erfassten Gebiet. An diesen Punkten können die Höhen aus den vorliegenden Oberflächenmodellen interpoliert und

Tab. 1: Vertikaldifferenzen an den Vergleichspunkten der kinematischen Trajektorie

	PhotoScan	UASMaster
Mittelwert [m]	-0.085	-0.060
Max. Abw. [m]	0.410	-0.468
Std. Abw. [m]	0.052	0.115
RMS [m]	0.100	0.129



Abb. 3: Vertikaldifferenzen an den Vergleichspunkten der kinematischen Trajektorie (DOM PhotoScan)

anschließend mit den Höhen aus der kinematischen GNSS-Trajektorie verglichen werden. Tab. 1 gibt einen Überblick über die berechneten Mittelwerte, maximalen Abweichungen, Standardabweichungen und RMS-Werte. Die Abb. 3 zeigt für das DOM aus der PhotoScan Auswertung die Vertikaldifferenzen in Abhängigkeit von der Lage der Punkte. Je größer der dargestellte Kreisdurchmesser, umso höher die Differenz. Bei der Interpretation der Ergebnisse muss die geringere Genauigkeit der kinematischen Trajektorie berücksichtigt werden, daher wird hier auch nur von Vergleichs- nicht Referenzpunkten gesprochen. Die größeren Abweichungen treten vor allem in den durch Abschattung (Vegetation, Geländeeinschnitt) beeinflussten Bereichen auf, in den Gebieten mit freier Sicht sind die Differenzen kleiner. Dieses weist deutlich auf Ungenauigkeiten in der kinematischen GNSS-Trajektorie hin. Im äußersten Blockrandgebiet tauchen ebenfalls größere Differenzen auf, was sich mit der geringeren Überdeckung und damit auch geringeren Genauigkeit der 3D-Punktvolke aus der photogrammetrischen Auswertung begründen lässt. Die Genauigkeit der PhotoScan bzw. UASMaster Auswertungen ist nahezu identisch, wenn man die RMS-Werte vergleicht. In Bezug auf die Streuung ist das Ergebnis der Punktvolke aus PhotoScan aber besser. Im Vergleich zu den Ergebnissen der 3D-Einzelpunktbestimmung ist die Höhengengenauigkeit insgesamt schlechter, und liegt eher im Bereich von 3 – 4 pix (RMS) was aber auch durch die geringere Qualität der kinematischen Vergleichspunkte verursacht ist.

Die Abb. 4 zeigt die Differenz der beiden Oberflächenmodelle aus der PhotoScan bzw. UASMaster Auswertung (DOM PhotoScan minus DOM UASMaster). Da es sich um die gleiche Oberfläche handelt, abgeleitet aus dem gleichen Bildmaterial, sollten die Differenzen klein sein.

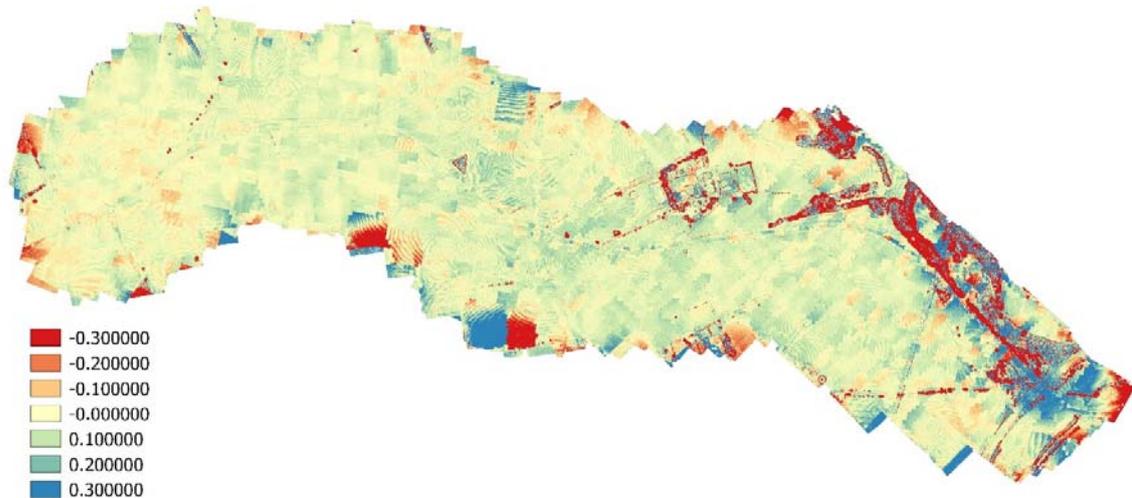


Abb. 4: Differenz der DOMs PhotoScan – UASMaster

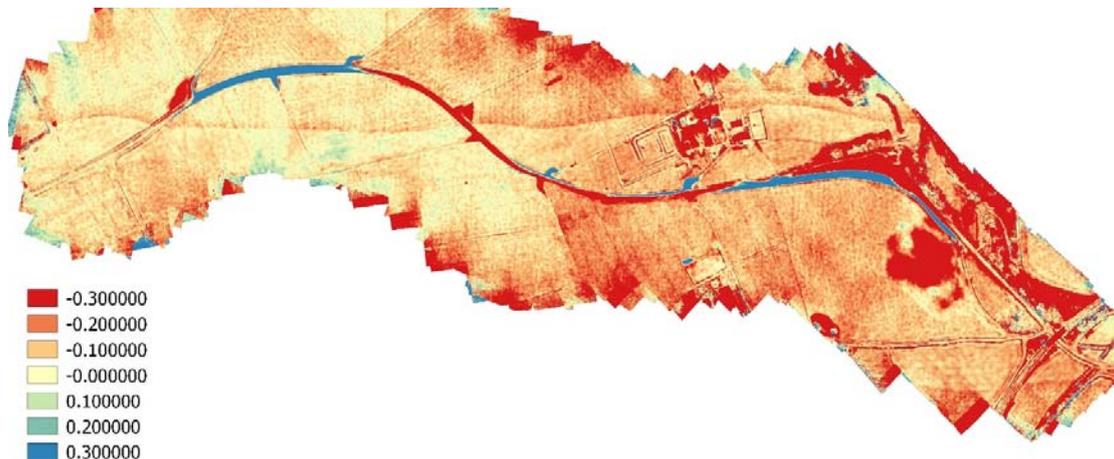


Abb. 5: Differenz zwischen Laserscanner DGM – PhotoScan DOM

Die Grafik zeigt Übereinstimmung überwiegend im Bereich von -10cm – 10cm. Größere Differenzen tauchen vor allem am Blockrand und im Bereich von Vegetation und an Gebäudekanten auf, was aber zu erwarten ist. Es lassen sich jedoch zusätzliche systematische Effekte in den Differenzen erkennen. Es handelt sich um eine Art Kachelung, die offensichtlich auf die Behandlung der jeweiligen Modelle in der Punktwolkengenerierung zurückzuführen ist. Ein solcher Effekt taucht nur in dem von UASMaster gerechneten DOM auf. Die dichte Punktwolke wird zunächst modellweise erzeugt. Diese modellweise generierten 3D-Punkte werden dann zu der gesamten Punktwolke zusammengeführt. Offensichtlich wird in PhotoScan dabei eine andere Anpassung der Modelle durchgeführt. Die in der Aerotriangulation für alle Bilder bestimmten Orientierungen / Unbekannten sind optimal für den gesamten Block geschätzt. Lokal, d.h. bezogen auf einzelne Bildpaare, können sich aber noch Restparallaxen zeigen. Auch dieser Effekt ist aus der klassischen AT bekannt: Wenn die Orientierungen aus einer Bündeltriangulation für das nachfolgende Stereoplotting übernommen werden sollen, kann es notwendig sein, Restparallaxen in dem Modell durch eine Nachorientierung über die zuvor

bestimmten Verknüpfungspunkte zu korrigieren. COLOMINA (2002) bezeichnet dieses mit dem „point interface“ (Punkt-Schnittstelle) in der Photogrammetrie. Trimble arbeitet aktuell an einer verbesserten Angleichung der Einzelmodelle zur Gesamtpunktwolke.

In der Abb. 5 wird das DOM aus der PhotoScan Auswertung verglichen mit dem Geländemodell aus der 2002 durchgeführten Laserscannerbefliegung der Landesvermessung. Da die Höhendaten zu unterschiedlichen Epochen erfasst wurden – das Laser-DGM dokumentiert den Zustand vor der Neutrassierung der Landstraße – müssen sich in dieser Differenz die Veränderungen der Landschaft durch die Straßenbaumaßnahme niederschlagen, was deutlich zu erkennen ist. Differenzen ergeben sich nicht nur an der eigentlichen Straßentrasse, im östlichen Bereich lässt sich unmittelbar neben der Straße auch ein größerer Bereich identifizieren, der offensichtlich aufgefüllt wurde. Darüber hinaus enthält die Differenz natürlich auch den systemimmanenten Unterschied zwischen Geländeoberfläche (Laser-DGM) und sichtbarer Oberfläche (photogrammetrisches DOM aus PhotoScan), vor allem für die mit Vegetation bestandenen Bereiche. Die gezeigten Differenzen sind tendenziell eher negativ, was so zu erwarten war, da hier das Oberflächenmodell von dem „tiefer liegenden“ Geländemodell abgezogen wurde. Die Differenzen liegen im Bereich von -10 – -30 cm. Auffällig sind sich quer durch das Gelände ziehende Versätze zwischen den Höhenmodellen, deutlich zu erkennen im mittleren Bereich des Untersuchungsgebiets und ebenfalls im süd-östlichen Bereich. Hierbei handelt es sich um Restsystematiken aus der Streifenanpassung der damaligen Laserscannerauswertung, die nun durch die hohe Genauigkeit des photogrammetrischen Oberflächenmodells aufgedeckt werden können. Derartige Effekte sind dem LGL BW auch aus anderen Projekten bekannt.

5 Diskussion der Ergebnisse

Die Analyse der Oberflächenmodelle aus der UAS-Befliegung bestätigt das hohe Potenzial der photogrammetrischen Ableitung von dichten Punktwolken. Bei einer Bodenpixelgröße GSD von 3 cm kann das resultierende Oberflächenmodell mit 3 – 4 pix Genauigkeit realisiert werden. Die Genauigkeit der Oberflächenpunkte ist jedoch schlechter als die Vertikalgenauigkeit der 3D-Einzelpunktbestimmung. Die 3D-Einzelpunktgenauigkeit (RMS) liegt bei etwa 1 pix (horizontal) und ca. 3 pix (vertikal). Im Vergleich zu dem Hessigheim-Projekt (CRAMER et al. 2013) sind die bodenpixelbezogenen RMS-Werte schlechter. Dies liegt am Einfluss der individuellen Blockgeometrien und ggf. auch an der kleineren GSD im L 583 Projekt. Hinzu kommt der Umstand, dass bei der Verwendung von Passpunkten am Boden die Neupunkte im gegebenen Passpunkttrahmen interpoliert werden, wohingegen die direkte Georeferenzierung (GNSS-gestützte Aerotriangulation ohne Passpunkte) immer einen extrapolierenden Charakter hat. Dadurch können sich theoretische Restfehler (z.B. in der Gesamtsystemkalibrierung) leichter in der Objektpunktbestimmung niederschlagen. PRYZBILLA et al. (2015) untersuchen ebenfalls das Potenzial des MAVinci RTK Systems, berücksichtigen aber in ihren Auswertungen zusätzliche 4 Geländepasspunkte. Diese RTK-Konfiguration mit 4 Passpunkten erreicht für das vorgestellte Projekt die gleiche Genauigkeit wie eine klassische Bündeltriangulation mit 18 Passpunkten. Hervorzuheben ist, dass sich im L 583 Projekt keine systematischen Aufwölbungen des Oberflächenmodells erkennen lassen. Dieses zeigt den stabilisierenden Einfluss der präzisen GNSS-Projektionszentrumskoordinaten, die den Block vollständig kontrollieren, auch wenn

(außer der Referenzstation) keine Passpunktinformation am Boden verwendet wird. Im Gegensatz zu den bisherigen klassisch passpunktbasierten Untersuchungen in Hessigheim bzw. Talheim-Mössingen bedeutet dieses eine signifikante Effizienzsteigerung, da keinerlei Aufwand in eine zusätzliche Einmessung und Signalisierung von Passinformation am Boden erforderlich ist. Diese Effizienzsteigerung ist aber immer mit dem Extrapolationscharakter der direkten Georeferenzierung abzuwiegen. Die hier erzielten Vertikalgenauigkeiten erfüllen mit etwa 10 cm (RMS) jedoch die gestellten Anforderungen der Landesvermessung.

Für die Integration der UAS-basierten Oberflächenmodelle in das existierende Geländemodell – die eigentliche Motivation für derartige UAS-Projekte seitens des LGL BW – müssen nun aber noch verschiedene, teils manuelle Arbeitsschritte durchgeführt werden, um die Standards und die Qualität des digitalen Geländemodells des LGL BW gewährleisten zu können. In der mittels digitaler Bildkorrelation erzeugten dreidimensionalen Punktwolke werden alle sichtbaren Landschaftselemente abgebildet, auch Vegetation und weitere, das DGM störende Objekte. Zudem können bei automatischen Korrelationsverfahren die allgemein bekannten Probleme, z.B. in Schattenbereichen und texturschwachen Gebieten, entstehen. Die erzeugte Punktwolke wird daher zunächst automatisch in Bodenpunkte und Nichtbodenpunkte gefiltert (siehe Abb. 6).

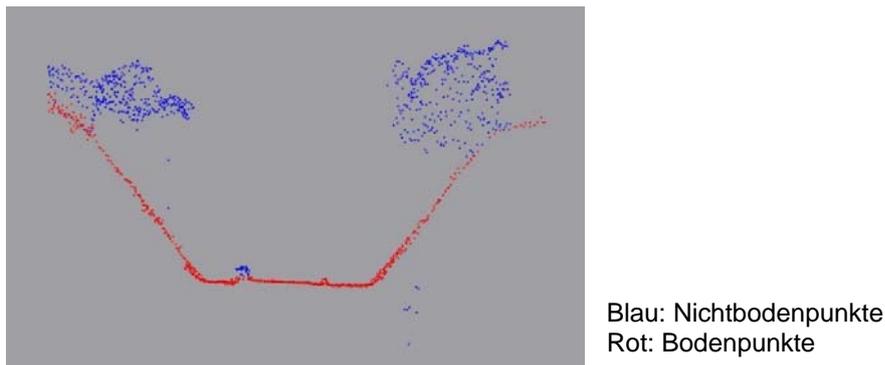


Abb. 6: Gefilterte Punktwolke (Klassifikation der UAS-Oberflächenpunkte)

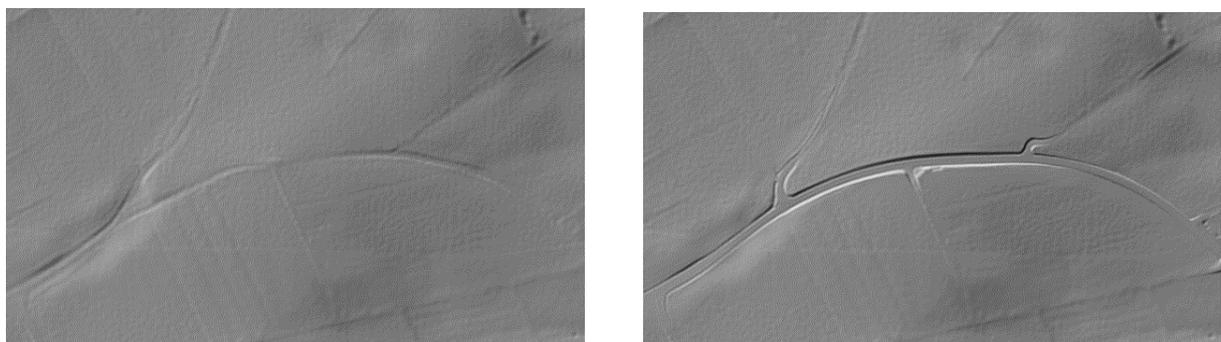


Abb. 7: Laserscanner-DGM (Zustand vorher, links) modifiziert durch die Daten aus der UAS-Befliegung (Zustand nachher, rechts), Ausschnitt aus dem Projekt L 583

Im Stereomodell kann durch den Vergleich der Bodenpunkte aus der aktuellen UAS-Befliegung mit dem veralteten DGM der veränderte Bereich abgegrenzt werden, in dem dann die weitere Bearbeitung erfolgt. Beim LGL BW liegt das Digitale Geländemodell in einer Gitterweite von 1 m landesweit flächendeckend vor. Somit wird auch aus den korrelierten Bodenpunkten der UAS-Befliegung für den veränderten Bereich eine regelmäßige, lückenlose 1 m-Gitterstruktur mittels linearer Prädiktion abgeleitet. In diesem Bereich wird der alte Datenbestand durch den neuen ersetzt. Dabei wird beim Übergang vom alten zum neuen DGM in einem 5 m breiten Streifen eine gleitende Anpassung gerechnet, so dass bei eventuellen geringen Höhendifferenzen keine Stufe entsteht.

Um die Qualitätsanforderungen an das DGM zu gewährleisten, wird das Ergebnis im Stereomodell kontrolliert und eventuell an den oben genannten Schwachstellen der Punktwolke manuell korrigiert. Dabei wird das DGM-Gitter durch Re-Interpolation optimal an den Geländeverlauf angepasst. Das Resultat dieser Anpassungen ist in Abb. 7 für einen Ausschnitt des Testgebiets gezeigt. Ein Nachteil der Stereoauswertung von UAS-Flügen gegenüber Standardbildflügen mit großformatigen Flächensensoren ist die Vielzahl und die unregelmäßige Anordnung der kleinformatischen Bilder. Ein daraus resultierender ständiger Wechsel der Stereomodelle wirkt hier bei der Bearbeitung sehr erschwerend. Der Einsatz stabilisierter Kameraaufhängungen könnte sich daher eventuell positiv auswirken. Hierzu sind weitere Untersuchungen geplant.

6 Zusammenfassung

Die hier vorgestellte Projektstudie belegt erneut das große Potenzial der UAS-gestützten Oberflächenmodellgenerierung für Anwendungen der Landesvermessung. Die RTK-Projektionszentren ermöglichen eine sehr effiziente Datenerfassung. Die Lagerung des photogrammetrischen Bildverbands erfolgt durch die präzisen GNSS-Projektionszentrumskoordinaten. Die bekannten Referenzstationskoordinaten im Befliegungsgebiet ermöglichen darüber hinaus eine etwaige Korrektur von blockkonstanten Offsets. Das photogrammetrische DOM kann so mit ca. 10 cm (RMS) bestimmt werden, was den Genauigkeitsanforderungen der Landesvermessung entspricht. Die photogrammetrischen Höhendaten können nach einer Filterung und Anpassung zur Fortführung des bestehenden landesweiten Geländemodells verwendet werden.

Literatur

- ACKERMANN, F., 1997: Geo-Kodierung ohne Passpunkte. GIS - Geo-Informationen-Systeme **10** (2), 28-32. <http://www.ifp.uni-stuttgart.de/publications/1997/Ack97.pdf> (letzter Zugriff: 28.04.2016)
- BÄUMKER, M.; PRZYBILLA, H.-J. & ZURHORST, A., 2013: Enhancements in UAV flight control and sensor orientation. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **40** (1/W2), 33-38. <http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-1-W2/33/2013/isprsarchives-XL-1-W2-33-2013.pdf> (letzter Zugriff: 28.04.2016)

- COLOMINA I., 2002: Modern sensor orientation technologies and procedures. Integrated Sensor Orientation, OEEPE Official Publication No. **43**, Heipke C., Jacobsen K. & Wegmann H. (Eds.), 59-70. <http://bono.hostireland.com/~eurosdr/publications/43.pdf> (letzter Zugriff: 28.04.2016)
- CRAMER, M., HAALA, N., ROTHERMEL, M., LEINSS, B. & FRITSCH, D., 2013: UAV-gestützte Datenerfassung für Anwendungen der Landesvermessung – das Hessigheim-Projekt Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **22**, E. Seyfert (Hrsg.), 33. Wissenschaftlich-technische Jahrestagung DGPF, 27. Februar – 1. März 2013, 450-469.
- GÜTLINGER, M. 2015: UAS-Einsatz zur DGM-Aktualisierung, in Bundesanstalt für Gewässerkunde (Hrsg.): Wasserstraßenbezogene geodätische Anwendungen und Produkte der Fernerkundung. Kolloquium am 18./19. November 2014 in Koblenz. – Veranstaltungen 2/2015, Koblenz, März 2015, 11-19, http://doi.bafg.de/BfG/2015/Veranst2_2015.pdf (letzter Zugriff: 28.04.2016)
- HAALA, N., 2013: The Landscape of Dense Image Matching Algorithms. Photogrammetric Week '13, Ed. D. Fritsch, Wichmann, Berlin/Offenbach, 271-284. <http://www.ifp.uni-stuttgart.de/publications/phowo13/240Haala-new.pdf> (letzter Zugriff: 28.04.2016)
- PRZYBILLA, H.-J., REUBER, C., BÄUMKER, M. & GERKE, M., 2015: Untersuchung zur Genauigkeitssteigerung von UAV-Bildflügen. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **24**, T. Kersten (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 16.-18. März 2015 an der Universität zu Köln, Tagungsband auf CD-ROM, 45-54. http://www.dgpf.de/src/tagung/jt2015/proceedings/papers/06_DGPF2015_Przybilla_et_al.pdf (letzter Zugriff: 28.04.2016).

UAV Sensor Orientation with Pre-calibrated Redundant IMU/GNSS Observations: Preliminary Results

PHILIPP CLAUSEN¹, MARTIN REHAK¹ & JAN SKALLOUD¹

Zusammenfassung: In unserem Beitrag stellen wir eine Drohne vor, welche für eine hochpräzise Kartographie mit Positions- und Orientierungssensoren ausgestattet ist. Der Einsatz der Sensoren als absolute und relative Messung kann die Genauigkeit des Mappings erhöhen, falls die Sensoren korrekt kalibriert sind. Dies zeigen wir an einem Fallbeispiel, in dem eine Blockstruktur befliegen wurde, in welcher Passpunkte (Ground Control Points) nur lokal angelegt wurden. Eine redundante Sensor-Konfiguration zeigt dann auf, wie gut und genau die Kalibrierung und dessen Resultat sind.

1 Introduction

1.1 General Concepts

Unmanned aerial vehicles (UAVs) are an important tool for surveyors, construction engineers and scientists worldwide. Thanks to their affordability and recent advances in guidance, autonomy, and easiness of use, they spread among wide public. Accurate georeferencing plays a key role in successful terrain reconstruction and artefact-free orthophoto generation; it is particularly of great relevance in applications such as corridor mapping, road and pipeline inspections, and mapping of large areas with homogeneous surface structure, e.g. forests or agricultural fields.

The presented experiments focus on new approaches of aerial control. The accuracy in the aforementioned mapping scenarios can be quantitatively improved by new approaches of aerial position and relative attitude control. By differencing the observations, the process of system calibration is simplified and some of the systematic errors in the observations are ruled out. This is interesting in the context of mapping with micro aerial vehicles (MAVs) that have limited payload capacity and flight endurance. The asset of aerial relative control is further increased by using redundant attitude observations.

1.2 Challenge

We investigate the proposed methods of georeferencing on a real data set acquired with a custom-built MAV fixed-wing plane. It is equipped with a geodetic-grade GNSS receiver and a redundant Inertial Measurement Unit (R-IMU) that is based on Micro-Electro-Mechanical-Systems (MEMS). This R-IMU was specifically developed to be employed on MAVs. To achieve accurate results, a state-of-the-art sensor calibration is presented here and performed. This procedure estimates the deterministic errors present in the sensors. The camera system is pre-calibrated for the lens-distortions, whereas a self-calibration is performed for the principal point and the distance. We show methods of handling the redundant observations. First, several independent sets of exterior orientation (EO) parameters are calculated and then used in a Bundle Block Adjustment

¹ École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, E-Mail: {philipp.clausen, martin.rehak, jan.skaloud}@epfl.ch

(BBA). These observations are treated in an absolute and relative manner, giving rise to different solutions.

We evaluate the proposed calibration and orientation methods during a real mapping project. The accuracy of the mapping solution is assessed via a set of independent check points that were accurately surveyed. We confirm that thanks to on-board observations of UAV-position and attitude a centimetre mapping accuracy is achievable using proper pre-calibration and processing techniques in scenarios where classical methods of sensor orientation fail (i.e. deliver non-acceptable lower accuracy by a factor superior to 10).

1.3 Paper Structure

The photogrammetric system used on the MAV is described in Section 2. Section 3 details the inertial navigation system with its calibration and properties. Section 4 introduces the custom-built MAV and the dataset we collected with it. The final results are reported and discussed in Section 5. The conclusion in Section 6 then rounds up the paper.

2 Photogrammetric System

The imaging system is composed of the Sony NEX 5R camera with a 16 mm Sony lens. It is mounted next to the in-house developed GECKO4Nav-Board (KLUTER 2012), which is shown in Figure 1. The specifications of the installed IMUs are explained in detail in Section 3. The IMUs and the camera system are rigidly mounted so that their relative orientation stays the same over the whole period of testing. This is assured by the means of a stable carbon structure.

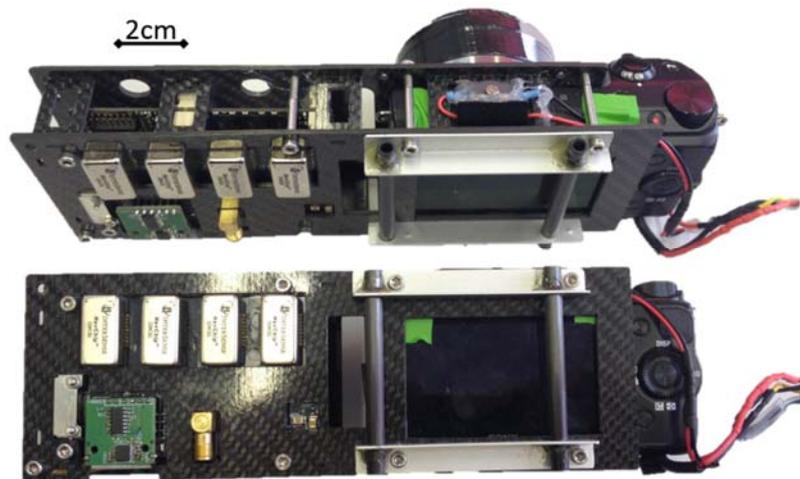


Fig. 1: GECKO4Nav-Board on the left side and the Sony Camera on the right side rigidly connected through a carbon structure.

Precise timestamping of the IMUs to the GPS-time is assured by a geodetic grade GNSS receiver, which is directly connected to the GECKO4Nav-Board. The PPS signal serves as synchronization among all the components. It synchronizes the individual IMUs to the common clock. The camera is synchronized to the GNSS receiver by sending a pulse every time the shutter opens. Thus, the

exact time of each picture is very well known. The weight of the photogrammetric payload shown in Figure 1 is around 550 g.

The IMUs mounted next to the photogrammetric system are MEMS-IMUs with an interesting noise characteristics (INTERSENSE 2015). The most pertinent error-characteristics are shown in Table 1. The increased noise characteristics (in comparison to a tactical grade IMU) are compensated with their weight of only 6 g. This makes them perfectly usable in MAVs, although some electronics have to be designed for synchronization, timestamping, and power control.

Tab. 1: Properties of one Navchip-IMU.

Property	Value
size	24x14x9 mm
weight	6 g
power requirement	0.2 W
frequency	250 Hz
gyro bias	10 °/hr

3 Redundant IMU

There are several advantages of using multiple IMUs at the same time (GUERRIER et al. 2012). One of the advantages is *redundancy*: if one sensor fails or delivers faulty observations, then the other sensor can detect it and compensate the error. Another advantage is related to the value of the noise level. By combining multiple sensors in a *synthetic* IMU the overall noise is first determined and then reduced (WAEGLI et al. 2010). This is advantageous, as the noise-level may vary due to vibrations and its exact value is required for the filtering/smoothing with GNSS data. Before using the R-IMU concept the sensors have to be calibrated for deterministic errors.

A total number of four Navchip-IMUs are mounted next to each other (Figure 1). The IMU-sensor-system is placed at different attitudes where static measurements l are taken for several seconds. The amplitude $\|l\|$ is compared to the reference gravity signal $\|g\|$ taking advantage of the compensation process explained by Syed et al. (SYED et al. 2007). The bias b , the mean scale factor S , and the non-orthogonality θ between the individual sensor axes x, y and z can be estimated. The found properties are then applied to compensate the measurements of the accelerometers l according to the formula:

$$\begin{pmatrix} l_x \\ l_y \\ l_z \end{pmatrix}_{\text{compensated}} = \begin{bmatrix} 1+S_x & 0 & 0 \\ \theta_{xy} & 1+S_y & 0 \\ \theta_{xz} & \theta_{yz} & 1+S_z \end{bmatrix} \begin{pmatrix} l_x \\ l_y \\ l_z \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} \quad (1)$$

Figure 2 shows the norm of the accelerometer measurements before and after calibration for one IMU. The wrong norm in different positions is due to the fact that each axis has an individual bias, and that the axis are coupled through the non-orthogonality. The corrected signal can then be used in the next processing steps explained in the following subsection. These procedure of calibration is used separately on the four IMUs.

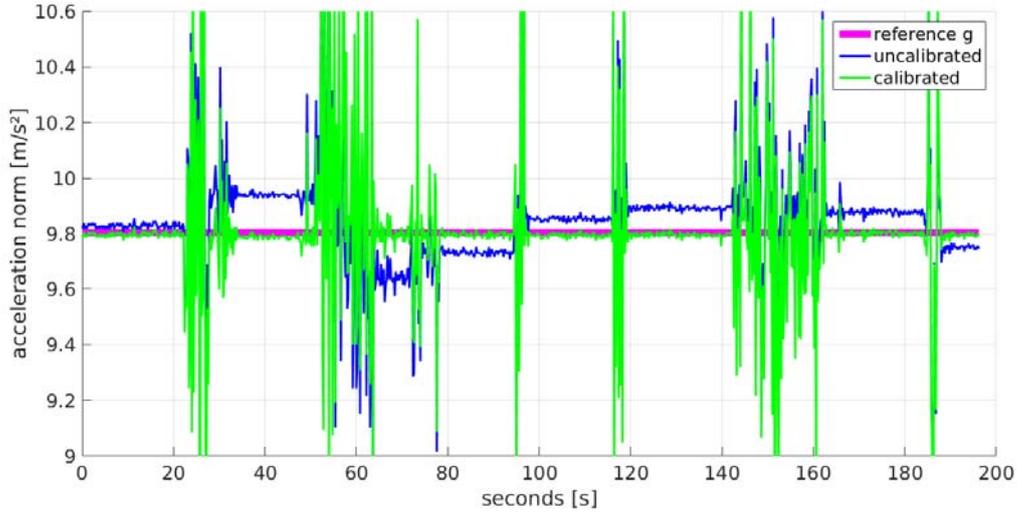


Fig. 2: Example dataset before (blue) and after (green) calibration of an IMU showing the differences in the acceleration norm at different attitudes. The reference gravity value g is shown in red colour.

The resulting biases from the calibration process are shown in Table 2, where one can see that each IMU has its own error characteristics. Therefore, a thorough calibration is an absolute necessity for each sensor as the results are not similar at all. Wrongly calibrated accelerometer sensors can for instance lead to wrong attitude determination.

Tab. 2: Sensor error bias example is shown for all four IMUs.

Property	IMU 0	IMU 1	IMU 2	IMU 3
b_x [μg]	6764	5945	11397	-997
b_y [μg]	16225	4198	1538	28220
b_z [μg]	-871	-2507	-2058	-5372

4 MAV Platform and Test Data

4.1 MAV Platform

A custom-developed MAV platform was used to collect data for this study. The fixed-wing platform has a wingspan of 1630 mm and a length of 1170 mm. The maximal payload capacity is around 800 g. The operational weight varies between 2200-2800 g. The aircraft is made of expanded polypropylene foam (Figure 3). The flexible nature of the construction material makes the platform resistant to damage. The plane is easy to assemble and repair with ordinary hobby-grade tools. The cost of the system components is significantly lower with respect to size and endurance compared with other platforms (MAVINCI 2015). The endurance with 600 g payload is approximately 40 minutes. The plane is controlled by a Pixhawk autopilot that has been intensively developed over the last few years (MEIER et al. 2012).



Fig. 3: Custom-built MAV platform flying.

Tab. 3: Summary of acquired data for the flight of Figure 4.

Property	Value
camera	Sony Nex 5R
lens	Sony 16 mm
flying height	120-150 m
mean GSD	4.5 cm/pix
overlap forward/lateral	80-60 %
number of Photos	207
number of GCPs	5
number of ChPs	15

4.2 Test data

This study was conducted over agricultural fields and roads. The testing area has a size of approximately 70 ha and is equipped with 20 permanent markers whose location is accurately surveyed. Regarding the accuracy assessment, 15 points were used as independent Check-Points (ChPs) and 5 points were used as ground-control-points (GCPs). The GCPs cover only 1/3 of the area. This is done to simulate a case where the majority of the mapping-area is badly accessible. The data set contains data from a block consisting of 7 parallel lines and 7 lines perpendicular to them, flown in two separate flight heights (Figure 4). This separation is important when performing camera self-calibration since it helps with decorrelating internal and external camera parameters. The most pertinent facts about the flown mission are summarized in Table 3.

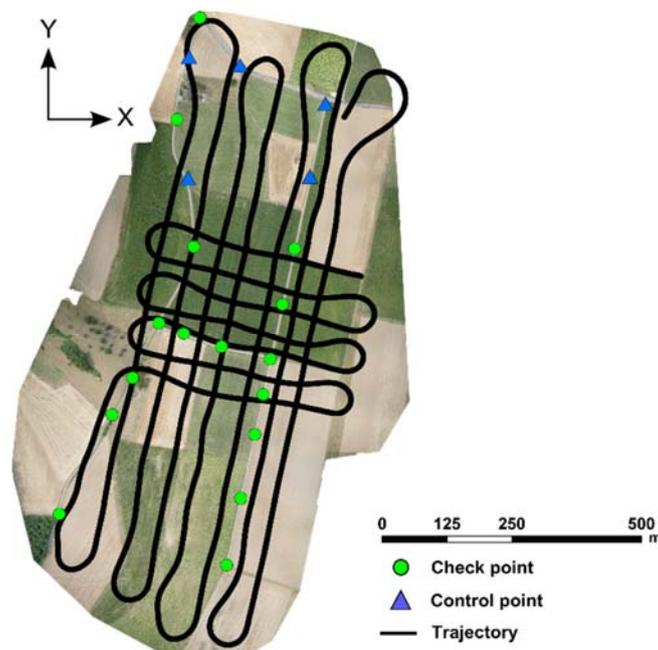


Fig. 4: Flown mission in Vufflens, Switzerland in 2015. The GCPs (blue triangles) are “badly” placed in order to simulate inaccessible area for GCPs. The black flight line shows the perpendicular flying path at two different heights.

4.3 Processing strategy

The data recorded during the flights were pre-processed in a way similar to mature mapping systems. The GNSS data was processed in a professional software package. Thanks to the precise time synchronization between the camera and the GNSS receiver, the exact acquisition time of each image is directly known. The calculated antenna positions were subsequently fused in an Extended Kalman filter with the IMU at a frequency of 250 Hz. The GNSS/INS-derived attitude was then corrected for the boresight misalignment, which was calibrated in previous experiments (REHAK & SKALOUD 2015). Nevertheless, the inclusion of relative attitude eliminates the need of such calibration. The image observations of the tie-points were automatically measured in the images using Pix4D mapper (PIX4D 2015). The observations of GCPs and ChP were obtained manually using its rayCloud engine.

The observations were then processed in a custom-developed Bundle Block Adjustment (BBA) software. The latter was developed because the offer of commercially available software for BBA allowing redundant and relative position and attitude observations is very limited. In addition, custom implementation gives us full control of observation stochastic modelling that is important when using accurate aerial control. Several BBA projects were created with different inputs.

5 Assessment of Approach: Redundancy in EO observations

Redundancy of EO parameters is very interesting with MAVs. Particularly of interest is the concept of relative orientation as it eliminates the need for boresight calibration of each sensor and decreases the noise level of the whole system. The redundancy in IMU data can be treated in several ways. A first option is the joint utilization of several IMUs as explained by WAEGLI et al. (2008). The second option is based on computing several trajectories and inputting them as additional observations into the BBA. However, the introduction of absolute attitude control would require separate boresight calibration of all IMUs as well as handling the correlations among them. Here, the relative orientation has a big potential since the boresights are eliminated and the residual correlations become very small. Several BBA projects were computed with different inputs and kinds of aerial control. The first test includes the following projects using always 5 GCPs:

- (1) Indirect Sensor Orientation (SO)
- (2) Integrated SO (ISO) with *uncalibrated* IMU using absolute position and attitude
- (3) ISO with *precalibrated* IMU using absolute position (A_p) and absolute attitude (A_a)
- (4) ISO with *precalibrated* IMU using absolute position (A_p) but relative attitude (R_a)

The results of this first test is given in Table 4. The Indirect SO solution (1) can be massively improved by using the absolute position and absolute attitude of the camera system (2). This solution (2) is using the uncalibrated IMU and can then be further improved by utilizing the precalibrated IMU (3). The RMS value decreases slightly for the Z-axis. However, the impact would be larger for situations with fewer image observations. The results in the Z-axis can be improved even further by employing the approach (4) with the relative attitude update. This test evidences the influence of the calibrated IMU data on the relative attitude updates. Detailed investigations on relative attitude observations are presented in REHAK & SKALOUD (2016).

Tab. 4: Accuracy assessment at independent ChPs. The block has 5 GCPs and 15 ChPs. The IMU number “0” was used with absolute position (Ap), absolute attitude (Aa), and relative attitude (Ra).

Dataset	Accuracy					
	Mean ChP [mm] [px]			RMS ChP [mm] [px]		
	X	Y	Z	X	Y	Z
(1) Indirect SO	68 1.5	8 0.2	-664 14.8	16 0.4	145 3.2	1171 26
(2) ISO Ap Aa <i>uncalib</i>	13 0.3	26 0.6	74 1.6	26 0.6	37 0.8	87 1.9
(3) ISO Ap Aa <i>precalib</i>	14 0.3	25 0.6	64 1.4	28 0.6	35 0.8	78 1.7
(4) ISO Ap Ra <i>precalib</i>	14 0.3	24 0.5	45 1.0	27 0.6	36 0.8	65 1.4

The second test was conducted to verify if the IMU calibration combined with the subsequent use of attitude updates is consistent on all four IMUs of the Gecko4Nav-Board. The following tests were made:

- (1) ISO with precalibrated IMU 0 using absolute position (Ap) and absolute attitude (Aa)
- (2) ISO with precalibrated IMU 1 using absolute position (Ap) and absolute attitude (Aa)
- (3) ISO with precalibrated IMU 2 using absolute position (Ap) and absolute attitude (Aa)
- (4) ISO with precalibrated IMU 3 using absolute position (Ap) and absolute attitude (Aa)

The results are provided in Table 5. The results are all similar in the horizontal axes (around 0.6 px in X and 0.8 px in Y). Larger variations in the RMS are present in Z-direction (1.6 px to 2.2 px) among individual solutions. In comparison to indirect SO (that has around 26 px) the improvement factor is around 10. Small differences can arise from the Kalman Filter solution, since the same stochastic model has been used for each of the IMUs.

Tab. 5: Accuracy assessment at independent ChPs. The block has 5 GCPs and 15 ChPs. All precalibrated IMUs “0” to “3” were used with absolute position (Ap) and absolute attitude (Aa).

Dataset	Accuracy					
	Mean ChP [mm] [px]			RMS ChP [mm] [px]		
	X	Y	Z	X	Y	Z
(1) ISO Ap Aa IMU 0	14 0.3	25 0.6	64 1.4	28 0.6	35 0.8	78 1.7
(2) ISO Ap Aa IMU 1	7 0.2	26 0.6	85 1.9	21 0.5	37 0.8	98 2.2
(3) ISO Ap Aa IMU 2	8 0.2	25 0.6	66 1.5	26 0.6	35 0.8	82 1.8
(4) ISO Ap Aa IMU 3	6 0.1	27 0.6	56 1.2	23 0.5	37 0.8	73 1.6

6 Conclusion and perspectives

We presented the advantages of using position and attitude aerial control over the indirect sensor orientation. The importance of attitude increases with a decreasing number of image observations, and it depends on the texture of the mapping area and on an unfavourable image geometry (i.e. low overlap or corridor). The pre-calibration of low-cost IMUs is recommended since it improves the accuracy in the Z-direction. Relative attitude updates mitigate the residual misalignment and eliminate the need for precise boresights between camera and IMU. Depending on the demanded accuracy, areas with difficult access for placing GCPs can be skipped because few GCPs are sufficient in the ISO.

We will further investigate different methods of introducing redundant IMU measurements directly into the BBA.

7 References

- GUERRIER, S., WAEGLI, A., SKALLOUD, J. & VICTORIA-FESER, M.-P., 2012: Fault detection and isolation in multiple mems-imus configuration. *IEEE Transactions on Aerospace and Electronic Systems* **48**, 2015-2031.
- KLUTER, T., 2012: GECKO4NAV Technical Reference Manual Revision 1.0.
- MAVINCI, 2015: MAVinci – Unmanned Aerial Systems. URL <http://www.mavinci.de/>. Accessed 2015-12-10.
- MEIER, L., TANSKANEN, P., HENG, L., LEE, G., FRAUNDORFER, F. & POLLEFEYS, M., 2012: PIXHAWK: A micro aerial vehicle design for autonomous flight using onboard computer vision. *Autonomous Robots* **33**, 21-39.
- INTERSENSE, 2015: Intersense Navchip. URL <http://www.intersense.com/pages/16/246/>. Accessed 2015-12-10.
- PIX4D, 2015: Pix4Dmapper. URL <http://pix4d.com/>. Accessed 2015-12-10.
- REHAK, M. & SKALLOUD, J., 2015: Fixed-Wing Micro Aerial Vehicle for Accurate Corridor Mapping. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2** (1/W1), 23-31.
- REHAK, M. & SKALLOUD, J., 2016: Applicability of new Approaches of Sensor Orientation to Micro Aerial Vehicles. *The International Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences ISPRS ICWG III/I Annals of ISPRS Congress*, Prague.
- SYED, Z.P., AGGARWAL, P., GOODALL, C., NIU, X. & EL-SHEIMY, N., 2007: A new multi-position calibration method for MEMS inertial navigation systems. *Measurement Science and Technology* **18**, 1897-1907.
- WAEGLI, A., GUERRIER, S. & SKALLOUD, J., 2008: Redundant MEMS-IMU integrated with GPS for performance assessment in sports. *Proceedings of the IEEE/ION PLANS 2008*, Monterey, CA, USA.
- WAEGLI, A., SKALLOUD, J., GUERRIER, S. & PARES, M., 2010: Noise reduction and estimation in multiple micro-electro-mechanical inertial systems. *Measurement Science and Technology* **21**, 065201-065212.

Ableitung von Eisen(II, III)oxid in Fließgewässern mittels Multispektraldaten

CHRISTOPH ULRICH¹, LUTZ BANNEHR¹ & ANGELA LAUSCH²

Zusammenfassung: Aufgrund der weltweiten Bergbauaktivitäten in unmittelbarer Nähe von Seen und Flüssen stellt die Verockerung (Verschlammung durch Eisen(III)hydroxid) von Gewässern ein latentes globales Umweltproblem dar. Der Verockerungsprozeß entsteht durch eisenhaltiges Mienenabwasser und Grundwasserableitungen während der Bergbauaktivitäten.

Zur Ableitung von Eisen(II, III)oxiden in Fließgewässern in Tagebauregionen werden Spektralindizes herangezogen, die Indikatoren für die Konzentration von Eisen(II) und Eisen(III) darstellen. In der Literatur wurden diese Indizes bislang zur Bodenanalyse herangezogen. Als Datenbasis für die Multispektralanalyse dienen Daten vom Landsat 8 OLI. Im November 2015 ist infolge eines Dammbrochs in Brasilien eine hohe Eisenoxidkonzentration im Fluss „Rio Doce“ eingeschwemmt worden. Die bisherigen Ergebnisse zeigen, dass die Spektralindizes durchaus auf Fließgewässer übertragbar sind und einen guten Indikator für hohe Eisenoxidkonzentrationen darstellen.

1 Einleitung

Weltweit führen Bergbauaktivitäten in unmittelbarer Nähe von Seen und Flüssen zu Umweltproblemen. Durch die Absenkung des Grundwasserspiegels, der Lagerung des Abraums und den Bergbauabflüssen kommt es zu einem erhöhten Eintrag gelöster Metalle (Fe, Al, Mn, Cu) und Säure (SÁNCHEZ 2008). Diese Eintragungen führen zu einem Artensterben im Gewässer. Besonders dramatisch war dies im November 2015 in Brasilien zu sehen, wo große Mengen an Mienenabwasser aufgrund zweier Dammbroche den Fluss „Rio Doce“ verseuchten (BRÜHWILLER 2015).

Im geringeren Ausmaß sind die Folgen des Tagebaus auch in der deutschen Lausitz anhand dem Fluss „Spree“ und dessen Zuflüsse zuerkennen (AKTIONSBÜNDNIS KLARE SPREE E.V. 2013). In beiden Fällen hat sich der Fluss ockerartig verfärbt, was auf einen hohen Eisengehalt schließen lässt.

Unter normalen und natürlichen Bedingungen ist Eisen in vielen Gewässern vorhanden, jedoch in geringerer Konzentration. So beträgt der Gesamteisengehalt in Flüssen etwa 0,5 – 1 ppm und im Grundwasser ca. 100 ppm. Im Vergleich dazu beträgt der Eisengehalt im Meereswasser nur etwa 0,001 – 0,003 ppm (LENNTECH 2015).

Elementares Eisen sowie zahlreiche Eisenverbindungen sind unter normalen Umständen wasserunlöslich (LENNTECH 2015). Ausnahmen bilden das reduzierte, zweiwertige Eisen (Eisen II (Fe^{2+})) und das dreiwertige Eisen (Eisen(III) (Fe^{3+})). Während Eisen(II) nur unter reduzierten

¹ Hochschule Anhalt, Institut für Geoinformation und Vermessung, Bauhausstraße 8, D-06846 Dessau, E-Mail: [c.ulrich, l.bannehr]@afg.hs-anhalt.de

² Helmholtz-Zentrum für Umweltforschung GmbH – UFZ, Department Landschaftsökologie, Permoserstraße 15, D-04318 Leipzig, E-Mail: angela.lausch@ufz.de

Bedingungen mit neutralen pH-Wert gelöst vorliegt, ist es Eisen(III) nur in sauren Gewässern (ENDER et al. 2012).

Durch Hydrolyse von Eisen(III) entsteht Eisen(III)hydroxid. Dieses fällt bei einem neutralen pH-Wert des Gewässers als Schlamm aus und führt zu einer ockerartigen Verfärbung des Gewässers (Uhlmann et al. 2013). Dieses Phänomen wird als Verockerung bezeichnet (Ministerium für Wirtschaft und Energie des Landes Brandenburg 2015). Verockerung resultiert unter anderem aus den Eisenfrachten der Bergbaue.

Im Allgemeinen beinhaltet die Bestimmung der Schwermetallkonzentration in Gewässern mehrere Arbeitsschritte:

1. Systematisches Sammeln von Gewässerproben,
2. Laboranalysen und
3. Interpolation der punktuellen Ergebnisse (CHOE et al. 2008).

Dies ist jedoch zeit- und kostenintensiv. Eine Alternative bietet der physikalisch-analytische Ansatz der multi- und hyperspektralen Fernerkundung. In der Vergangenheit wurde versucht, mittels Methoden der Fernerkundung, flächendeckend kostengünstige Algorithmen zur Ableitung der Wasserinhaltsstoffe zu entwickeln (CHOE et al. 2008). Die ersten Verfahren entstammen der Bildspektroskopie in den Achtzigerjahren des letzten Jahrhunderts. Gleichzeitig erfolgte die Entwicklung der ersten hyperspektralen Fernerkundungssensoren, die sich seitdem rasant weiterentwickelt haben und besonders für die Identifikation von Schwermetallen im Boden eignen (MELENDENZ-PASTOR et al. 2011).

Im Weiteren wird die Anwendung verschiedener Spektralindizes zur Ableitung von Eisen(II) und Eisen(III) auf das Gebiet vom „Rio Doce“ angewendet. Bislang wurden diese Indizes für Bodenanalysen herangezogen. Der vorliegende Fall zeigt, dass sich Spektraloperationen auch sehr gut auf verockerte Fließgewässer anwenden lassen. Zur Beurteilung der Anwendbarkeit der Indizes werden der Fluss „Rio Doce“ mittels der Sensoren vom Landsat 8 OLI analysiert.

2 Untersuchungsregion - Fluss „Rio Doce“ in Brasilien

Der Fluss „Rio Doce“ war ein artenreicher Fluss, bis im November 2015 zwei Staumauern eines Eisenerzbergwerkes brachen. Die ausgelöste Schlammlawine ergoss sich in den Fluss, der die Abflüsse der Eisenerzmine 660 km bis an die Atlantikküste transportierte (UPADEK 2016). Die Abflüsse des Bergbaus enthielten verschiedene Schwermetalle, beispielsweise Arsen, Aluminium, Blei, Kupfer und Eisenoxide (EISENHAMMER 2015; FLÖRKE 2015). Folglich kam es zu einem großen Artensterben sowie zur rotbraunen Verfärbung des Flusses (UPADEK 2016). Ein Überblick über das betroffene Gebiet ist in Abb. 1 dargestellt.

Zur Ableitung von Eisenoxiden wurden für das Flussgebiet entlang des „Rio Doce“ betroffene Gebiet im Zeitraum der Naturkatastrophe Landsat 8 – Satellitendaten genutzt. Andere Satellitendaten wie Sentinel 2 waren für diesen Zeitraum nicht verfügbar. Die Analysen wurden zum einen an der betroffenen Mine selbst und zum anderen an einem Staudamm westlich der Stadt „Colatina“ durchgeführt.



Abb. 1: Darstellung der Untersuchungsgebiete in Brasilien; Datenquelle: Openstreetmap; Hintergrundkarte: Bing Aerial

Die untersuchte Eisenerzmine liegt nahe der Bergbaustadt „Mariane“. Dort brachen zwei Dämme der Rückhaltebecken aufgrund des Abraums und der Abwässer. Dies löste eine Schlammlawine aus, die sich in den Fluss „Rio Doce“ ergoss (FLÖRKE 2015). Der Bereich rund um die Eisenerzmine ist interessant, um die Methodiken anhand der trockenen Abraumsedimente zu prüfen. Ferner ist der Stausee aufgrund der großen Gewässerfläche geeignet. Da die Bodenpixelauflösung vom Landsat 8 OLI 30 m beträgt, ist die normale Flussbreite für eine genaue Analyse nicht ausreichend.

3 Methodik

Die hier aufgeführten Methoden wurden schon mehrfach von den Geologen zur Ableitung von Eisenoxiden im Boden eingesetzt. Im Folgenden wird versucht, mit diesen Methoden Rückschlüsse auf Eisen(II)- bzw. Eisen(III)oxid in Fließgewässern zu schließen.

Im Allgemeinen lassen sich Wasserinhaltsstoffe mittels Fernerkundungsmethoden nur im sichtbaren Bereich des Lichtes (VIS) bestimmen (GEGE 2004). Wie in Abb. 2 zu erkennen ist, absorbiert das Wasser zum großen Teil das solare Spektrum, Ausnahme bildet jedoch der Spektralbereich des VIS von 390 nm bis 780 nm (KELLER 1999).

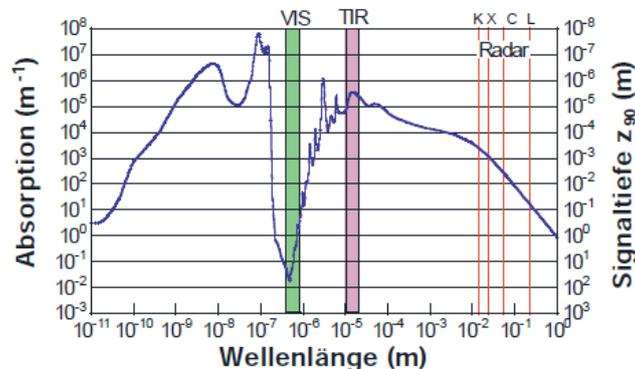


Abb. 2: Absorption von reinem Wasser und die daraus berechnete Dicke der Wasserschicht (GEGE 2004: 73).

Aufgrund des Reflektionsvermögens elektromagnetischer Energie von Gewässern, können im VIS verschiedene Gewässerinhaltsstoffe fernerkundet werden. Diese werden allgemein in drei Klassen untergliedert: Phytoplankton, gelöste organische Substanzen und nicht-organische Schwebstoffe (Gege 2004).

Aufgrund der Inhaltsstoffe reflektieren alle Gewässer unterschiedlich und werden deshalb in Case-I - (offene Ozeane) und Case-II - (Küsten- und Binnengewässer) Gewässer untergliedert (Morel und Antoine 2007). In Abb. 3 wird dies illustriert.

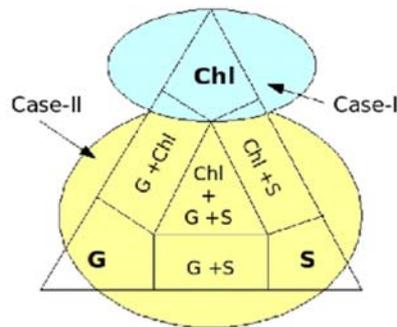


Abb. 3: Gewässereinteilung nach den drei Sichtbaren Inhaltsstoffen Chlorophyll-a (Chl), Gelbstoffe (G) und Schwebstoff-Gehalt (S) in Case I (hellblau) und Case II (gelb) (FRANK 2008: 16).

Während bei Case I-Gewässern vor allem durch Phytoplankton die optischen Eigenschaften bestimmt werden, so können bei Case II-Gewässern auch Gelb- und Schwebstoffe Einfluss auf diese Eigenschaften ausüben. Dadurch lassen sich Case II-Gewässer in sechs Untergruppen gliedern (FRAUENDORF 2002).

3.1 Falschfarbenbilder

Zur Bildinterpretation werden unterschiedliche Kanalkombinationen durchgeführt. In der Geologie wurde daher häufig beim Landsat 7 TM die RGB-Kombination (TM 7-4-2) genutzt (MWANIKI et al. 2015). Diese Kanalkombination entspricht beim Landsat 8 OLI den Kanälen (OLI 7-5-3). Dadurch lassen sich unterschiedliche Gesteins- und Bodentypen gut erkennen (MWANIKI et al. 2015).

3.2 Spektralindizes

Zur Ableitung von Eisenoxiden mittels eines Spektralindizes werden in der Literatur einige Methoden vorgestellt (siehe Tab. 2). Diese beziehen sich in den meisten Fällen auf die Detektion von Eisenoxiden im Boden mittels eines Fernerkundungssatelliten wie beispielsweise Landsat 7 TM oder ASTER (VAN DER MEER et al. 2014). Da für den Zeitraum der Untersuchungen nur Landsat 8 OLI-Daten zur Verfügung standen, wurden die Spektralindizes vom Landsat 7 TM- und ASTER-Satelliten an die Kanäle des Landsat 8 OLI-Satelliten angepasst. In Tab. 1 sind die Bänder der drei genannten Fernerkundungssatelliten entsprechend der Wellenlänge aufgelistet. Wie zuerkennen ist, überlappen sich einige Kanäle der Landsat-Satelliten mit denen vom ASTER-Satelliten. Während die Landsat-Baureihe eher auf den VIS-Bereich ausgelegt ist, besitzt ASTER einer starke Differenzierung im SWIR-Bereich (VAN DER MEER et al. 2014).

Aufgrund der Überlappung der Kanäle vom Landsat 7 TM und ASTER zum Landsat 8 OLI können die Spektralindizes entsprechend an den Landsat 8 angepasst werden. Dieser Sachverhalt ist in Tab. 2 dargestellt.

Tab. 1: Auflistung der benötigten Kanäle der Fernerkundungssatelliten Landsat 7 TM, Landsat 8 OLI und ASTER.

Beschreibung	Landsat 7 TM ³		Landsat 8 OLI ³		ASTER ⁴	
	Band	Zentrale Wellenlänge [nm]	Band	Zentrale Wellenlänge [nm]	Band	Zentrale Wellenlänge [nm]
Aerosole	-	-	1	440	-	-
Blau	1	485	2	480		
Grün	2	560	3	560	1	560
Rot	3	660	4	655	2	660
Nahes Infrarot	4	835	5	865	3N	820
Cirruswolken	-	-	9	1.370	-	-
SWIR I	5	1.650	6	1.610	4	1.650
SWIR II	7	2.220	7	2.200	5	2.165
Panchromatic	8	710	8	590	-	-

Die Anpassung der Kanäle an den Landsat 8 OLI erfolgt problemlos, da Landsat 7 und 8 nahezu gleich aufgebaut sind und der ASTER-Satellit zur Ableitung von Eisenoxiden nur die Kanäle eins bis fünf benötigt.

Tab. 2: Darstellung der Spektralindizes zur Bestimmung von Eisen(II)- und Eisen(III)oxide.

Spektralindizes	Landsat TM 7		Landsat 8 OLI
Eisen(II)- u. Eisen(III)oxide	$\frac{TM_3}{TM_1}$	(van der Meer et al. 2014; Mwaniki et al. 2015)	$\frac{OLI_4}{OLI_2}$
Eisen(III)	$\frac{TM_3}{TM_1}$	(Durning et al. 1998)	$\frac{OLI_4}{OLI_2}$
Eisen(II)	$\frac{TM_5}{TM_4}$	(Durning et al. 1998; Mwaniki et al. 2015)	$\frac{OLI_6}{OLI_5}$
Spektralindizes	ASTER		Landsat 8 OLI
Eisen(III)	$\frac{Aster_2}{Aster_1}$	(KALINOWSKI und OLIVER 2004)	$\frac{OLI_4}{OLI_3}$
Eisen(II)	$\frac{Aster_5}{Aster_3} + \frac{Aster_1}{Aster_2}$	(KALINOWSKI und OLIVER 2004)	$\frac{OLI_7}{OLI_5} + \frac{OLI_3}{OLI_4}$

4 Ergebnisse

4.1 Eisenerzmine nahe der Bergbaustadt „Mariane“

Anhand der Falschfarbenkombination (Landsat 8 OLI, Kanal 7-5-3) der Eisenerzmine sind die einzelnen Bodentypen gut zu differenzieren. Diese Methodik ist jedoch nicht geeignet, um die Eisen(II)- und Eisen(III)oxide im Gewässer abzuleiten. Denn diese werden in Abb. 4 (*rechts*) blau dargestellt, obwohl in Abb. 4 (*links*) eine deutliche Verockerung zuerkennen ist.

³ Landsat-Kanäle nach Satellit: http://landsat.usgs.gov/best_spectral_bands_to_use.php

⁴ ASTER-Kanäle: http://asterweb.jpl.nasa.gov/content/03_data/04_Documents/aster_user_guide_v2.pdf

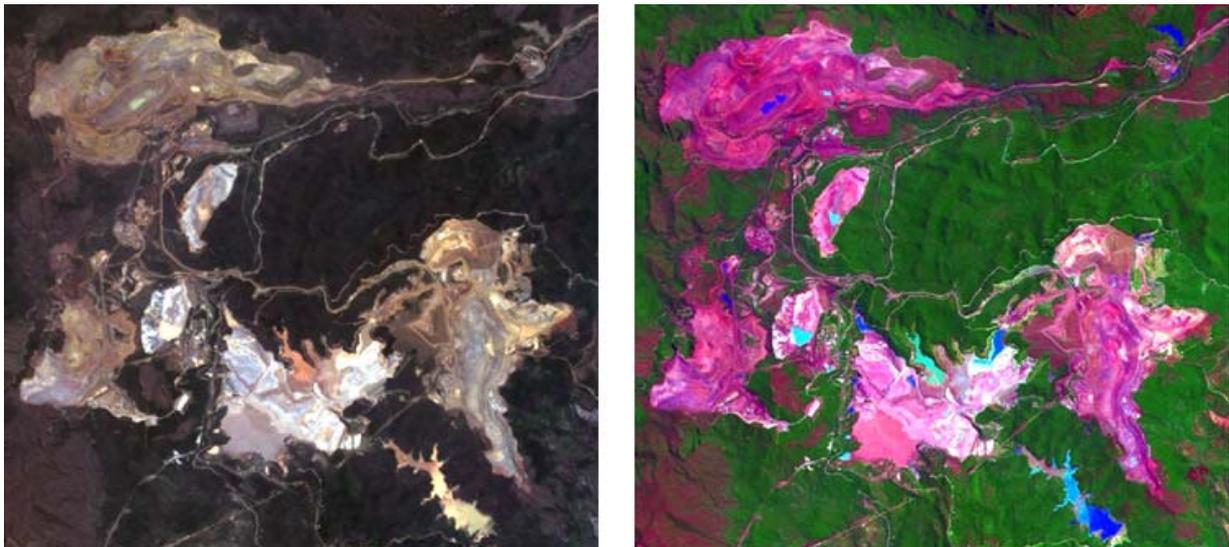


Abb. 4: Eisenerzmine vor dem Unglück am 5.11.2015. Links als RGB-Bild (OLI 4-3-2) und rechts als Falschfarbenbild (OLI 7-5-3). Im Falschfarbenbild ist Vegetation grün, Gewässer blau und Boden in rötlichen Variationen.

In Abb. 5 ist eine Indexdarstellung von Eisen(II)- und Eisen(III)oxiden im Bereich der Eisenerzmine vor den Damnbrüchen dargestellt. Das linke Bild der Abb. 5 stellt das Resultat mittels der Spektralindizes vom Landsat 7 TM und das rechte mittels der ASTER-Indizes dar. Die Skala beider Berechnungen ist gleich, sodass die Ergebnisse vergleichbar sind. In Abb. 5 sind nur die Farben innerhalb des angegebenen Wertebereiches dargestellt. Es ist zu erkennen, dass die Resultate der Spektralindizes stellenweise dieselben Orte aufzeigen, lediglich der Umfang ist unterschiedlich groß. Des Weiteren sind die Eisenoxide räumlich deutlich voneinander getrennt. Während Eisen(II)oxide nur an trockenen Stellen vorhanden sind, befinden sich die Eisen(III)oxide in den Gewässern.

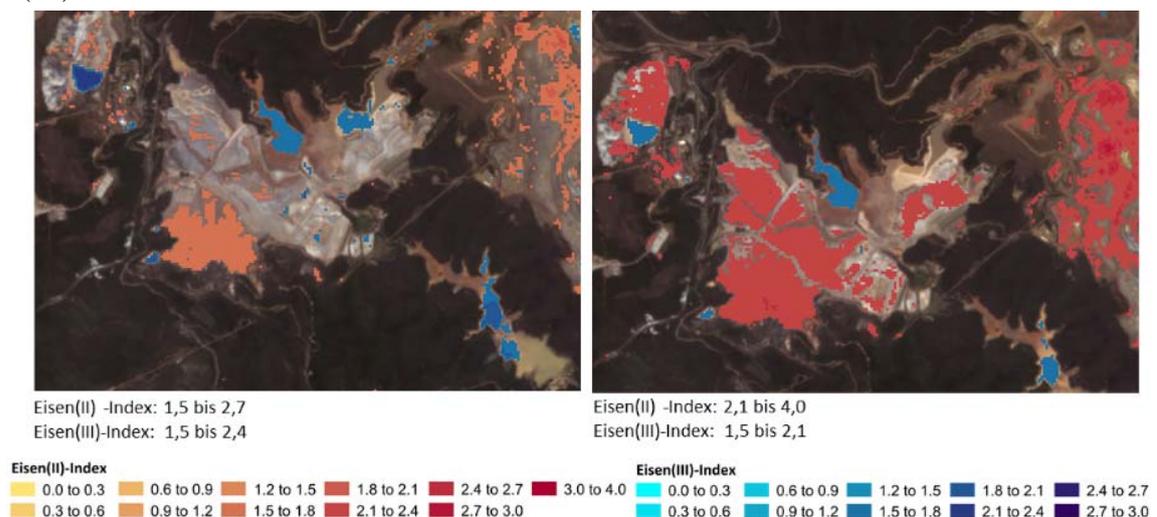


Abb. 5: Darstellung der Eisen(II)- und Eisen(III)oxid - Menge als Index vor dem Minenunglück. Übertragung der Spektralindizes von Landsat 7 TM (links) und ASTER (rechts) auf Landsat 8 OLI.

Die ASTER-Indizes liefern eine bessere Differenzierung von Eisen(II)oxiden. Dem gegenüber ist der Landsat 7 TM-Index zur Ableitung von Eisen(III)oxiden besser geeignet, da der ASTER-Satellit keinen blauen-Kanal besitzt.

Abb. 6 stellt die betroffene Eisenerzmine nach den Staudammbrüchen dar. Anhand der starken Durchmischung der Eisen(II)- und Eisen(III)oxide, sind die Folgen der Schlammlawine offensichtlich zu erkennen. Allerdings befindet sich das Eisen(II)oxid noch auf den Abraumflächen und das Eisen(III)oxid im Flussbett. Ferner ist anhand des Eisen(III)-Indizes des Landsat 7 TM ein Rückgang der Konzentration zu erkennen. Folglich wurde Eisen(III)oxid mit der Schlammlawine in den „Rio Doce“ eingeschwemmt.

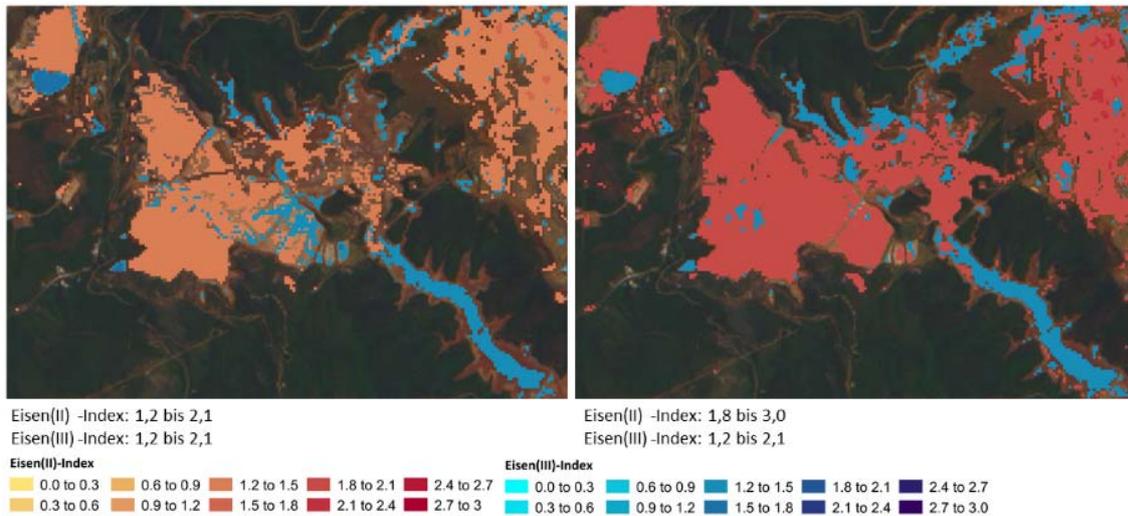


Abb. 6: Darstellung der Eisen(II)- und Eisen(III)oxid - Menge als Index nach dem Minenunglück. Übertragung der Spektralindizes von Landsat 7 TM (links) und ASTER (rechts) auf Landsat 8 OLI.

4.2 Stausee westlich von Colatina

Übersichtshalber ist in Abb. 7 der Stausee vor und nach dem Minenunglück zum einen als RGB-Bild dargestellt und zum anderen als Ratio-Bild. Dazu wurden drei Spektralindizes des Landsat 7 TM $\left(\frac{TM_3}{TM_1}, \frac{TM_5}{TM_7}, \frac{TM_5}{TM_4}\right)$ auf den Landsat 8 OLI übertragen. Mit Hilfe des Index $\frac{TM_5}{TM_7}$ können Karbonate, Mineralien mit einer Hydroxylgruppe sowie Vegetationen abgeleitet werden. Als RGB-Kombinationen stellen die Indizes $\left(\frac{TM_3}{TM_1} - \frac{TM_5}{TM_7} - \frac{TM_5}{TM_4}\right)$ Eisen(III) rot, Eisen(II) blau, Gewässer schwarz und Vegetation grün dar (DURNING et al. 1998). Hier ist deutlich zu erkennen, dass der „Rio Doce“ im Bereich des Stausees viel Phytoplankton besaß, was in diesem Fall ein Indikator für einen intakten Lebensraum ist. Die gravierenden Folgen der Dammbürche der Eisenerzmine sind anhand der Verockerung und somit an der hohen Eisen(III)oxid Konzentration ersichtlich.



Abb. 7: Darstellung des Stausees vor und nach dem Minenunglück als RGB- und Ratio-Bild (Landsat TM 7 3/1-5/7-5/4).

Wie stark die Eisen(III)oxid Konzentration zugenommen hat, ist in Abb. 8 zuerkennen. Während der Eisen(III)-Index vor der Schlammlawine nur geringe Mengen an Eisen(III)oxide aufzeigt, so hat im Nachhinein die Konzentration deutlich zugenommen.

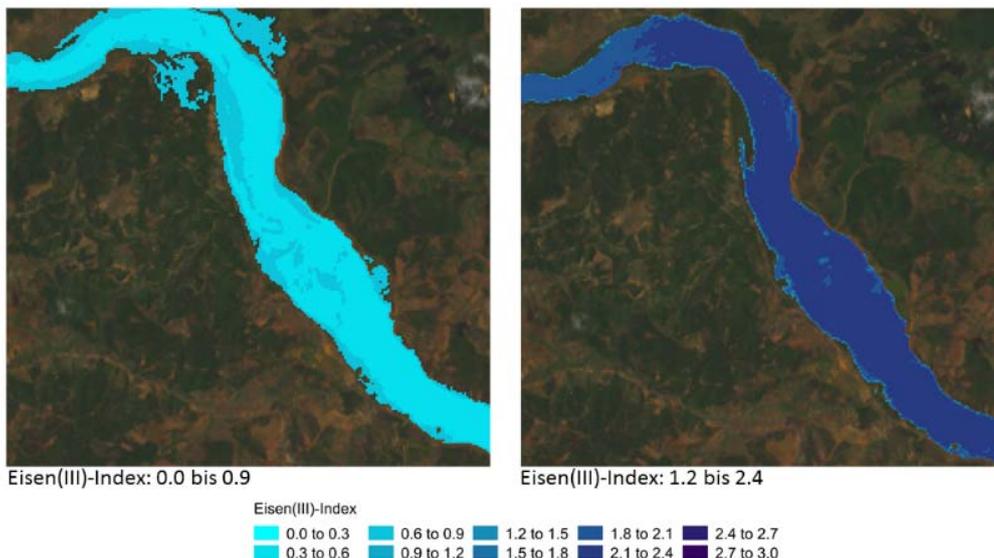


Abb. 8: Darstellung des Eisen(III)-Indexes am Stausee westlich von Colatina vor den Dambrüchen an der Eisenerzmine (links) und nach den Dambrüchen (rechts).

5 Fazit & Ausblick

Es hat sich gezeigt, dass die im VIS Bereich liegenden spektralen Kanäle von Landsat 8 OLI genutzt werden können, um Informationen über die Eisen(II)- und Eisen(III)oxid Konzentration abzuleiten. Die Spektraloperationen machen das Ausmaß der Katastrophe der Dammbrechung in Brasilien deutlich sichtbar. Die Spektralindizes des Landsat 7 TM- und ASTER-Satelliten lieferten am „Rio Doce“ unterschiedliche Ergebnisse. Der zur Ableitung von Eisen(II)oxide eingesetzte

ASTER-Index lieferte eine bessere Differenzierung als der Index des Landsat 7 TM. Für die Ableitung von Eisen(III)oxide verhielt es sich genau andersherum, da führte der Spektralindex des Landsat 7 TM zu einem differenzierteren Ergebnis.

Das Falschfarbenbild (OLI 7-5-3) lieferte bei der Ableitung von Eisen(II, III)oxide keine positiven Ergebnisse.

Probleme der Verockerung sind weltweit zu finden, wenn auch meist im geringeren Ausmaß. So findet in der Lausitz, infolge des offenen Tagebaus, eine stetige Verockerung der Spree und der künstlich geschaffenen Gewässer statt. Zur quantitativen Bestimmung der verschiedenen Eisenkonzentrationen werden derzeit Algorithmen basierend auf hyperspektralen Fernerkundungsdaten entwickelt.

6 Literaturverzeichnis

- AKTIONSBÜNDNIS KLARE SPREE E.V., 2013: Aktionsbündnis „Klare Spree“. Blog. Allgemeine Informationen. <http://klare-spree.de/de/informationen-zur-verockerung/allgemeine-informationen.html> (23.09.2015).
- BRÜHWILLER, T., 2015: Folgen eines Dammbrochs: Brasiliens toter Fluss - NZZ Unglücksfälle und Verbrechen. Nachrichten. Neue Zürcher Zeitung. <http://www.nzz.ch/panorama/ungluecksfaelle-und-verbrechen/brasiliens-toter-fluss-1.18648108> (09.02.2016).
- CHOE, E., VAN DER MEER, F., VAN RUITENBEEK, F., VAN DER WERFF, H., DE SMETH B. & KIM, K.-W., 2008: Mapping of heavy metal pollution in stream sediments using combined geochemistry, field spectroscopy, and hyperspectral remote sensing: A case study of the Rodalquilar mining area, SE Spain. *Remote Sensing of Environment*, **112** (7), 3222-3233.
- DURNING, W.P., POLIS, S.R., FROST, E.G. & KAISER, J.W., 1998: Integrated Use of Remote Sensing and GIS for Mineral Exploration. http://www.gis.usu.edu/docs/data/nasa_arc/nasa_arc97/SDSU/LaCuesta.pdf.
- EISENHAMMER, S., 2015: Mud from Brazil dam burst is toxic, U.N. says. Reuters (26.11.2015). <http://www.reuters.com/article/us-brazil-damburst-mud-idUSKBN0TE2O520151126> (30.03.2016).
- ENDER, R., FLEISCHHAMMEL, P., GRÜNEWALD, U., SCHAPP, A., SCHOENHEINZ, D., SCHÜMBERG, S., SEILER, D., UHLMANN, W. & ZIMMERMANN, W., 2012: „Perspektive See“ - Zum Stand der Entwicklung der Wasserbeschaffenheit in den Lausitzer Bergbaufolgeseen. Abschlussbericht Projektzeitraum 2008-2012. Cottbus, Dresden. http://www.lmbv.de/tl_files/LMBV/Dokumente/Wassermanagement/Forschung%20zu%200Seen/Perspektive_See_Abschlussbericht2008-2012.pdf.
- FLÖRKE, S., 2015: 660 Kilometer Fluss, 60 Millionen Kubikmeter Gift. Welt Online (26.11.2015). <http://www.welt.de/vermishtes/article149323787/660-Kilometer-Fluss-60-Millionen-Kubikmeter-Gift.html> (30.03.2016).
- FRANK, T., 2008: Chlorophyll-a-Kartierung des Kivusees (Ruanda, Kongo) mit Fernerkundungsdaten. Diplomarbeit, Zürich.
- FRAUENDORF, J., 2002: Entwicklung und Anwendung von Fernerkundungsmethoden zur Ableitung von Wasserqualitätsparametern verschiedener Restseen des Braunkohletagebaus in

- Mitteldeutschland. Dissertation. Halle (Saale). <http://sundoc.bibliothek.uni-halle.de/diss-online/02/02H226/prom.pdf> (11.01.2015).
- GEGE, P., 2004: Fernerkundung von Wasserinhaltsstoffen. Laufener Seminarbeiträge 2/2003. Laufen/Salzach 2004, 73-78. http://www.anl.bayern.de/publikationen/spezialbeitraege/doc/lsb2003_02_011_gege_fernerkundung_wasserinhaltsstoffe.pdf (10.12.2015).
- KALINOWSKI, A. & OLIVER, S., 2004: ASTER Mineral Index Processing Manual. Remote Sensing Applications, Geoscience Australia, 36 p.
- KELLER, I., 1999: Invariante Textur-Klassifikation in der Fernerkundung. Dissertation zur Erlangung des akademischen Grades des Doktors der Naturwissenschaften, Berlin. http://userpage.fu-berlin.de/geoiss/ress/15_Keller_Ivo.pdf (10.12.2015).
- LENNTECH, 2015: Eisen und Wasser. Eisen (Fe) und Wasser. <http://www.lenntech.de/pse/wasser/eisen/eisen-und-wasser.htm> (22.01.2016).
- VAN DER MEER, F.D., VAN DER WERFF, H.M.A. & VAN RUITENBEEK, F.J.A., 2014: Potential of ESA's Sentinel-2 for geological applications. *Remote Sensing of Environment* **148**, 124-133.
- MELENDENZ-PASTOR, I., NAVARRO-PEDREÑO, J., GÓMEZ, I. & ALMENDRO-CANDEL, M.B., 2011: The use of remote sensing to locate heavy metal as source of pollution. *Advances in Environmental Research*, **7**, 225-233.
- MINISTERIUM FÜR WIRTSCHAFT UND ENERGIE DES LANDES BRANDENBURG, 2015: Verockerung der Spree. Energieland Brandenburg. <http://www.energie.brandenburg.de/sixcms/detail.php/bb1.c.323611.de> (04.01.2015).
- MOREL, A. & ANTOINE, D., 2007: ATBD 2.9 - Pigment Index Retrieval in Case 1 Waters. *Laboratoire d'Océanographie de Villefranche* (4), 25.
- MWANIKI, M.W., MOELLER, M.S. & SCHELLMANN, G., 2015: A comparison of Landsat 8 (OLI) and Landsat 7 (ETM+) in mapping geology and visualising lineaments: A case study of central region Kenya. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (7/W3), 897-903.
- SÁNCHEZ, J., 2008: Acid Mine Drainage in the Iberian Pyrite Belt: an Overview with Special Emphasis on Generation Mechanisms, Aqueous Composition and Associated Mineral Phases. Band **10**, Sevilla, 34-43. http://www.ehu.es/sem/macla_pdf/macla10/Macla10_34.pdf (22.01.2016).
- UHLMANN, W., THEISS, S., NESTLER, W. & LAUS, T., 2013: Studie zu den Auswirkungen des Grundwasserwiederanstieges auf die Beschaffenheit der Oberflächengewässer in den Sanierungsgebieten B1 (Seese/Schlabendorf) und B2 (Greifenhain/ Gräbendorf). Im Auftrag der LMBV. Dresden. <http://www.braunespreewatch.de/images/studien/13-03-29Studie-Eisenbelastung-Nordraum-Spree-2013.pdf> (06.01.2015).
- UPADEK, C., 2016: Umweltkatastrophe am Rio Doce - Das Leben nach der Schlammflut. Nachrichten. Deutschlandfunk. http://www.deutschlandfunk.de/umweltkatastrophe-am-rio-doce-das-leben-nach-der-schlammflut.799.de.html?dram:article_id=343403 (30.03.2016).

Towards Integrated 3D Reconstruction and Semantic Interpretation of Urban Scenes

MAROŠ BLAHÁ¹, CHRISTOPH VOGEL², AUDREY RICHARD¹, JAN D. WEGNER¹,
THOMAS POCK² & KONRAD SCHINDLER¹

Abstract: We describe recent progress made towards automated extraction of semantically annotated 3D city models from aerial imagery. The idea of semantic 3D reconstruction is to reconstruct the 3D geometry of an observed scene while at the same time also interpreting the scene in terms of semantic object classes (such as buildings, vegetation etc.) - similar to a human operator, who also interprets the image content while making measurements. The advantage of jointly reasoning about shape and object class is that one can exploit class-specific a-priori knowledge about the geometry: on the one hand the type of object provides information about its shape, e.g. walls are likely to be vertical, whereas streets are not; on the other hand, 3D geometry is also an important cue for classification, e.g. in our example vertical surfaces are more likely to be walls than streets. Recent work has developed computational models that allow one to jointly infer geometry and class. For simple priors, such as preferred surface orientations, these models even have favourable mathematical properties like convexity of the optimisation. However, they rely on a dense, volumetric discretisation of 3D space, therefore the computation is memory-hungry and slow. We have developed an adaptive, hierarchical formulation of semantic 3D reconstruction, which makes it possible to process scenes of much larger, realistic size. The intuition is that both high spatial resolution and high numerical precision are only required in regions that are likely to contain a surface. Technically, our scheme amounts to repeatedly solving a constrained, convex optimisation problem, while iteratively removing low-confidence constraints. In our experiments the adaptive reconstruction incurs no loss in accuracy, but offers up to 98% lower memory consumption and up to 95% shorter computation time.

1 Introduction

Photogrammetric mapping encompasses two visual tasks: to *geometrically reconstruct* the observed scene in 3D, and to *semantically interpret* the data. These two tasks are not independent of each other, and human operators naturally solve them in conjunction, utilising their a-priori knowledge about the shape of different types of objects like buildings, water surfaces etc. On the contrary, computer systems for automated photogrammetric analysis treat the two tasks in isolation and sequentially. Either geometry is reconstructed as a generic surface model, which can then serve as input to extract objects like buildings, the ground (DTM) etc. Or the images are first interpreted to detect objects, which can then be individually reconstructed.

In which order should one go about these two steps? Arguably one should address them jointly, because a-priori knowledge about the world acts in both ways. A building probably has vertical walls; but conversely, a concrete-grey structure is much more likely to be a building wall if it is

¹ ETH Zürich, Photogrammetry and Remote Sensing, Stefano-Francini-Platz 5, CH-8093 Zürich, E-Mail: [maros.blaha, audrey.richard, jan.wegner, schindler]@geod.baug.ethz.ch

² TU Graz, Computer Graphics and Vision, Inffeldgasse 16, A-8010 Graz, E-Mail: [christoph.vogel, pock]@icg.tugraz.at

vertical. However, treating both the 3D geometry and the semantic class labels as unknowns and inferring them together is technically challenging. Only recently models have emerged that capture the problem in a principled manner (HÄNE et al. 2013, SAVINOV et al. 2015), but they are computationally expensive and therefore limited to small regions of interest and/or few images.

Here, we describe a way to make semantic 3D reconstruction a lot more efficient, such that it scales up to larger regions and image sets. The application we are interested in is the generation of interpreted 3D city models from (nadir and oblique) aerial images. We follow the pioneering work of (HÄNE et al. 2013), where semantic reconstruction is formulated as a multi-class labelling problem on a voxel grid. Like in volumetric surface reconstruction (CURLISS & LEVOY 1996) the 3D space is discretised into voxels. But rather than only being labelled as *freespace* or *solid*, voxels can take on multiple labels like *freespace*, *building*, *ground*, *vegetation*, etc.

Our work builds on the elementary insight that instead of a regular voxel grid the labelling can be done with variable volumetric resolution. Large parts of the volume – in particular *freespace*, but also the inside of buildings, areas under the ground, etc. – only need to be modelled at coarse resolution. A fine discretisation and high numerical precision are only required near the boundary surfaces. We start from a coarse voxel grid and adaptively refine the reconstruction only near (predicted) label transitions. On the one hand this reduces memory consumption, so that at a given target resolution one can reconstruct larger scenes. On the other hand, it also speeds up the computation, because after every refinement an approximate solution is already available from the previous level, like in a multi-scale pyramid. In our experiments we observe up to 40× lower memory footprint and 22× shorter runtime, without any loss of quality.

2 Related Work

Automated 3D city reconstruction is a classical problem of photogrammetry. For purely geometric surface reconstruction, volumetric representations are rare and explicit surface representations are preferred. In this context (LAFARGE & MALLET 2012) have already proposed to first estimate semantic labels for 3D points and use those to support geometric reconstruction (but not vice versa). Early attempts to merge geometric and semantic reconstruction into a one-shot process started with depth maps (LADICKÝ et al. 2010), later work moved to volumetric representations (HÄNE et al. 2013; BAO et al. 2013; KUNDU et al. 2014; SAVINOV et al. 2015; VINEET et al. 2015) or, rarely, meshes (CABEZAS et al. 2015). All these works are variants of the same conceptual idea, to do the semantic labelling in 3D, such that it is inherently consistent across images, while at the same time enforcing class-specific priors rather than isotropic smoothing. (HÄNE et al. 2013) employ a non-metric regulariser, developed within the linear programming relaxation of a multi-label Markov random field (ZACH et al. 2014) in 3D voxel space. (KUNDU et al. 2014) is perhaps the closest work to ours. They also employ the octree data structure to save memory, but fix the local resolution at the beginning, based on an initial guess of the geometry (whereas we adaptively refine it). Also, like all other works mentioned above it uses only street-level imagery and models only scene parts visible from the streets, whereas we use aerial images and reconstruct the entire scene.

Before semantic (multi-label) 3D reconstruction, volumetric representations were already used for generic surface modelling (CURLLESS & LEVOY 1996), where a distance field or indicator function is defined on the voxels. From that representation one can extract an explicit surface by finding the zero level set (e.g., LORENSEN & CLINE 1987; KAZHDAN et al. 2007). Many volumetric techniques work with regular voxels (CREMERS & KOLEV 2011; KOLEV et al. 2012, KOSTRIKOV et al. 2014). The data term, corresponding to a voxel’s probability of lying behind the surface, is typically a signed distance generated from image correspondences (e.g., ZACH et al. 2007, ZACH 2008). Some authors even model a pixel’s visibility along the complete ray (LIU & COOPER 2010, SAVINOV et al. 2015), which however leads to higher-order potentials over all voxels intersected by each ray, such that these methods do not scale to larger datasets. Instead of the regular voxels (LABATUT et al. 2007; JANCOSSEK & PAJDLA 2011) tessellate the space by Delaunay tetrahedralisation of the initial multi-view point cloud. Tetrahedrons are labelled as empty or occupied, and triangles on the interface between the two labels form the object surface.

Algorithms for surface reconstruction from point clouds, like the Poisson method (KAZHDAN et al. 2006), also use octrees; and in some cases also multi-grid solvers (BOLITHO et al. 2007). This is similar in spirit to our adaptive multi-scale approach, but the least-squares nature of Poisson reconstruction is susceptible to outliers. Our model allows for robust error functions, at the cost of a more complicated optimisation. Also, our octree structure is not determined once and for all by the input data, but refined adaptively. A different view on our model is to see it as a coarse-to-fine reconstruction on a volumetric pyramid (ZACH et al. 2007; ZACH 2008), in which the refinement is applied selectively (e.g., HORNUNG & KOBBELT 2006).

3 Method

To simplify the technical description, we first describe the basic model with a regular voxel grid, and then extend it to the irregular, adaptive scheme. Throughout, the description stays on a conceptual level, for mathematical details the interested reader is referred to (BLAHÁ et al. 2016).

3.1 Basic Model

Let the region of interest Ω be discretised into regular, equally sized voxels $s \in \Omega$. In this discrete representation, joint geometric and semantic reconstruction can be cast as a labelling problem: at each voxel, determine the most likely semantic class label. By including a class *freespace*, one at the same time estimates also which voxels do not belong to any object, and thus implicitly reconstructs the 3D surfaces that separate objects from freespace. At each voxel an indicator vector $x^i \in [0,1]$ is stored, which is 1 for the assigned class and 0 for all other classes. The trick of (HÄNE et al. 2013) is to additionally store pseudo-marginals x^{ij} for each pair of classes and each grid direction, such that class transitions and their orientations are made explicit. Finding the best labelling then amounts to minimising the energy function

$$E(x) = \sum_{s \in \Omega} \left(\sum_{classes\ i} \rho_s^i x^i + \sum_{class_pairs\ i,j} \varphi^{ij} (x_s^{ij} - x_s^{ji}) \right),$$

subject to appropriate constraints which ensure that the variables are non-negative, that they sum to 1 appropriately, and that the x^i and x^{ij} are consistent. The ρ^i are conventional data terms that encode the likelihood of different labels at voxel s . The convex and 1-homogeneous functions φ^{ij} encode the individual a-priori likelihoods of different class transitions, taking into account the orientation $(x^{ij} - x^{ji}) \in [-1,1]^3$ of the boundary surface. The energy can be seen as a generalised form of the standard Markov Random Field energy (in its linear programming relaxation).

The **data term** ρ^i combines depth values and semantic class probabilities observed in the images. It favours configurations where the transition from freespace to another class occurs at the predicted depth along a pixel’s viewing ray, and penalises deviations from that depth with a truncated linear penalty. At the same time, ρ^l maps the class likelihoods from the images into the 3D volume, by applying them at a voxel along the ray, slightly behind the predicted surface. The data cost in a voxel is computed by summing over all rays that pass through it. We point out that truncating the data cost means that we require the volume to be empty only in a limited interval in front of the predicted depth, rather than along the entire ray. This is clearly an approximation, but has the advantage that the data evidence can be encoded as a single unary term per voxel. In contrast, taking into account the full length of the viewing rays would lead to a higher-order potential that links all voxels along the ray, for each single pixel (LIU & COOPER 2010; KUNDU et al. 2014; SAVINOV et al. 2015). An efficient treatment of such potentials remains future work.

The class-specific **priors** φ^{ij} penalise class transitions, but other than Potts-type smoothing they account for the type of transition and the corresponding surface orientation. The prior is modelled as a sum of two terms, an isotropic part that encodes the likelihood (frequency) of the transition from class i to class j ; and an anisotropic part that increases the penalty if the transition occurs in an improbable direction (e.g., an overhanging building wall). A convenient way to encode such a non-metric (w.r.t. the label space) and direction-dependent regulariser is through the indicator function of an appropriately chosen convex set, the so-called *Wulff* shape (ZACH et al. 2014).

3.2 Adaptive Multi-scale Extension

The basic model described so far links semantic interpretation and shape reconstruction in a principled manner, but it needs a lot of memory and processing power. Instead of exhaustively storing variables for the finest grid resolution, we reduce the set of unknowns by storing voxels (and their indicator variables) at an adaptive resolution: finer close to the surface boundaries, coarse further away from them. Any voxel stores only one set of pseudo-marginals x^{ij} , regardless of its size, so that computational resources are saved for non-expanded voxels. For the new, adaptive discretisation Ω^l of the space, we seek an energy E_l that approximates the original energy E (in which all voxels have the finest resolution) as tight as possible.

It turns out that one can see the energy E_l over the reduced set of unknowns (induced by the adaptive discretisation) as a constrained version of the original energy E , subject to additional equality constraints. To ensure that E_l and the constrained E are identical, the regularisation φ^{ij} must change as a function of voxel size, respectively refinement level. Intuitively speaking, one has to compensate for the fact that, as the refinement proceeds, transition penalties are increasingly “concentrated” in a smaller fraction of the overall volume and on fewer surfaces of the boundary voxels. The transition cost Φ_l^{ij} at a certain refinement level is a weighted sum of the “virtual costs” that would apply at the highest voxel resolution. The overall energy now reads

$$E_l(x) = \sum_{s \in \Omega^l} \left(\sum_{\text{classes } i} \rho_s^i x^i + \sum_{\text{class_pairs } i,j} \Phi_l^{ij} (x_s^{ij} - x_s^{ji}) \right).$$

It sums the data costs over the voxels $s \in \Omega^l$ of the current refinement stage, and applies the corresponding level-dependent transition penalties Φ_l^{ij} . Non-negativity and normalisation constraints remain the same, consistency constraints must take into account that a voxel now may, at any of its faces, meet a single voxel of the same size, multiple smaller voxels, or part of a bigger voxel. For technical details and equations please see (BLAHA et al. 2016).

In our algorithm the scene is initially reconstructed on a very coarse grid and adaptively subdivided only close to the (putative) class boundaries. The refined variable set for the new, smaller voxels is initialised from the intermediate solution at the previous level. The alternation between energy minimisation and refinement is repeated, until the smallest voxels have reached the final target resolution. Note, due to the implicit representation the surface topology can change, e.g., a narrow street might open between two previously connected buildings.

The last missing piece is a criterion to decide which voxels to refine for the next round of energy minimisation. We simply refine all voxels that, at the current resolution, would be assigned to a different class than any of their neighbours. Moreover, we limit the resolution difference between adjacent voxels to at most 1, so splitting a voxel may trigger additional splits. In every iteration of the adaptive refinement, the energy function is, by construction, convex and can be minimised with standard tools. We convert it to primal-dual form with the help of Lagrange multipliers for the constraints and use the numerical scheme of (CHAMBOLLE & POCK 2011).

4 Experiments

As test dataset we use a block of 510 oblique and nadir images (102 exposures, Maltese cross configuration) from the city of Enschede, Netherlands. After orientation, per-image evidence for the depth is generated with semi-global matching (HIRSCHMÜLLER 2008), and class likelihoods are estimated with a MultiBoost classifier (BENBOUZID et al. 2012), using both RGB intensities and local shape features (CHEHATA et al. 2009). Figure 1 illustrates these pre-processing steps.

We impose two types of class-specific priors (Wulff shapes). One prefers horizontal surfaces with upward-pointing normal, and gradually increases the cost as the normal vector is tilted. This prior describes the boundaries *ground-freespace*, *ground-building*, *ground-vegetation*, *building-roof*, *roof-freespace*. The second prior has a strong preference for vertical transitions *building-freespace* and *building-vegetation*. The strength of the priors (tolerance to deviations from the preferred orientation) is set individually per label pair.

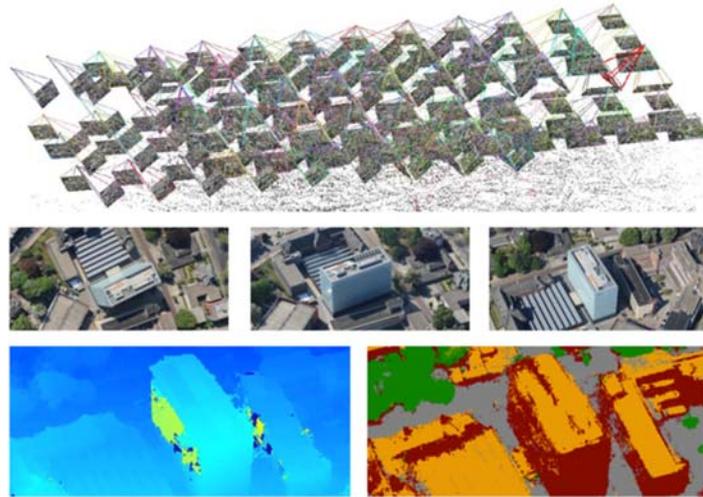


Fig. 1: Input data: Image orientations (*top*), depth maps (*bottom left*) and class likelihoods (*bottom right, illustration shows the most likely labels per pixel*).

In a first experiment, we compare the adaptive model to a fixed voxel grid with the same target resolution. Unfortunately, no 3D ground truth was available. As a pragmatic compromise, we evaluate semantic correctness in image space: Semantic segmentations are hand-labelled for two representative images. Both reconstructions are back-projected into those images and quantitatively compared on a per-pixel basis. Overall the differences between adaptive and non-adaptive reconstruction are tiny (<0.7 percent points) and mostly due to aliasing, see Figure 2. We conclude that the hierarchical scheme does not incur any loss of accuracy. In this context we point out a remaining systematic error of the current model (with or without adaptive refinement). The approximate data term demands that behind a surface a few voxels are occupied along each ray, which leads to fattening at thin surfaces and silhouette edges. The effect is best visible at the transition from *building* to *roof*. It could be mitigated by correctly modelling visibility along entire rays (SAVINOV et al. 2015), with much higher memory consumption.

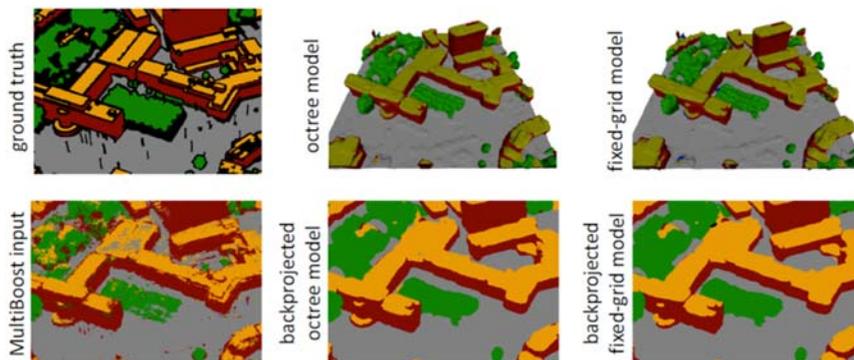


Fig. 2: Comparison of classification accuracy in 2D.

The benefits of joint reconstruction are bigger in the other direction, when semantic knowledge improves the surface geometry in regions where matching fails, see examples in Figure 3. Unfortunately, we do not have reference data to quantify this improvement.



Fig. 3: Reconstruction with and without class-specific priors. Left to right: example image; semantic reconstruction; semantic reconstruction back-projected to the image; and generic volumetric surface reconstruction with the same input (Zach 2008).

We go on to measure how much memory and computation time can be saved with the adaptive scheme. All experiments are run on a machine with 64 GB of RAM and a hexa-core *Intel i7* CPU. Due to the huge memory consumption of the baseline (non-adaptive) method we have to limit this comparison to small sub-regions of the data. To ensure the comparison is fair we only voxelise a tight bounding box around the data. The tight box is a lot smaller for Enschede, because the terrain is flat. In mountainous cities the savings would be even higher. The top level of the octree has voxel size 13.5 m^3 , the final resolution after five refinements is 0.4 m^3 . In all tests the adaptive scheme saves ca. 95%, see Table 1. For the two smaller scenes 3 and 4 we have run another refinement down to 0.2 m^3 voxel size. At that resolution we can no longer run the baseline, since it would require $> 108 \text{ GB}$ of memory, ca. $35\times$ more than the adaptive scheme. Figure 4 illustrates the evolution of the adaptive refinement. At the top, one can see how the accuracy and detail of the reconstruction gradually improves. The bottom row displays the voxel size in a vertical slice, colour-coded from blue (13.5 m^3) to yellow (0.2 m^3).

Table 1: Adaptive vs. non-adaptive volumetric reconstruction on two different test scenes.

	<i>runtime # 0.4m [sec]</i>		<i>memory @ 0.4m [GB]</i>		<i>memory @ 0.2m [GB]</i>	
	<i>scene 1</i>	<i>scene 2</i>	<i>scene 1</i>	<i>scene 2</i>	<i>scene 1</i>	<i>scene 2</i>
<i>voxel grid</i>	91'982	92'893	13.6	13.6	108.5	108.5
<i>octree</i>	5'488	4'984	0.7	0.7	3.3	2.7
<i>ratio</i>	16.8	18.6	19.4	19.4	32.9	40.2

Finally, we come to the target of reconstructing a large urban area and process all the 510 images at once to cover the city centre of Enschede (ca. 3 km^2) with target voxel size 0.8 m^3 ($1/2048$ of the bounding volume), see Figures 5 and 6. For this reconstruction the adaptive scheme uses a modest 28 GB of memory and runs 40 hours on one PC. To process the same dataset without adaptive computation ($2048\times 2048\times 128$ voxels), one would need 434 GB of memory.

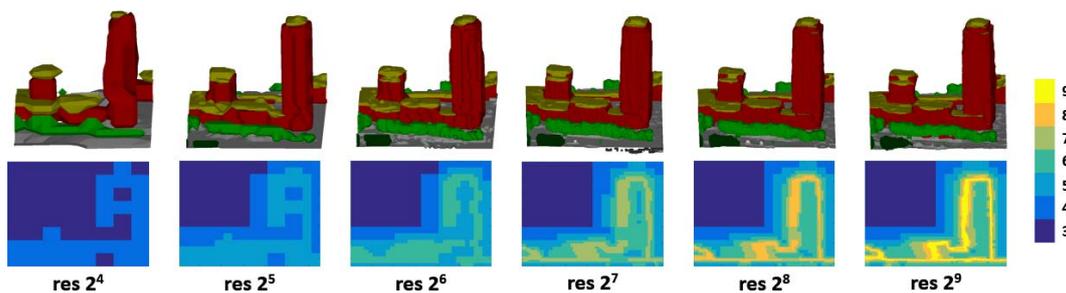


Fig. 4: Evolution of adaptive reconstruction over five refinement steps.

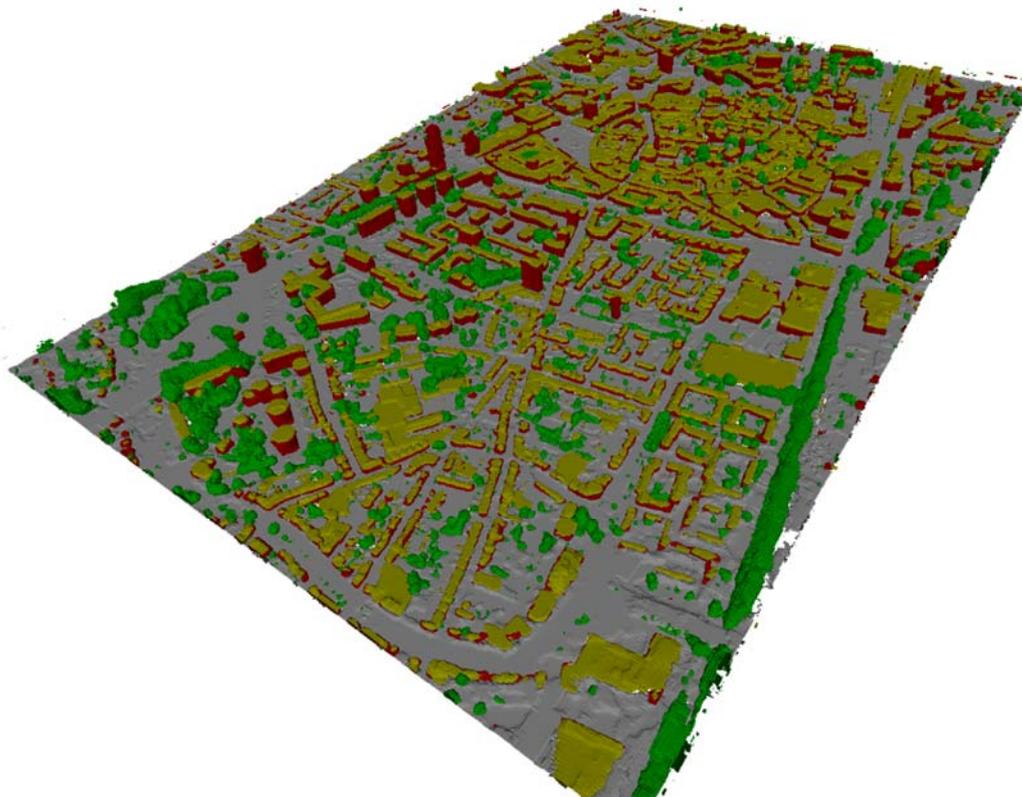


Fig. 5: Semantic reconstruction of Enschede city centre.

5 Conclusion

We have described an adaptive multi-resolution framework to jointly infer the 3D geometry and a semantic segmentation of a scene from multi-view imagery, taking into account interactions between surface geometry and object type. The framework greatly improves the computational efficiency of integrated geometric/semantic 3D reconstruction, such that one can process scenes of realistic size and resolution. In future work we will investigate how to transfer the idea of adaptive spatial refinement to irregular space tessellations like the recently popular Delaunay tetrahedralisation, which by construction already adapt to the point distribution of the dataset.

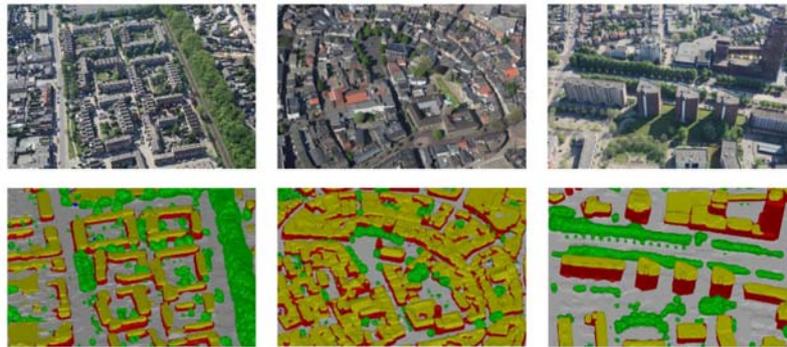


Fig. 6: Visual comparison to aerial images with the same viewpoint.

6 Acknowledgements.

We thank Christian Häne and Marc Pollefeys for source code and discussions. This work was supported by SNF grant 200021_157101.

7 Bibliography

- BAO, Y., CHANDRAKER, M., LIN, Y. & SAVARESE, S., 2013: Dense object reconstruction using semantic priors. *IEEE Conference on Computer Vision & Pattern Recognition*, 1264-1271.
- BENBOUZID, D., BUSA-FEKETE, R., CASAGRANDE, N., COLLIN, F.-D. & KÉGL, B., 2012: MULTIBOOST: a multi-purpose boosting package. *Journal of Machine Learning Research* **13**(1), 549-553.
- BLAHÁ, M., VOGEL, C., RICHARD, A., WEGNER, J., POCK, T. & SCHINDLER, K., 2016: Large-scale semantic 3d reconstruction: an adaptive multi-resolution model for multi-class volumetric labeling. *Computer Vision and Pattern Recognition (CVPR)*.
- BOLITHO, M., KAZHDAN, M., BURNS, R. & HOPPE, H., 2007: Multilevel streaming for out-of-core surface reconstruction. *5th Eurographics Symposium on Geometry Processing*, 69-78.
- CABEZAS, R., STRAUB, J. & FISHER III, J. W., 2015: Semantically-aware aerial reconstruction from multi-modal data. *International Conference on Computer Vision (ICCV)*, 2156-2164.
- CHAMBOLLE, A. & POCK, T., 2011: A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision* **40**(1), 120-145.
- CHEHATA, N., GUO, L. & MALLET, C., 2009: Airborne LiDAR feature selection for urban classification using random forests. *ISPRS Archives* **38**(3/W8).
- CREMERS, D. & KOLEV, K., 2011: Multiview stereo and silhouette consistency via convex functionals over convex domains. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(6), 1161-1174.
- CURLISS, B. & LEVOY, M., 1996: A volumetric method for building complex models from range images. *ACM SIGGRAPH*, 303-312. □
- HÄNE, C., ZACH, C., COHEN, A., ANGST, R. & POLLEFEYS, M., 2013: Joint 3d scene reconstruction and class segmentation. *Computer Vision and Pattern Recognition (CVPR)*.
- HIRSCHMÜLLER, H., 2008. Stereo processing by semiglobal matching and mutual information.

- IEEE Transactions on Pattern Analysis and Machine Intelligence **30**(2), 328-341. □
- HORNUNG, A. & KOBELT, L., 2006: Hierarchical volumetric multiview stereo reconstruction of manifold surfaces based on dual graph embedding. *Computer Vision and Pattern Recognition (CVPR)* **1**, 503-510.
- JANCOSEK, M. & PAJDLA, T., 2011: Multi-view reconstruction pre-serving weakly-supported surfaces. *Computer Vision and Pattern Recognition (CVPR)*, 3121-3128. □
- KAZHDAN, M., BOLITHO, M. & HOPPE, H., 2006: Poisson surface reconstruction. *Eurographics*.
- KAZHDAN, M., KLEIN, A., DALAL, K. & HOPPE, H., 2007: Unconstrained isosurface extraction on arbitrary octrees. *5th Eurographics Symposium on Geometry Processing*, 125-133.
- KOLEV, K., BROX, T. & CREMERS, D., 2012: Fast joint estimation of silhouettes and dense 3D geometry from multiple images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**(3), 493-505.
- KOSTRIKOV, I., HORBERT, E. & LEIBE, B., 2014: Probabilistic labeling cost for high-accuracy multi-view reconstruction. *Computer Vision and Pattern Recognition (CVPR)*, 1534-1541.
- KUNDU, A., LI, Y., DELLAERT, F., LI, F. & REHG, J., 2014: Joint semantic segmentation and 3d reconstruction from monocular video. *European Conference on Computer Vision (ECCV)*, Springer International Publishing, 703-718.
- LABATUT, P., PONS, J.-P. & KERIVEN, R., 2007: Efficient Multi-View Reconstruction of Large-Scale Scenes using Interest Points, Delaunay Triangulation and Graph Cuts. *IEEE 11th International Conference Computer Vision (ICCV)*, 1-8.
- LADICKÝ, L., STURGESS, P., RUSSELL, C., SENGUPTA, S., BASTANLAR, Y., CLOCKSIN, W. & TORR, P., 2010: Joint optimisation for object class segmentation and dense stereo reconstruction. *British Machine Vision Conference (BMVC)*.
- LAFARGE, F. & MALLET, C., 2012: Creating large-scale city models from 3D-point clouds: a robust approach with hybrid representation. *International Journal of Computer Vision* **99**(1), 69-85.
- LIU, S. & COOPER, D. B., 2010: Ray Markov random fields for image-based 3d modeling: Model and efficient inference. *Computer Vision and Pattern Recognition (CVPR)*, 1530-1537.
- LORENSEN, W.E. & CLINE, H.E., 1987: Marching cubes: A high resolution 3d surface construction algorithm. *ACM SIGGRAPH Computer Graphics* **21** (4), 163-169.
- SAVINOV, N., LADICKÝ, L., HÄNE, C. & POLLEFEYS, M., 2015: Discrete optimization of ray potentials for semantic 3d reconstruction. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5511-5518.
- VINEET, V., MIKSIK, O., LIDEGAARD, M., NIEBNER, M., GOLODETZ, S., PRISACARIU, V.A., KAEHLER, O., MURRAY, D.W., IZADI, S., PEREZ, P. & TORR, P.H., 2015: Incremental dense semantic stereo fusion for large-scale semantic scene reconstruction. *2015 IEEE International Conference on Robotics and Automation*, 75-82.
- ZACH, C., 2008: Fast and high quality fusion of depth maps. *Proceedings of the International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)* **1**, 2.
- ZACH, C., HÄNE, C. & POLLEFEYS, M., 2014: What is optimized in convex relaxations for multilabel problems: Connecting discrete and continuously inspired MAP inference. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **36**(1), 157-170.
- ZACH, C., POCK, T. & BISCHOF, H., 2007: A globally optimal algorithm for robust TV-L1 range image integration. *IEEE 11th International Conference on Computer Vision (ICCV)*, 1-8.

Semantic 3D City Models Serving as Information Hub for 3D Field Based Simulations

BRUNO WILLENBORG¹, MAXIMILIAN SINDRAM¹ & THOMAS H. KOLBE¹

Abstract: A major challenge in the context of global urbanization is to make cities “smarter” in order to satisfy their increasing requirements to safety, environmental conditions and quality of living. This requires both the environmental and social processes taking place in cities to be perspicuous and predictable. Today, simulations from various disciplines like wind field or flood simulations, noise propagation, pollutant dispersion or explosion simulations have become essential tools for decision making in urban planning and analytics. A common simulation approach is the “Finite Volume”-method, where the observation area is partitioned into a regular grid of cells to solve partial differential equations numerically. 3D city models serve as reliable representation of the real world objects.

In this study we present an approach for the integration of such field-based simulations with object-based, semantic 3D city models. It allows for automatically deducing a finite volume-based representation of arbitrary objects in 3D city models. The city model is used as an information hub between the systems. It serves both as the data source for the geometry representation of the simulation and its result storage. For this, the corresponding region is divided into equal cubes (voxels) and each voxel is tested against intersection with the city objects. Thus, each voxel inherits the semantics of its intersecting city object, which ensures that a) the voxel holds the characteristics of its counterpart and b) the results of the simulation runs can be aggregated per object and stored back in the city model.

Visualization and result analysis is of crucial importance, since highly tailored simulation tools do not offer interactive access to the simulation results. In contrast, the frameworks for 3d city models provide comprehensive tools for both visualization and intuitive analysis. They allow even for non-expert users to assess the information generated by the simulation tools.

The proposed approach is introduced and evaluated on the example of a blast simulation and implemented using a 3D city model according to the international standard CityGML. It is realized in an Open Source plug-in for the 3D geodatabase software 3DCityDB.

1 Introduction and problem statement

Urbanization is one of the biggest challenges of our time. While today, already more than 50% of the global population lives in urban areas, this figure is expected to increase to over 70% by the year 2050 (CHEN et al. 2012). For this reason, many disciplines deal with the question of how urban regions can be made smarter. The aim of smart cities initiatives is to assess the state of the city as a system and to understand the processes and thus develop tailored and sustainable optimization strategies. For this reason, environmental, ecological and social simulation tools support the urban planners in decision making. The tools require a representation of the simulation area that reflects the real world objects in a most realistic way. Semantic 3D city models have been successfully used in current smart city projects. They serve as reliable source

¹ Lehrstuhl für Geoinformatik, Technische Universität München, Arcisstr. 21, 80333 München, E-Mail [b.willenberg, maximilian.sindram, thomas.kolbe]@tum.de

of information and provide a common information hub with the international standard CityGML as an interoperable representative of semantic 3d city models.

A major challenge is to bring the object based city model together with the field based simulation system in 3D space, as both follow different modelling paradigms. On the one hand, according to the ISO 19109 definition of geographic objects, the city model objects are decomposed following logical criteria which can be observed in the real world (ISO 19109 2005). Their shape, orientation and location in the model is derived from their real world counterparts. All spatial objects carry their inherent semantics. The simulation system on the other hand, divides the observed area into a regular grid of finite volume elements so called volume pixels (voxels) without any thematic information. Real world objects are approximated by an accumulation of these cells.

An essential advantage of semantic 3D city models that is also mandatory for field-based simulation tools is that they provide information about the characteristics of each city object and its parts. On the left side of Fig. 1 a CityGML building is shown in its object-based boundary representation. The building is semantically subdivided into two roof surfaces, 4 wall surfaces and one ground surface. Each of these thematic surfaces can be enriched by thematic information (e.g. area, material, color, ...). This object based representation is overlain by its field based representation that is approximating the shape of the building by a voxel grid. Currently, the semantic information is lost during the geometric derivation of the voxel model from the city model objects.

The central research question of this paper is to find a way to relate both systems – the object based and the field based - to allow information exchange between them. This comprises three parts: (1) the derivation of geometrically and topologically sound voxel models from semantic 3d city models and (2) persisting semantics in the voxels of the field based simulation model as illustrated with the blue voxels in Fig. 1 and (3) the back referencing of simulation results to the city model. The general objective of the discussed research work is the development of an

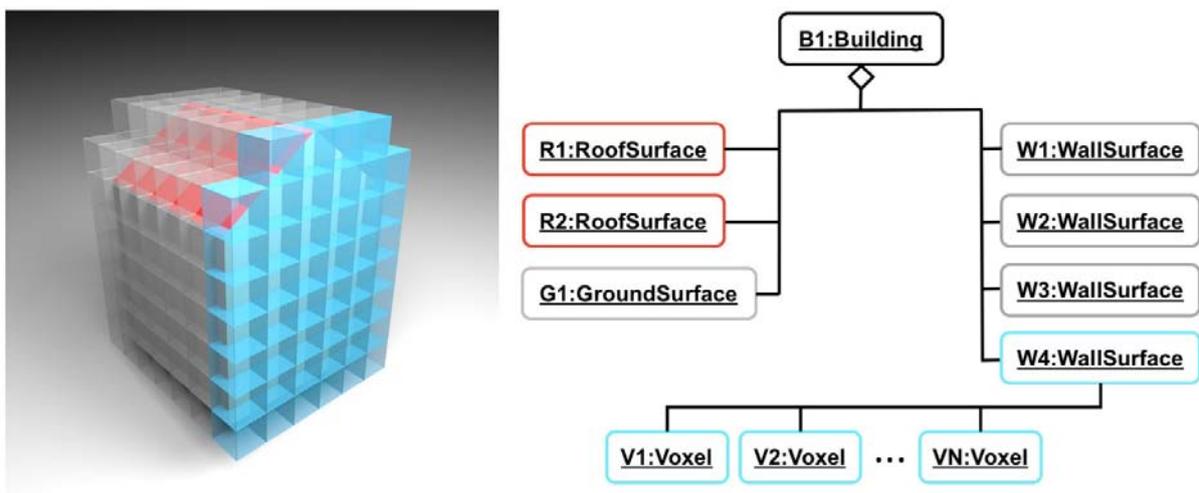


Fig. 1: Object based CityGML building model with its corresponding field based voxel representation. The schema on the right highlights the logical relation between WallSurface 4 and the voxels related to it.

automated workflow for coupling CityGML city models and field based simulation tools, enabling non-expert users to perform simulations and assess their results.

2 Theoretical basis

Today, Computational Fluid Dynamics (CFD) simulation tools support planning and decision making in various fields of applications in an urban context. Examples are given in FADL & KARADELIS (2013) describing a CFD simulation for wind comfort and safety in pedestrian areas or at HANNA et al. (2006) investigating wind flow and dispersion of pollutants in urban space. Research on the topic of 3D noise mapping has been done by STOTER et al. (2008), an example on the detailed 3D simulation of flooding for the city of Munich is presented by VARDUHN et al. (2015). In the field of explosive safety and building construction CFD simulations allow strategic and conceptual preparation for individual blast scenarios (RUTNER et al. 2008; KLOMFASS & THOMA 2012; TROMETER & MENSINGER 2014). These highly tailored applications allow for efficient simulation from different domains, but lack interactive access to simulation results. 3D city models and their frameworks offer comprehensive tools for visualization and result analysis even for non-expert users. They represent a reliable data source for both geometries and semantics in an urban context.

2.1 CityGML

The CityGML standard is an open data model for the representation and exchange of virtual 3D city and landscape models. The model includes information on geometry, appearance, semantics, and topology of urban objects organized in a multi scale approach with four Levels of Detail (LOD). It is a format based on the Extensible Markup Language (XML) and the ISO 19100 standards family. CityGML is an application schema of the Geography Markup Language version 3.1.1 (GML3) which implements the ISO 19107 standard (HERRING 2001). In 2012, its latest issue, version 2.0.0, was released as an official standard of the Open Geospatial Consortium (OGC) (OGC 2012).

The CityGML model not only represents the shape and graphical illustration of city objects. The objects' semantics, their thematic properties, taxonomies, aggregations and interrelations are taken in to account (KOLBE 2009). This enables CityGML to serve as an information model for a broad range of applications (BILJECKI et al. 2015). The rich semantic properties of CityGML's urban objects represent a potential source for input parameters for field based simulations like for example construction materials of buildings. The standard allows the creation of additional attributes and objects, which are not part of the predefined classes and thus, allows its extension for the discussed application context (KOLBE 2009). In the recent years, a framework of tools has evolved around the standard. This gives developers access to well proven and maintained software components for data storage, management and distribution, visualization and assessment tasks.

Moreover, CityGML datasets become increasingly available and have improved in both quality and coverage. For example, about two third of all buildings in Bavaria, Germany, are already available in LoD2 (LDBV 2016). Hence, CityGML models are an attractive data source for simulation tools, as time and cost intensive data acquisition is likely not to be necessary.

Especially simulations for short term response or emergency scenarios like evacuations depend upon existing data.

3 Derivation of Voxel models from CityGML geometries

The host system for the derivation of the voxel representation from CityGML geometries is the 3DCityDB (www.3dcitydb.org). The 3DCityDB maps the CityGML data model onto a relational database schema. It operates on the commercial spatial database management system Oracle Spatial and the Open Source relational database management system PostgreSQL using PostGIS as extension for spatial data types and operations (KOLBE et al. 2009; STADLER et al. 2009; POSTGIS 2015; POSTGRESQL 2016). The whole calculation process is performed on the database where all operations are encapsulated in PL/pgSQL SQL-Procedural Language database functions. This approach is beneficial, as no data but lightweight function calls have to be transferred between the database and the model generation application avoiding network traffic as a bottleneck. The model generation is performed within the spatial reference system the underlying CityGML dataset is presented in. Hence, a projected spatial reference system with e.g. a metric system is required.

Only a few input parameters are needed for the conversion process. First, the thematic layers of the CityGML city model which should be converted to a voxel representation need to be selected. Depending on the given use case, certain aspects of the city model can be ignored to reduce the amount of data that has to be processed.

Second, the simulation domain within the city model is confined by a 3D bounding box (BBox). Additionally, the edge length of the voxels needs to be specified. This parameter defines the resolution of the voxel grid and should be selected according to the quality requirements of the simulation that will be conducted. It strongly influences the performance of both the model

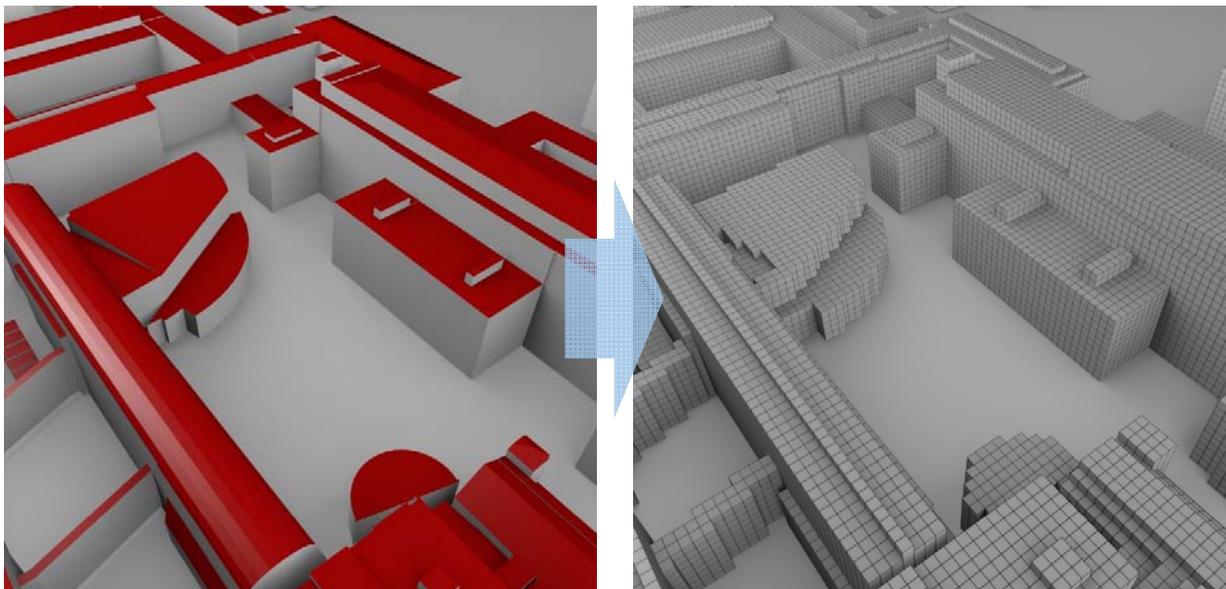


Fig. 2: CityGML buildings of the TUM campus and the derived voxel representation at a voxel edge length of two meter.

generation and the simulation run.

The derivation of the voxel model is performed in two step, the creation of the voxel geometry representation and the identification of voxels intersecting objects of the city model. For the geometry creation the number of voxels required to fill the whole simulation domain for each coordinate axis is calculated first based on the specified simulation BBox and the voxel edge length. Thus, all voxels can be organized in an integer IJK coordinate system, which is oriented as the XYZ Cartesian coordinate system of the city model, as depicted in Fig. 3. Each voxel is therefore uniquely identified by its IJK coordinates. By passing the voxels IJK coordinates, the XYZ origin of the coordinate system, which is the lower left bottom of the simulation domain, and the voxel edge length to a database function, the voxel geometry can be computed.

In the function a Well-Known-Text (WKT) representation (OGC 2011) of the voxels geometry is created which is scaled by the voxel edge length and translated to its position in the voxel grid relative to the given origin by multiplying the IJK coordinates with the voxel edge length. The voxels geometry is returned as a PostGIS PolyhedralSurface spatial object by the voxel creation function (OGC 2010, 2011).

Now, the voxel geometry objects interacting with city model objects have to be identified. They are tested for intersection with geometries of the city model using 3D intersection operations of PostGIS. Thereby, the GiST index structures provided by the database system are used (HELLERSTEIN et al. 1995). They implement an R-Tree spatial index which increases query performance using a tree data structure for bounding box comparisons (GUTTMAN & STONEBRAKER 1983). An example of a voxel model derived from a CityGML model of the Campus of Technical University Munich (TUM) is illustrated in Fig. 2.

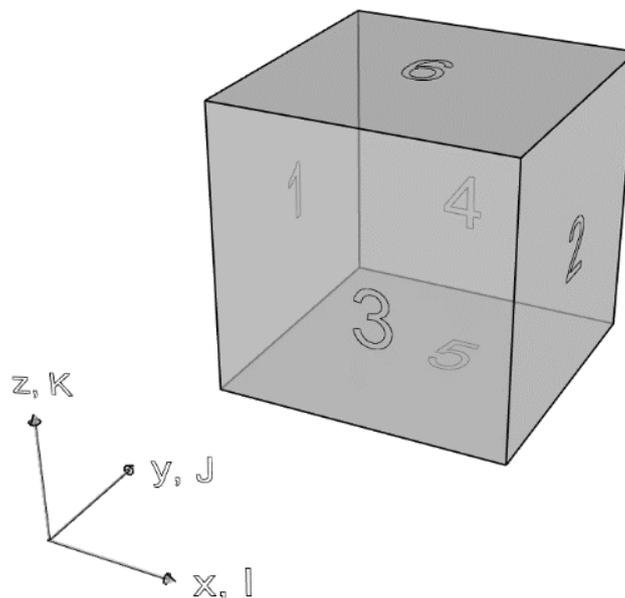


Fig. 3: Voxel with Cartesian XYZ coordinates system, integer IJK coordinate system and numbering schema for voxel faces. Face one is located on the minimum side of yz plane, face two is located on the maximum side of the yz plane. The other faces are numbered accordingly.

4 Information exchange between city and voxel model

When identifying voxels with spatial relation to the city model geometries, only voxels with a positive result for the intersection test are saved. These voxels are stored in a VOXEL database table with their IJK coordinates, their geometry and a unique identifier (UID_VOXEL). In another database table named INTERSECTION the logical link between the voxels and the city model geometries is stored as tuples of the UID_VOXEL and the SURFACE_GEOMETRY_ID, the unique identifier of the city models' geometries. With this n:n relationship it is possible to put single city model geometries or whole city objects in relation with single voxels. The numbering schema depicted in Fig. 3 even allows a unique identification and relation of individual voxel faces. Hence, the information exchange between city model entities and the voxel model discussed in section 1 is now feasible using standard database joins. The semantic information contained in the city model can be query from the database for use in a simulation application that is operating on the derived voxel representation.

For allowing interactive visualization and analysis within the city model the simulation results need to be stored in the city model objects. This can be done using CityGML's extension mechanisms. Using the above mentioned logical relation between both models, simulation results that are returned per voxel or voxel face, can be aggregated per related city object and persistently stored with it as generic attribute. The enriched city model objects are now usable for further analytic or visualization tasks based on the simulation results.

The presented semantic enrichment of the city model with simulation results poses an added value for both the city model data and simulation tool, as linking the simulation results with the already existing information in the model opens new possibilities for assessment of the data.

5 Example usage scenario: Blast simulation with the Apollo Blastsimulator

In the following section the proposed approach for the integration of CFD simulation tools and 3D city models is evaluated on the example of a blast simulation. The work described in this section is inspired by the Master Thesis of WILLENBORG (2015).

5.1 Apollo Blastsimulator

The Apollo Blastsimulator is a CFD simulation tool developed at Fraunhofer Institute for High-Speed Dynamics, Ernst-Mach-Institute (EMI), in Freiburg, Germany. According to KLOMFASS et al. (2013) and KLOMFASS & HERZOG (2010), when analyzing blast effects mostly two kinds of calculation models are available. First, there are strongly simplified approaches based on TNT equivalences and scaled distances. They work very fast, but have restrictions regarding accuracy and the spectrum of permitted applications. Second, there are general purpose CFD simulations delivering both accurate results and usability by a broad range of applications for the price of difficult handling and long setup and computation times. The Apollo Blastsimulator tries to fill the gap between both approaches and combines good usability, versatility and computational efficiency. This is achieved by tailoring the methodological concepts to the application of explosions and blast waves.

The computational mesh the application works on, needs to be computed for the observation area prior to the simulation run. The mesh consists of square finite volume elements that can be created from CityGML data with the presented approach. Every solid structure in the simulation domain needs to be approximated by a closed shell of voxels. After a simulation run, the Apollo Blastsimulator returns two types of results. Besides physical quantities a set of probability values for various damage categories is provided. For each voxel surface exposed to fluid (air) in the simulation domain a set of result values is returned.

5.2 System architecture and workflow

To provide a user friendly application the described approach for the voxel model creation has been implemented as a plugin for the 3DCityDB Importer/Exporter, the management software of the 3DCityDB. Besides its data import/export capabilities for various formats, a powerful Application Programming Interface (API) is shipped, enabling developers to extend the tool with new components, implemented as plugins. The API offers a wide range of ready-to-use core functionality like database interaction, GUI elements, parallelisation strategies or logging, making plugin development easy (NAGEL & KADEN 2011). The programmer can rely on a well-tested application backbone and focus on the logic of the use-case.

The core element of the workflow is the *Apollo Blast Plugin*, the plugin extension for the 3DCityDB Importer/Exporter implemented in this study. It operates as the control and data exchange unit between simulation, city model, visualization, and result analysis.

The plugin provides a GUI allowing the user to specify a simulation extent on an interactive 2D map, the explosion location, and the explosive device. Further, the calculation approach, either engineering method (simplified approach based on TNT equivalences and scaled distances) or CFD simulation, can be selected. For the CFD simulation the computational time and result quality can be configured by adapting the resolution of the voxel mesh. This is required for multi-stage analysis, typically conducted in risk analysis. First, a big amount of scenarios is analysed with the engineering method to identify potentially critical scenarios, which are re-evaluated with the high quality CFD simulation afterwards (KLOMFASS et al. 2013). All settings are persistently stored on disk in an XML configuration file to facilitate easy reuse and exchange. For the creation of the voxel model a parallelization strategies has been implemented. The IJK BBox of the simulation domain is split into slices along the I axis. This allows to process several slices on a given number database connections in parallel to leverage multiple cores of the database server and significantly reduce the time consumption of the conversion process.

In a typical workflow, the user first creates a scenario definition with the above settings in the Apollo Blast Plugins GUI. After that, the ASCII input files for the simulation tool, containing the computational mesh, are computed and passed to the simulation application. When the Apollo Blastsimulator has terminated, the returned ASCII results files are read and processed by the plugin and all affected CityGML objects are updated within the connected 3DCityDB.

5.3 Aggregation and storage of simulation results

After a successful run, the simulation results need to be transferred back to the city model. First, the Apollo Blastsimulator ASCII results file is read by the Apollo Blast Plugin. It contains a set of results values for each voxel surface exposed to fluid (air) in the simulation domain. Now, a

two level aggregation is performed. First, the maximum and average of all result parameters is aggregated per voxel. After that, all voxels related to a CityGML Thematic Surface (e.g. Wall-, Roof or GroundSurface) are aggregated based on the information stored in the intersection table during the generation of the voxel mesh. The summarized results are stored as generic attributes, one for each result parameter, with their corresponding city object in the 3DCityDB. Now all city objects in the simulation domain contain attributes for physical quantities (overpressure, overpressure impulse) and a set of probability values for damage classes like glass breakage, eardrum damage, destruction of masonry, or concrete walls and lethality. The semantically enriched city model now allows for thematic queries regarding the simulated blast event.

5.4 Visualization and result analysis

For visualization and result analysis a cloud-based 3D Webclient developed at the Chair of Geoinformatics, TUM is used. The browser based application uses the Cesium Earth Viewer (see <https://cesiumjs.org/>) to 3D visualize the city model. The city model is reduced both geometrically and semantically to the requirements of the specific application context (KOLBE et al. 2013; YAO et al. 2014; YAO & KOLBE 2014). Geometries are represented in the KML/COLLADA format, which is exported by the 3DCityDB Importer/Exporter KML/COLLADA exporter plugin (KOLBE et al. 2013). Thematic information, including the generic attributes created for the blast simulation results are exported to the Spreadsheet Generator Plugin and published via GoogleSpreadsheet allowing to apply spreadsheet calculations. The relation between both models is kept with a logical key (GMLID) (NADERI et al. 2012).

For demonstration purpose an artificial test scenario has been developed. We assume, that a

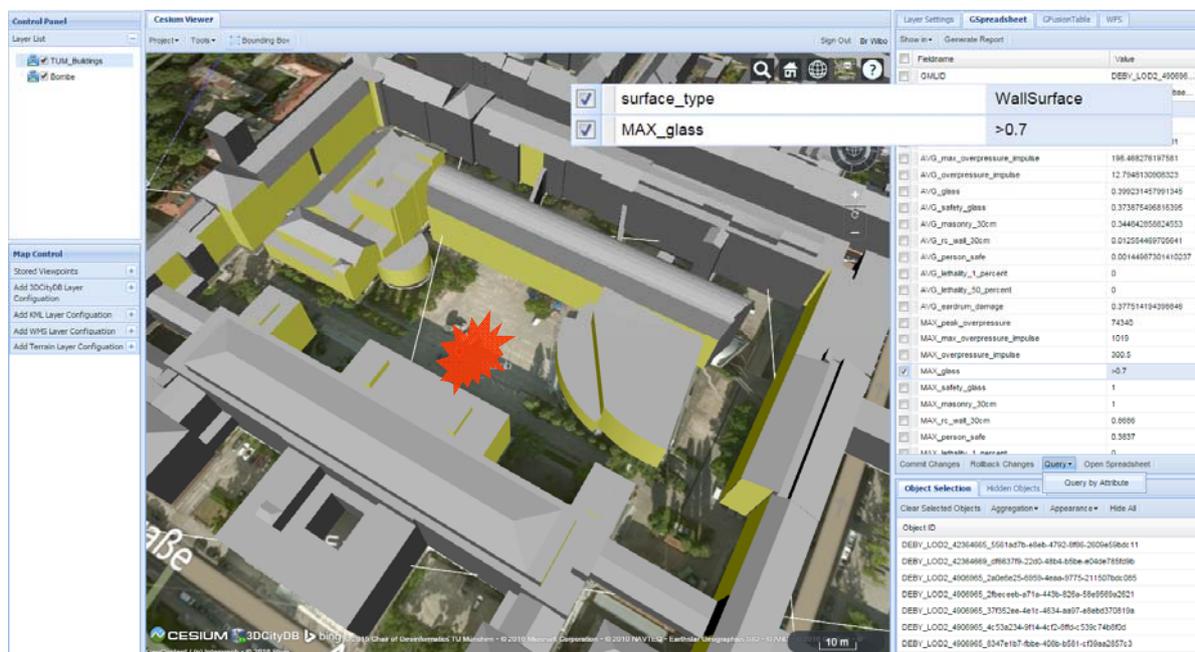


Fig. 4: Cloud-based 3D Webclient example of a fictive blast scenario on TUM campus: Wall surfaces with a glass breakage probability greater 70% are highlighted in the 3D view.

bomb from world war two has been discovered during ground working in the inner court of TUM. For the test scenario only the CityGML building layer has been used, all buildings have been translated to a plane on elevation zero meters.

To display the analytic capabilities of the Webclient we will try to identify walls, in which windows are likely to break if the bomb cannot be defused and needs to be detonated on site.

First, we need to setup the query for the required information. As depicted in Fig. 4, we type the filter criteria in the attribute view to the right (see enlarged entries). After issuing the query all matching surfaces are highlighted (yellow) in the 3D view. For further analysis aggregate operations on the selected objects can be performed directly in the client. For example, by summing up the area of all currently selected surfaces we can determine the total affected wall surface. By multiplication with factors for window area on a wall and window price per area we can provide a rough cost estimation for broken windows.

6 Discussion and evaluation

Besides its benefits mentioned in the previous sections, there are some limitations and problems with the proposed approach which are discussed in this section.

One issue concerns the quality of the derived voxel model. Since all geometries are reassembled by finite volume elements for the field based representation the orientation of the voxel grid relative to the input geometry greatly influence the model quality. As depicted in Fig. 5, if the grid is aligned with the input features a smooth voxel representation of the building wall surfaces is created while the transformation with no alignment causes a ripple effect on the surface approximation. Possible solutions are the alignment of the voxel grid with its input geometries. As this is not always possible in the context of 3D city models containing real world objects which are not always rectangular an increased grid resolution will be required to avoid ripple effects.

Another problem influencing the model quality are invalid city model geometries like non planar surfaces. They cause the intersection algorithms to fail consequently leading to holes in the voxel model. As CFD simulations require objects to be represented by a closed shell of voxels, this is a big issue. As the quality and validation of 3D city models is a matter of ongoing research, this is not further discussed here (ALAM et al. 2013).

Regarding the different modelling paradigms between field based simulations and object based 3D city models with boundary representation geometries there are some issues, that require further research. Due to different scales between single objects in city models and field based simulations (e.g. size voxel \ll size WallSurface) simulation results cannot be mapped in full detail to the city model. When aggregating the simulation results attached to several voxels to a single city object information is lost, results may even be falsified statistically. This issue could be addressed by the generation of raster textures which are supported by the CityGML appearance model. A combination of textures representing in depth details delivered by a simulation and aggregated attributes that can be subject to queries would foster analytic capabilities including a high level of detail for simulation results in 3d city models.

Another issue when trying to represent and store simulation results in CityGML is the absence of dynamic attributes. CFD simulations usually provided time series of results, which are not

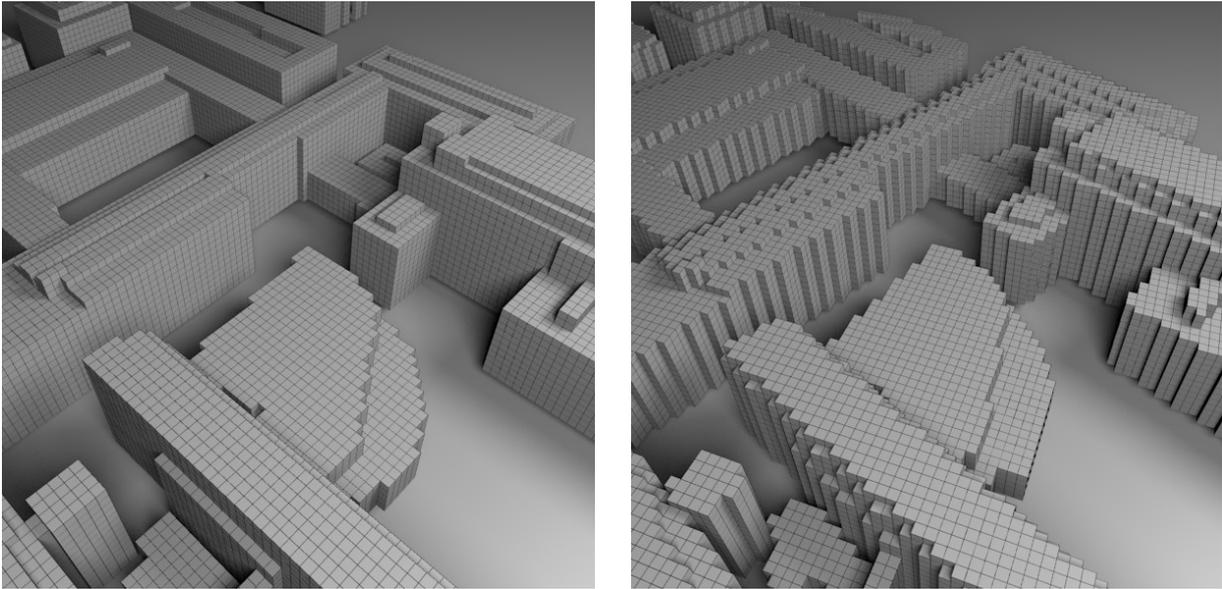


Fig. 5: Comparison of the model quality between CityGML buildings aligned with the voxel orientation before model creation (left image) and without alignment (right image).

included in the data model at this point. However, this topic is currently being researched by CHATURVEDI & KOLBE (2015) and is expected to be implemented for the next version of CityGML.

7 Conclusion

The central research question of this study was to relate semantic 3D city models and field based simulation models and their different modeling paradigms to allow information exchange between both of them. Based on the 3DCityDB an approach for deriving a field based geometry representation from a CityGML city model using spatial database operations provided by PostGIS was introduced. By mapping the results of 3D intersection tests between voxel and city model geometries to a relational database table information exchange between both systems with common database join operations could be realized. Accordingly, the rich semantic information of the city model is available to field based simulation tools. The other way round the back referencing of simulation results to their corresponding city objects is now feasible. Using the extension capabilities of CityGML a semantic enrichment of the city model with aggregated simulation results is performed based on the proposed approach. This permits visualization and assessment of simulation results within the city model and its framework. The generated information can be accessed in an interactive, cloud based environment allowing non-expert to perform analytical tasks. As simulation results are persistently stored with their city model objects they remain available for future analysis, possibly including new intelligence from various sources like simulations, and therefore pose a consistent added value for the data.

References

- ALAM, N., WAGNER, D., WEWETZER, M., FALKENHAUSEN, J. v., COORS, V. & PRIES, M., 2013: Towards automatic validation and healing of CityGML models for geometric and semantic consistency. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, **II-2/W1**, 1-6.
- BAYRISCHES LANDESAMT FÜR DIGITALISIERUNG, BREITBAND UND VERMESSUNG, 2016: LOD2 Arbeitsstand.
- BILJECKI, F., STOTER, J., LEDOUX, H., ZLATANOVA, S. & ÇÖLTEKIN, A., 2015: Applications of 3D City Models: State of the Art Review. *IJGI* **4** (4), 2842-2889.
- CHATURVEDI, K., KOLBE, T.H., 2015: Dynamizers - Modeling and Implementing Dynamic Properties for Semantic 3D City Models. *Eurographics Workshop on Urban Data Modelling and Visualisation*, Filip Biljecki & Vincent Turre (Hrsg.).
- CHEN, A.S., EVANS, B., DJORDJEVIC, S. & SAVIC, D. A., 2012: Multi-layered coarse grid modelling in 2D urban flood simulations. *Journal of Hydrology* **470**, 1-11.
- FADL, M.S. & KARADELIS, J., 2013: CFD Simulation for Wind Comfort and Safety in Urban Area: A Case Study of Coventry University Central Campus. *IJAEC* **2** (2), 131-143.
- GUTTMAN, A. & STONEBRAKER, M., 1983: R-trees. Electronics Research Laboratory, College of Engineering, University of California, Berkeley.
- HANNA, S.R., BROWN, M.J., CAMELLI, F. E., CHAN, S. T., COIRIER, W. J., KIM, S., HANSEN, O. R., HUBER, A. H. & REYNOLDS, R. M., 2006: Detailed Simulations of Atmospheric Flow and Dispersion in Downtown Manhattan: An Application of Five Computational Fluid Dynamics Models. *Bull. Amer. Meteor. Soc.* **87** (12), 1713-1726.
- HELLERSTEIN, J.M., NAUGHTON, J.F. & PFEFFER, A., 1995: Generalized Search Trees for Database Systems. *Proceedings of the 21. International Conference on Very Large Data Bases*.
- HERRING, J., 2001: OGC Abstract Specifications Topic 1 – Feature Geometry (ISO 19107 Spatial Schema).
- ISO 19109, 2005: Geographic information — Rules for application schema.
- KLOMFASS, A., HERZOG, O., 2010: A Reliable, Fast and Easy-to-use Simulation Tool for Blast Waves in Urban Scenarios. Fraunhofer Institute for High-Speed Dynamics, Ernst-Mach-Institute. Freiburg, Germany.
- KLOMFASS, A., THOMA, K., 2012: CFD and Holistic Methods For Explosive Safety and Risk Analysis. *Process and Plant Safety* **9**, 9-29.
- KLOMFASS, A., ZWEIGLE, T. & FISCHER, K., 2013: A New AMR/ALE-Solver for the efficient Simulation of Detonations and Blast Waves. *Intern. Symposium of Interaction of Effects of Munitions with Structures*.
- 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. (Hrsg.), 15-31.
- KOLBE, T.H., KOENIG, G., NAGEL, C. & STADLER, A., 2009: 3D-Geo-Database for CityGML. *Proceedings of the 3rd International Workshop on 3D Geo-Information, Lecture notes in Geoinformation and Cartography*, Lee, J. & Zlatanova, S. (Hrsg.).

- KOLBE, T.H., YAO, Z., HERRERUELA, J., NAGEL, C., KUNDE, F., WILLKOMM, P. & HUDRA, G., 2013: Addendum to the 3D City Database Documentation Version 2.0.1.
- NADERI, B., HERRERUELA, J. & KOLBE, T.H., 2012: Spreadsheet Generator Version 1 .0.0 Plugin for Importer/Exporter Version 1.4.0.
- NAGEL, C. & KADEN, R., 2011: Importer/Exporter - Plugin API - Slides for internal discussion. Institute for Geodesy and Geoinformation Science, Technical University Berlin.
- OPEN GEOSPATIAL CONSORTIUM, 2010: OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 2: SQL option.
- OPEN GEOSPATIAL CONSORTIUM, 2011: OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture.
- OPEN GEOSPATIAL CONSORTIUM, 2012: OGC City Geography Markup Language (CityGML) Encoding Standard. Version 2.0.0.
- POSTGIS PROJECT STEERING COMMITTEE, 2015: PostGIS 2.2.0dev Manual.
- RUTNER, M., MENSINGER, M. & TROMETER, S., 2008: Die Auswirkung von duktilen Anschlüssen auf die Explosionsresistenz lasttragender Stützen. *Stahlbau* **77** (6), 436-446.
- 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. (Hrsg.), 175-192.
- STOTER, J., KLUIJVER, H. de & KURAKULA, V., 2008: 3D noise mapping in urban areas. *International Journal of Geographical Information Science* **22** (8), 907-924.
- THE POSTGRESQL GLOBAL DEVELOPMENT GROUP, 2016: PostgreSQL 9.5.1 Documentation.
- TROMETER, S., MENSINGER, M., 2014: Simulation von Detonationsszenarien im urbanen Umfeld. *Geoinformationssysteme 2014, Tagungsband zur 1. Münchner GI-Runde*, Kolbe, T. H., Bill, R., Donaubaue, A. (Hrsg.), 150-164.
- VARDUHN, V., MUNDANI, R.-P. & RANK, E., 2015: Multi-resolution Models: Recent Progress in Coupling 3D Geometry to Environmental Numerical Simulation. *3D Geoinformation Science. Lecture notes in geoinformation and cartography*, Breunig, M., Al-Doori, M., Butwilowski, E., Kuper, P. V., Benner, J., Haefele, K.H. (Hrsg.), 55-69.
- WILLENBORG, B., 2015: Simulation of Explosions in Urban Space and Result Analysis Based on CityGML City Models and a Cloud Based 3D Web Client, Masterarbeit. Technische Universität München, München.
- YAO, Z., KOLBE, T.H., 2014: Cloud-basierter Webclient für kollaboratives Arbeiten auf semantischen 3D-Stadtmodellen. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e.V*, Band **23**, Beitrag 234.
- YAO, Z., SINDRAM, M., KADEN, R. & KOLBE, T.H., 2014: Cloudbasierter 3D-Webclient zur kollaborativen Planung energetischer Maßnahmen am Beispiel von Berlin und London. *Geoinformationssysteme 2014, Tagungsband zur 1. Münchner GI-Runde*, Kolbe, T.H., Bill, R., Donaubaue, A. (Hrsg.), 40-52.

Challenges for the Swiss Federal Office of Topography swisstopo for the Production of Digital Images and Altimetrical Models using the ADS80 Sensor for the Glacier Monitoring in the Swiss Alps

ROBERTO ARTUSO¹, NATHALIE BÉDERT¹ & FLORIAN GANDOR¹

Abstract: The Swiss Federal Office of Topography swisstopo produces since 2010 digital data on Swiss glaciers for the Federal Office for the Environment (BAFU) and the Laboratory of Hydraulics, Hydrology and Glaciology (VAW) of the Swiss Federal Institute of Technology of Zürich. The data comprise digital ADS80 (Airborne Digital Scanner) oriented aerial pictures with a Ground Sample Distance (GSD) of 25 cm, derived Digital Surface Models (DSM) with a 50 – 100 cm grid as well as orthophotos with a 25 cm GSD. These geodata are part of the national glaciers monitoring program GLAMOS (GLacier MOnitoring in Switzerland).

This paper describes specific aspects of the data acquisition (flight planning over glacier areas), the geometrical referencing with the help of Ground Control Points, as well as the DSM production using image matching. Glacier areas are characterized by an important topographical heterogeneity: a big reflectance over glacier areas and lots of shadows due to the mountains. These aspects are a specific challenge for image correlation and the production of accurate and complete DSM.

Details concerning the softwares used and tested by swisstopo are provided here, as well as which bands of the ADS80 scanner and which parameters are used for the production.

This paper provides an overview of swisstopo's tests and criteria (‘Benchmarking’) prior to the acquisition of new softwares for image matching using ‘Semi Global Matching’ algorithms. The comparison of the performances shows neat differences in data quality, data volume and time performances during the production of derivate products. Those are the criteria which are set for deciding which software offers the best quality and is the most cost-efficient for a future use for the production of DSMs with a higher resolution at a larger scale, for example for the whole swiss territory.

1 Introduction

Global warming causes a world-wide spread glacier withdrawal. The Swiss Alps are located in the south and southeastern part of the country. Their highest elevation reaches 4'500 m a.s.l. The Swiss Alps count numerous glaciers which are actively monitored since 1874, as the first measurement took place on the Rhône Glacier.

The GLAMOS project (GLacier MOnitoring in Switzerland) conducts each year glacier variations measurements on more than a hundred glaciers of Switzerland (FRANK et al. 2011). This project involves several partners, including the swiss federal office for the Environment BAFU, the Laboratory of Hydraulics, Hydrology and Glaciology (VAW) of the Swiss Federal Institute of Technology in Zürich (ETHZ), the Swiss Academy of Sciences, the Global Climate Observing System (GCOS) Schweiz and the Swiss Federal Office of Topography swisstopo.

¹ Bundesamt für Landestopografie swisstopo, Seftigenstrasse 264, CH-3084 Wabern,
E-Mail: [Roberto.Artuso, Nathalie.Bédert, Gandor.Florian]@swisstopo.ch

Fluctuations in glacier length and changes in volume and mass represent key parameters in environmental monitoring. Their measurement is achieved using multiple techniques, including geophysics for bedrock altitude determination (HAUCK et al. 2012) and aerial photogrammetry for determination of the surface altitude with a Digital Surface Model (DSM).

swisstopo provides ADS80 aerial images and derivate DSM material since 2010 for the needs of GLAMOS. The delivered products consist in a derived DSM and ortho-rectified aerial NRGB (16 bits) and RGB (8 bits) images called L2 with a 25 cm Ground Sample Distance (GSD). The actual DSM is produced using SocetSet NGATE 5.6.0 by BAE Systems with a 1 m grid. However, swisstopo is conducting a series of tests to produce a more detailed DSM which means that the resolution has to be enhanced.

Two specific softwares, XPro-DSM 6.3.1 by Leica Geosystems and Socet GXP 4.1.0 by BAE Systems are currently undergoing further investigations.

This study seeks also to initiate an improvement of the production for swissALTI^{3D}. swissALTI^{3D} is a 2 m grid Digital Terrain Model (DTM) covering the whole swiss territory produced by swisstopo and updated in a 6 years cycle. swissALTI^{3D} combines laser data, automatically correlated data in local areas and manually edited areas using 3D devices.

The specific aims of this study are:

- (1) To evaluate the performances of XPro-DSM 6.3.1 by Leica Geosystems and Socet GXP 4.1.0 by BAE Systems;
- (2) To determine for both softwares which parameters result in the most accurate and detailed Digital Surface Model and at what processing costs;
- (3) To discuss the possible applications for both GLAMOS and swissALTI^{3D}.

2 Study area

2.1.1 Study area for this paper

The area of Bis, Hohlicht and Turtmann, located in the canton of Valais was chosen for this specific study for the production of altimetrical models over glaciers (Fig. 1). This area was chosen for multiple reasons: it is a big surface characterized by heterogeneous altitudes ranging from 2'096 meters up to 4509 meters, including the Matternal Valley and the village of Zermatt.

The mountainous area is covered with glaciers with various expositions to the sun, which in some cases cause a very high light reflectance. Moreover, the mountains provide a lot of shadow areas. The Matternal Valley offers a good contrast to the glaciers mountainous glacier areas. The reflectance is smaller. These heterogeneous characteristics make this area ideal for testing various parameters, so as to ultimately apply them on bigger areas.

2.1.2 Study area for GLAMOS

The study area is located in the Swiss Alps, in the south – south eastern part of Switzerland. While the GLAMOS project includes over 100 glaciers, aerial imagery material is provided for 23 areas covering some of the biggest Swiss glaciers, as seen on Fig. 1 and Tab.1. Most of them are produced on a yearly basis.

Tab. 1: Areas of interest flown for GLAMOS

Area of interest	Flight frequency [year]	Footprint [km ²]	Number of flight lines
Finster- Ober- Lauter- Unteraar	1	150.1	4
Basodino	3-5	77.9	2
Birch	1	22.6	1
Bis Hohlicht Turtmann	1	252.6	2
Findel	1	36.4	1
Gauli	1	18.1	1
Gorner	1	34.9	1
Gries	1	60.8	1
Grosser Aletsch Süd	1	50.3	1
Gutz	1	51.9	1
Oberer, Unterer Grindelwald	1	73.1	2
Plaine Morte	1	203.3	3
Rhone	1	16.6	1
Silvretta	1	50.4	1
Trient-Dolent-Saleina	5	193.9	3
Trift (Gadmen)	1	15.3	1
Trift (Weissmies)	1	8.3	2

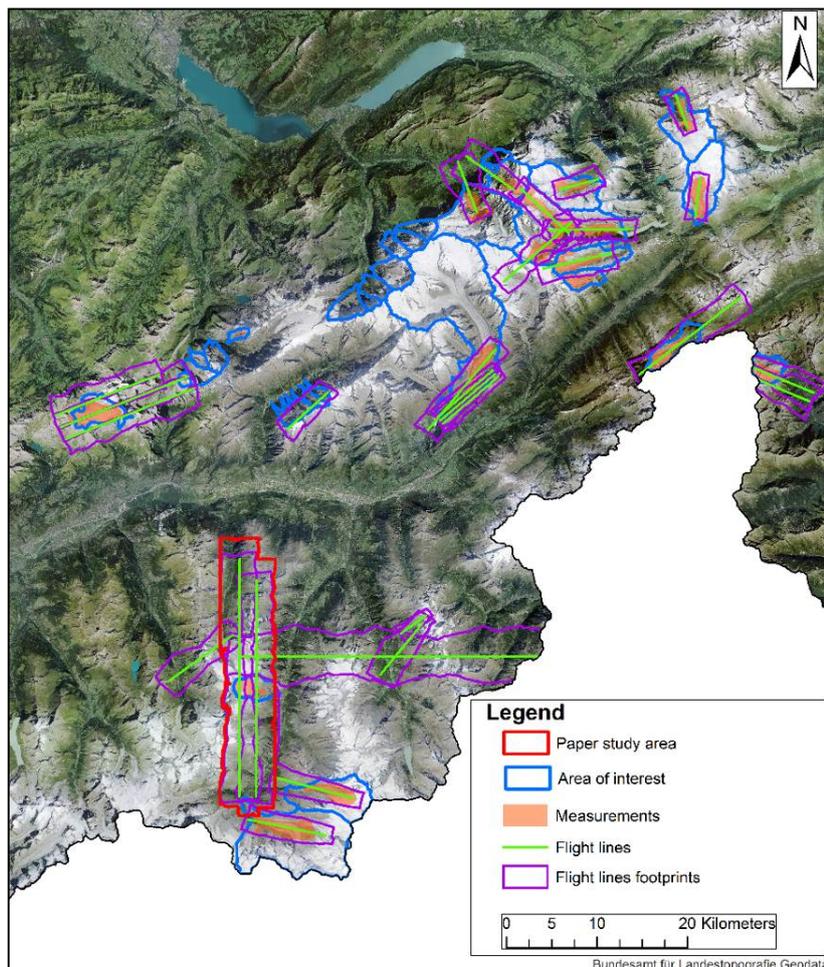


Fig. 1: Areas Of Interest flown by swisstopo for the needs of GLAMOS and for this paper

3 Material, methods and products

No material other than the current sensors were used for this study. This study aims only the software, material's performances are not evaluated. Airborne optical sensors allow large areas of interest to be covered at reasonable costs. Those big areas can only be efficiently processed thanks to automated processing using image matching methods. Derived products include L1 images, DSMs and L2 images.

3.1 ADS80 Stereo Aerial Images

The Leica Geosystems ADS80 camera is an Airborne Digital pushbroom Sensor for line scanning, in our case mounted with a 62.7 mm focal length Sensor Head called SH82. Two cameras have been acquired and used since 2009 by the Swiss Federal Office of Topography swisstopo. This sensor acquires five spectral bands with a radiometric resolution of 12 bits; the GNSS/IMU supports orientation of the image strips achieves a one GSD accuracy (HOBI & GINZLER 2012; WAGNER 2014).

The 12 bits bands are converted to 16 bits. By doing so, no information is added to the image and the quality remains unchanged.

Tab. 2: Technical details of the ADS80 – SH82 sensor head

Sensor	Dynamic range of the CCD	Spectral bands [nm]	CCD elements	Pixel size [µm]	Focal plate configuration
3 – line CCD scanner	12 bits	Panchromatic (PAN) 465-680 Blue (B) 428-492 Green (G) 533-587 Red (R) 608-662 Near infrared (N,NIR) 833-887	12 CCD lines with 12'000 pixel each	6.5	PAN F27° PAN F02° (A,B) PAN B14° Blue N0° Blue B16° Green N0° Green B16° Red N0° Red B16° NIR N0° NIR B16°

3.2 Pre-flight preparation

3.2.1 Flight planning for GLAMOS

The number of flight lines, their orientation and the flight altitude are determined by the glacier's ice tongue axe and the desired pixel size over the main focus area. Whilst a mean GSD of 25 cm is aimed at, the minimum and maximum GSD values may range from 12 cm to 47 cm due to obvious geographical heterogeneities in mountainous areas. The position and length of the flight lines are also determined according to stable areas offering good Ground Control Points (GCPs). The flights occur when the snow cap is at its minimum, from July until September and around midday if possible to minimize shadows.

3.2.2 Ground Control Points

The location of the GCPs used for the georeferencing process is determined and measured previously to the flights. The expected accuracy for the final GCPs field coordinate

measurements should be smaller than 3 cm, which can be achieved with a GNSS measurement. The GCPs must be easily photo-identifiable, visible and should be located in geologically stable areas. As the pictures for GLAMOS are flown with a 25 cm mean GSD in mountainous areas, 90 times 90 cm black and white squares are painted in stable, obstacle and landslide-free areas, as shown on the picture below. These squares are big enough and offer a good contrast to be recognizable on the raw pictures.

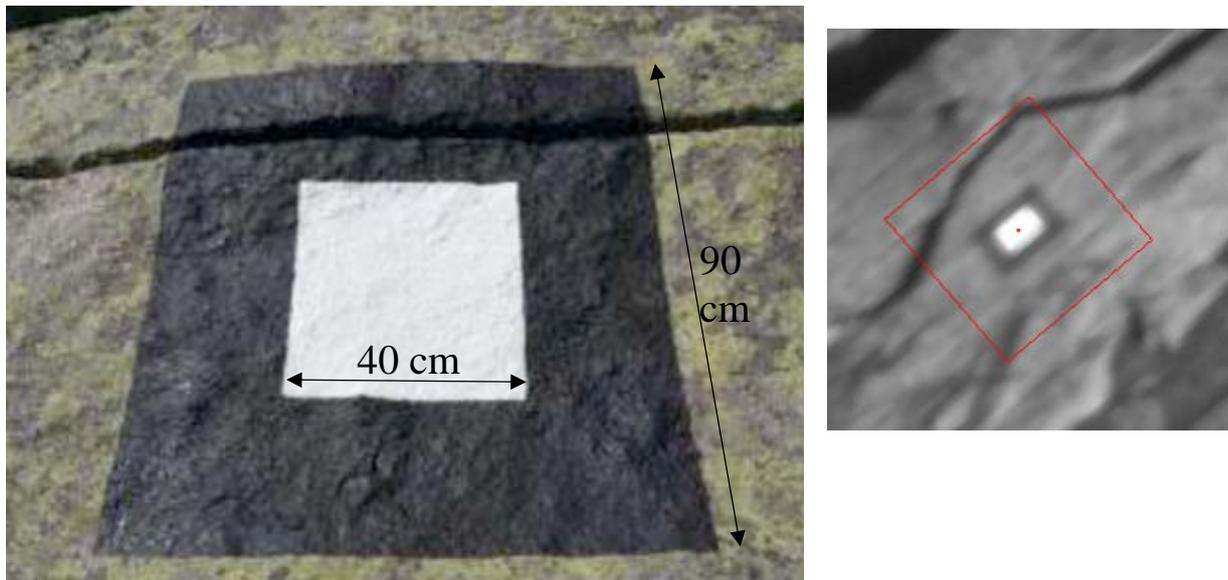


Fig. 2: Example of a typical GCP for GLAMOS on the terrain (left) and on the raw image (right)

3.3 Aerotriangulation

The aerotriangulation is a semi-automated process improving the *a priori* georeference of the images, provided by the GPS/IMU, thanks to tie points and GCPs.

Tie points are determined, in the case of the ADS camera, between the different viewing angles of the same flightline and between neighbor flightlines, in their overlapping area. This process is called relative external orientation and is carried out by Automated Point Matching (APM).

The GCPs are measured in the images. Since the coordinates of these GCPs are known, they provide the absolute external orientation. Having good GCPs is very important for the aerotriangulation. The absence of GCPs, GCPs with a bad precision or a bad geographical distribution of the GCPs on the area can have a negative impact on the precision of the aerotriangulation and therefore, on all the derived products too, such as the L1, DSMs and L2 products (see next paragraph).

Fig. 3 provides a schematic view of the production chain for ADS80 pictures and derived products.

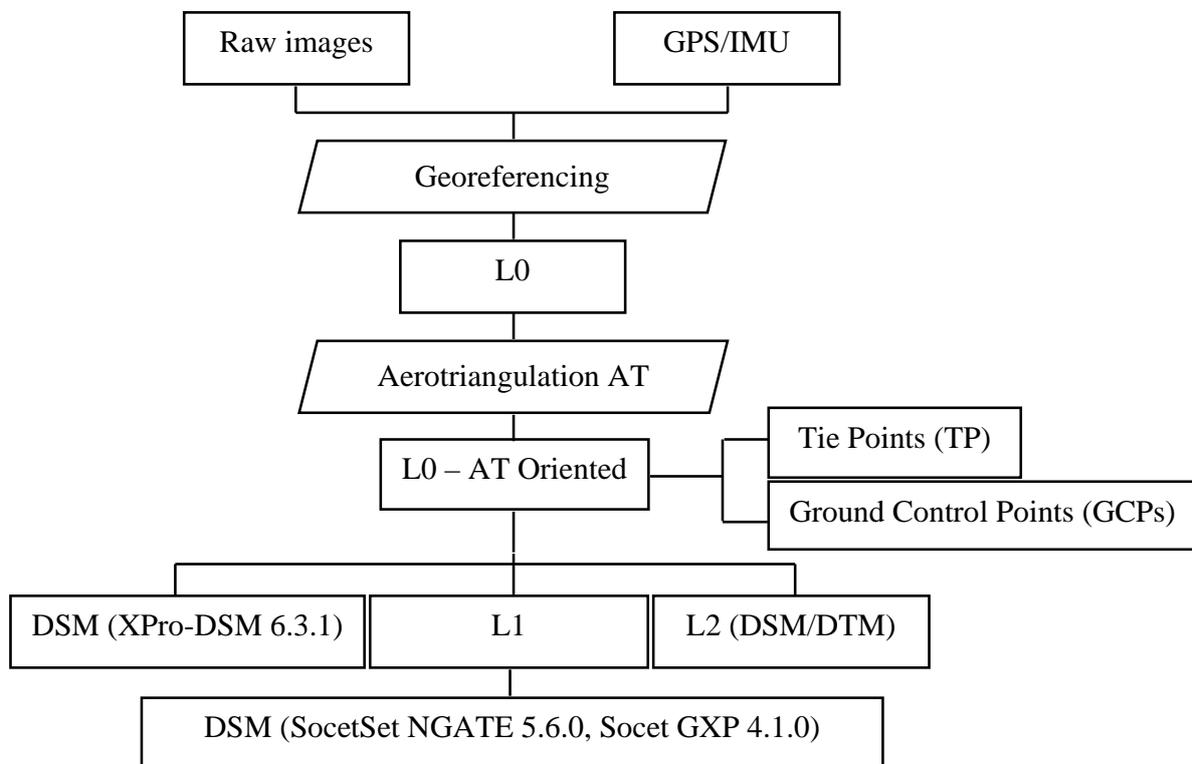


Fig. 3: Schematic production chain from the raw images to the L2 images

3.4 Derivate products

3.4.1 Images

Raw images from the ADS80/SH82 push-broom sensor, also called level-0 L0 images, consist of individual ground scan lines along a defined flight path. Although minimized by an active stabilization system, residual movements of the sensor still affect the raw images during the scanning process.

The system is mounted with a high-frequency IMU system that measures literally all the movements of the sensor for each line. Geometric distortions due to camera movements and present in the L0 raw pictures are corrected during the L1 image generation process with a given mean terrain height, internal and external orientation parameters. L1 oriented image contains 3D information which can be exploited with the stereo principles.

Level-2 L2 are 2D images ortho-rectified on the basis of a given Digital Terrain Model.

3.4.2 DSM Generation

In mountainous areas located at altitudes ranging from 2'000 to 4'500 m, forests and houses are scarce. It is thus assumed that the Digital Terrain Model (DTM) is equivalent to a DSM. DSM production is a highly automated process. Image-matching softwares identify corresponding pixels for an image pair and reconstruct their position using given extrinsic and intrinsic orientation (BÜHLER et al. 2012). swisstopo actually produces DSMs using SocetSet 5.6.0 by

BAE Systems with the “Next Generation Automatic Terrain Extraction” NGATE module. However, the needs for DSMs offering a higher resolution is increasing. A higher resolution comes along with a set of challenges, processing time being one of the biggest. Therefore, the alternative image matching techniques must be both exact and cost-efficient (DEVENECIA et al. 2007). swisstopo is currently testing softwares to compute ADS80-derived DSM, using Semi-Global Matching (SGM) and Automatic Spatial Modeler (ASM) techniques, which we describe hereafter.

The actual image matching is performed using SocetSet 5.6.0 by BAE Systems with the NGATE module. NGATE computes an elevation for every pixel using an imaging correlation method combining edge and image-matching algorithms. The input images are L1 oriented images. This solution has proven to be both effective and cost-efficient, offering a good compromise between image-matching quality and calculation time on big areas for both GLAMOS and swissALTI^{3D}. A fixed spacing of four times the GSD is defined to resample the internal irregular NGATE point cloud into the desired grid (1 m for GLAMOS, 2 m for swissALTI^{3D}). The near infrared – red – green (CIR) bands combination interpolation with the “low contrast strategy” produces the best quality in high alpine glacier regions and is used for the automated correlation.

Meanwhile, SocetSet tends to be outdated since it has been replaced by the more recent Socet GXP by BAE Systems, which also includes the NGATE algorithm.

Automatic Spatial Modeler (ASM) is based on the same method as NGATE. The former algorithm is however enhanced by using adaptive pixel aggregation, which enforced a local smoothness constraint (ZHANG 2014). To insure a time efficient correlation, the software can profit from the machine’s Graphics Processing Units (GPUs) for performant parallel computing. Socet GXP, by using ASM offers point cloud generation possibilities. Point clouds are useful as they provide non-regular spaced point information whereas grids are interpolated terrain heights to fit a regular user-defined grid.

In Semi-Global Matching (SGM) techniques, first, the image’s radiometry is calibrated and corrected, then the SGM performs scanline optimization along the horizontal scanline for each pixel at multiple angles. Each pixel is attributed a cost and a dispersion value and is in a matrix and then filtered by linear programming (HIRSCHMÜLLER 2005; HIRSCHMÜLLER 2011). This method provides the user with non-smoothed DSM with a high point density, available in ASPRS LAS format.

swisstopo has tested several DSM extraction softwares to replace SocetSet NGATE 5.6.0. Two of them, XPro-DSM 6.3.1 by Leica Geosystems and Socet GXP 4.1.0 by BAE Systems have provided satisfying first results. Table 3 presents the advantages and disadvantages of the selected softwares. Both softwares are interesting as they can be foreseen for different strategies. In the case of GLAMOS, most of the AOIs are covered by single flight lines, which can be efficiently processed using DSM-XPro. However, some of the AOI have to be covered by several flight lines, which makes Socet GXP interesting, since it is possible to calculate a DSM with multiple flight lines, as for swissALTI^{3D}. Moreover, Socet GXP offers efficient and user friendly DSM merging possibilities using the Figures Of Merit (FOM) information, which is not available with XPro-DSM. FOM gives a quantitative evaluation of the correlation which can also be used to choose the best source DSM for the end product (GINZLER & HOBI 2015).

Tab. 3: Advantages and disadvantages of SocetSet 5.6.0 by BAE Systems, of Socet GXP 4.1.0 by BAE Systems and of XPro-DSM 6.3.1 by Leica Geosystems

Software	Available correlation methods	Chosen correlation method	Advantages	Disadvantages
SocetSet 5.6.0 BAE Systems	ATE, NGATE	NGATE	<ul style="list-style-type: none"> • Figures Of Merit (FOM) • Processes simultaneously several overlapping lines • Merging handled • Seed DTM option • Automated processing possible • Interpolation for the no data area 	<ul style="list-style-type: none"> • No las output • Does not work with oriented L0 • No support available since replaced by Socet GXP • Works with L1 which must be previously oriented and processed
Socet GXP 4.1.0 BAE Systems	ATE, NGATE, ASM	ASM	<ul style="list-style-type: none"> • Figures Of Merit (FOM) • Processes several lines • Merging handled • Compatible with SocetSet NGATE files • Seed DTM option • Automated processing possible • Interpolation for the no data area 	<ul style="list-style-type: none"> • No las output • Does not work with oriented L0 • Works with L1 which must be previously oriented and processed
XPro-DSM 6.3.1 Leica Geosystems	SGM	SGM	<ul style="list-style-type: none"> • Las output • Seed DTM option • Automated processing possible • Works with L0 • Good results on both bright (glaciers) and dark (shadows) areas 	<ul style="list-style-type: none"> • Processes flights lines individually • Merging required • No FOM • No interpolation for the no data area

4 Results

Socet GXP 4.1.0 by BAE Systems and XPro-DSM 6.3.1 by Leica Geosystems are actually undergoing a series of tests to determine which one is the most accurate and cost-efficient, as well as which parameters and strategies should be used to satisfy GLAMOS' and swissALTI^{3D}'s needs. The tests are to be carried out until august 2016. In this paper, we present our first results, which up to this date have been essentially executed with XPro-DSM 6.3.1. Further tests and results with Socet GXP 4.1.0 will be available during summer 2016. All the tests are done on machines with similar characteristics, so that the performances for each calculation can be compared.

Hereafter are presented the results of tests using different viewing angle combinations, seed DTMs and comparisons of diverse image resolutions and extraction strategies (see Table 4).

Tab. 4: Methods and parameters tested on XPro-DSM 6.3.1

Software	Input image	Viewing Angle combination	Correlation method	Parameters
XPro-DSM 6.3.1	ADS80, L0, GSD 25cm	Green N0°- Green B16°	SGM	<ul style="list-style-type: none"> • No terrain option, normal exposure settings • No terrain option, bright settings • No terrain option, dark settings • Terrain option (swissALTI3D – 2m grid), normal exposure settings • Terrain option (GTOPO – 1 km grid), normal exposure settings
		Green N0°- Green B16°- PAN F27°		
		NIR N0° - NIR B16°		
		NIR N0° - NIR B16°- PAN F27°		
		PAN F27°- PAN F02°- PAN B14°		

4.1 Viewing Angle combination effect

Correlations executed with two viewing angle combination are overall 2.3 times faster as correlations with three viewing angle combination, with a mean speed of 9.55 km²/h for the two viewing angle combination and 4.0 km²/h for the three viewing angle combination. All the calculations were done using a machine equipped with 16 cores with a CPU frequency of 2.59 GHz.

The DSM coverage ranges from 92.5% to 96.2% for all the viewing angle combination except for the PAN F27°- PAN F02°- PAN B14° which ranges from 78.9% to 89.5%.

The results with the combinations NIR N0° - NIR B16°, NIR N0° - NIR B16°- PAN F27° and PAN F27°- PAN F02°- PAN B14° present more noise and artefacts and are thus qualitatively less accurate than the Green N0°- Green B16° and Green N0°- Green B16°- PAN F27° combinations.

4.2 Terrain option

In XPro-DSM 6.3.1, the quality of a DSM can be improved by using the “terrain option”, as shown below (Fig.4). When not using the “terrain option”, the software correlates a DSM based on a given mean altitude. However, in mountainous areas, the difference between the mean altitude and the real altitude can exceed 1’000 m. The “terrain option” allows the user to provide the software with a seed DTM with a more or less big grid spacing which is integrated during the DSM calculations. In addition to local DTM, global models are also available, such as GTOPO30 and SRTM3, both produced by the U.S. Geological Survey (USGS). GTOPO30 is a worldwide model derived from data issued by several sources, with a horizontal grid spacing of approximately 1’000 meters. SRTM3 is the result of a campaign called Shuttle Radar Topography Mission Global Coverage providing a global model from 56 degrees south latitude to 60 degrees north latitude, with a grid spacing of approximately 90 meters.

The results of a correlation using swissALTI^{3D} with a 2 meter grid spacing as seed DTM were compared to those obtained using GTOPO30 which has a 1'000 meter grid. Other tests are planned using SRTM3.

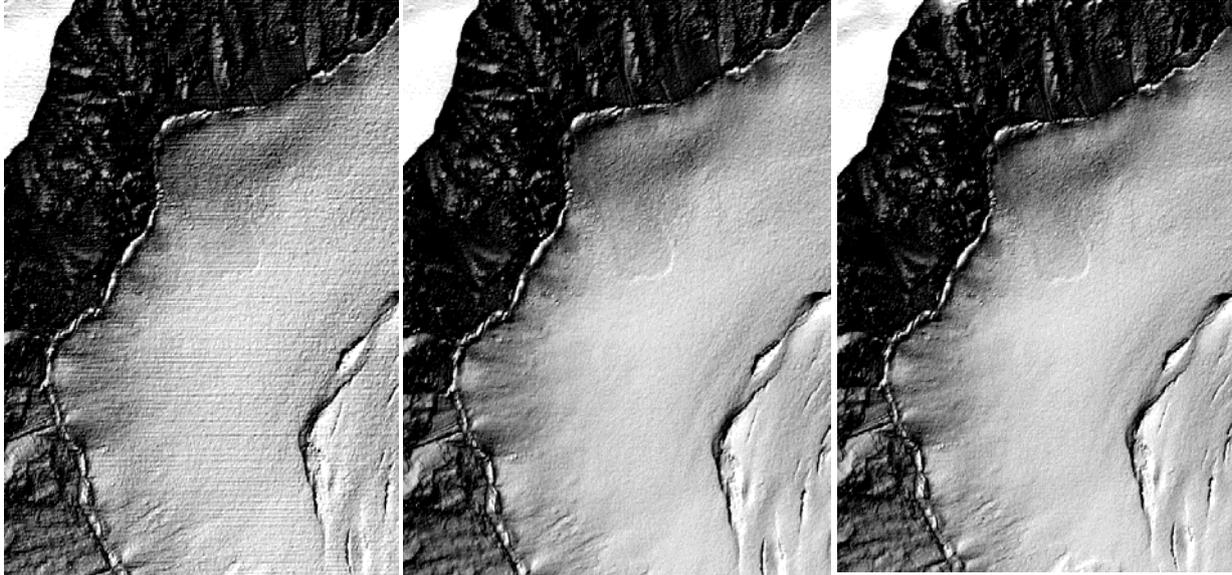


Fig. 4: DSM calculation with and without terrain option. On the left, there is no seed DTM. In the middle, swissALTI^{3D} with a 2m grid was used. On the right, GTOPO30 was used as seed DTM

Small wave artefacts with a 10 cm amplitude affecting the DSM with no terrain option are lessened by using a seed DTM, even with a 1'000 meter grid. The terrain option improves the quality, but does not necessarily provide a better accuracy. However, this quality improvement comes with very important time costs. Calculations time using swissALTI^{3D} with a 2 m grid as seed DTM are 30 times greater as calculations without the terrain option. Calculations with GTOPO30 as seed DTM are 7 times greater as without seed DTM. Qualitative results using swissALTI^{3D} 2m are comparable to those using GTOPO30. As the DSM quality definitely can be enhanced by the use of the terrain option, further investigations are currently been carried out using SRTM3 and swissALTI^{3D} with a bigger grid.

4.3 Normal, bright, dark extraction parameters

XPro-DSM comes with a set of strategies for specific scenes. For example, when set on “bright setting”, the software focuses on the bright pixels of the image, such as glaciers, providing a better quality on the bright areas. When set on “dark setting”, the software focuses on the dark pixels of the image and offers a better correlation in shadow areas. Results show that the coverage percentages using a normal, a bright or a dark extraction parameter are comparable (96.5%). The dark setting clearly does improve the interpolation quality in shadow areas, without deteriorating the quality over bright areas such as glaciers. Fig. 5 shows that void areas of the DSM calculated with the “bright” setting (circled in red) were successfully completed by the DSM calculated with the “dark” setting. Moreover, the correlation in shadow areas is improved by using the dark setting (green circle).

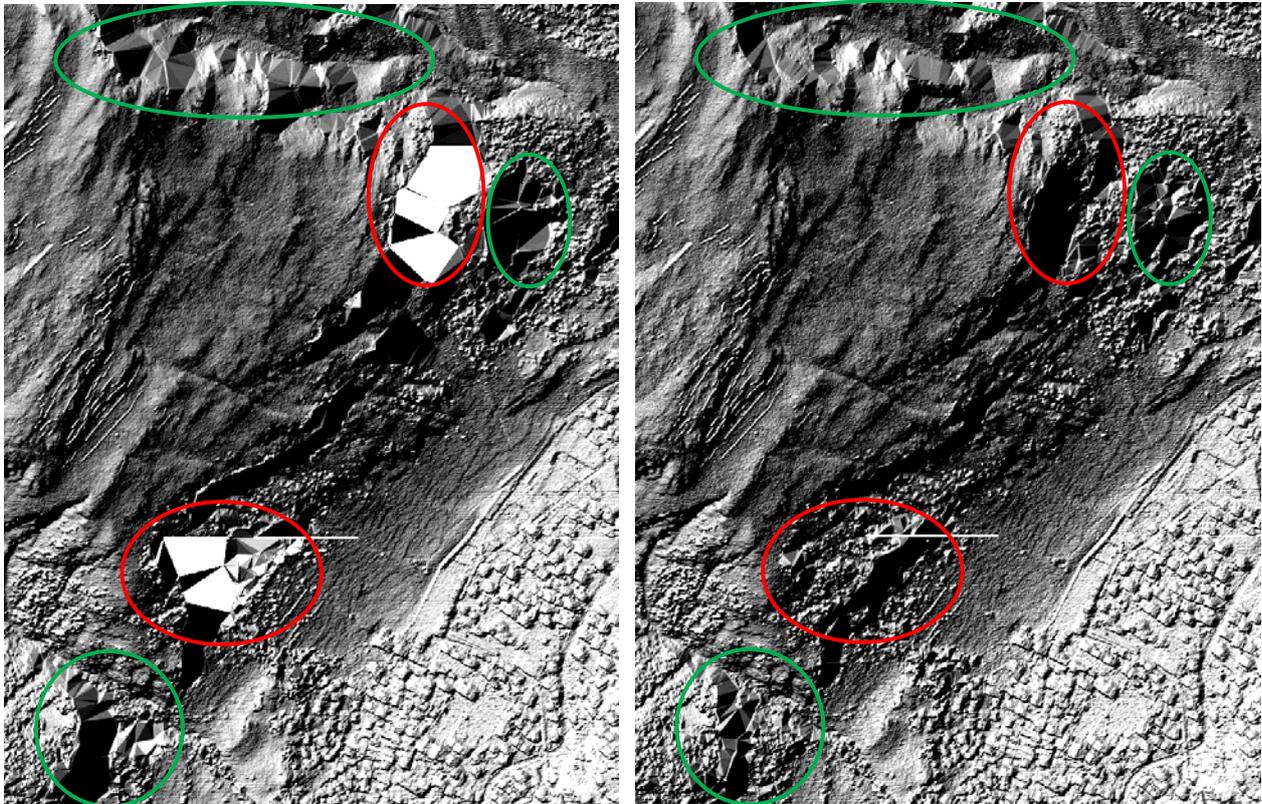


Fig. 5: DSM calculation with a bright setting on the left and with a dark setting on the right. A clear improvement can be observed as the dark settings allows void areas calculated with the bright setting to be filled (red circles). Moreover, the quality of the correlation in other shadow areas is improved by using the dark settings (green circles)

Processing time using a normal exposure setting, a bright or a dark setting are equivalent (Fig. 6).

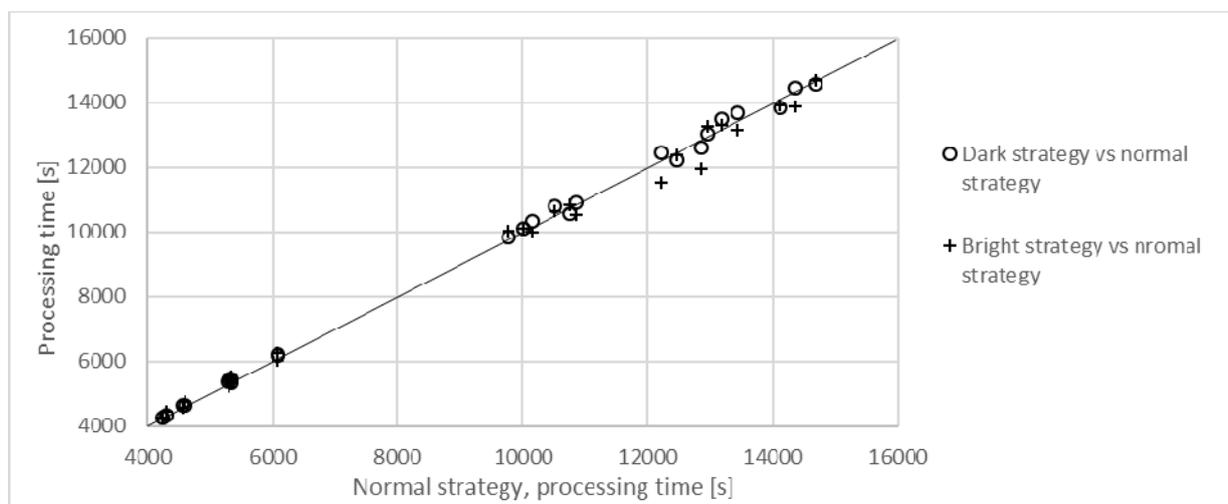


Fig. 6: Processing time [s] using normal, bright and dark settings

4.3.1 Resolution

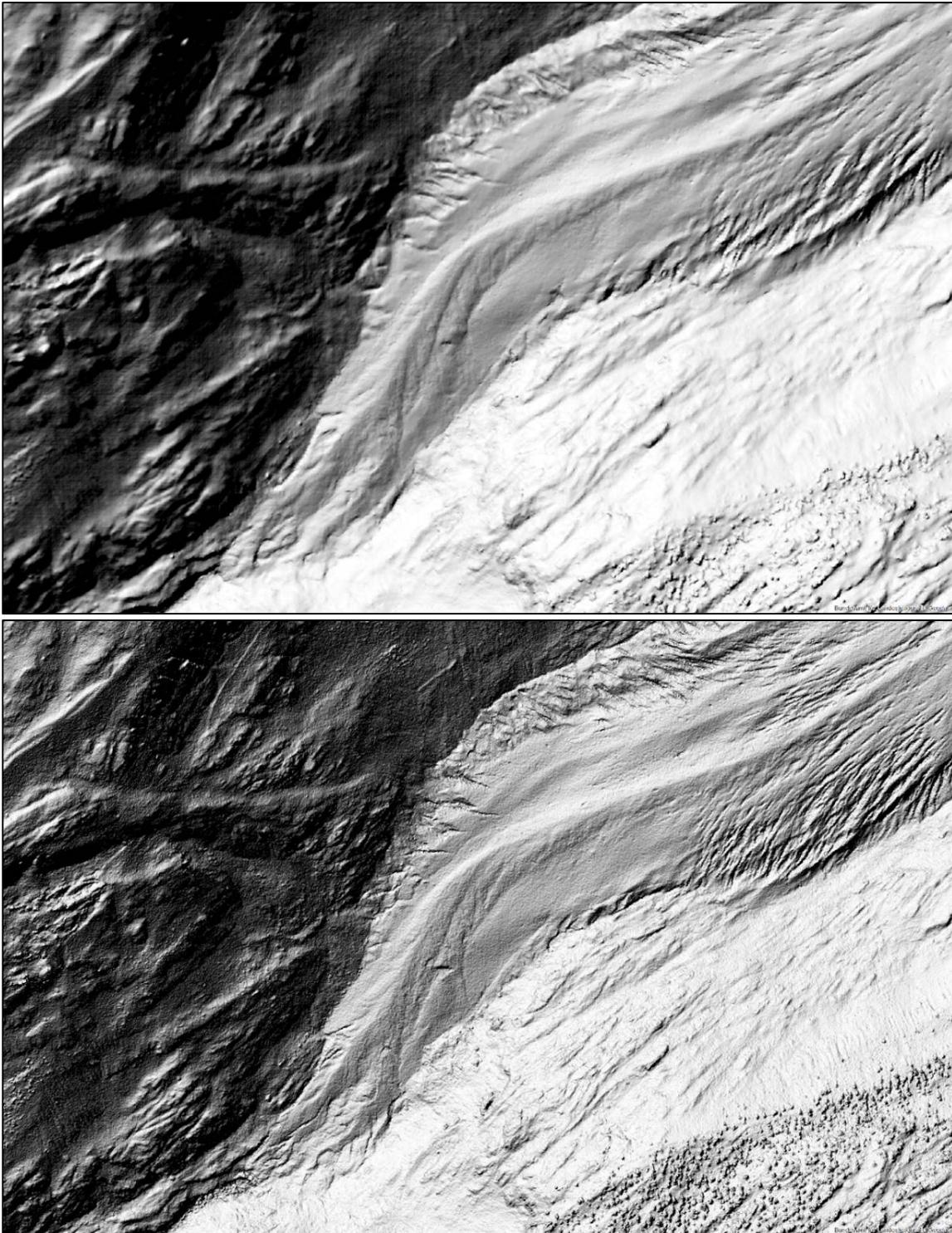


Fig. 7: The upper picture shows the DSM of the Aletsch ice tongue correlated with a 1 m grid spacing using SocetSet NGATE 5.6.0. The bottom picture shows the las output on the same area using XPro-DSM 6.3.1

XPro-DSM produces two sets of data, one with non-thinned xyz values and another one with thinned xyz values. Only the non-filtered xyz DSMs are considered. Fig. 7 shows two DSMs correlated over the Aletsch ice tongue. The data was collected during two flights, one in 2014 and the other in 2015, both with the ADS80 and with a 25 cm GSD. The upper picture shows a DSM extracted with SocetSet NGATE 5.6.0 with a 1 meter grid spacing with data from the 2014 flight. The bottom picture shows a DSM extracted with XPro-DSM 6.3.1.

Obviously, a higher resolution results provides more details especially for smaller objects. Glacier edges, structures and crevasses are more defined and detailed. Trees are easily identifiable from houses whereas they could in some cases be mistaken for one another using SocetSet NGATE 5.6.0 with a 1 meter grid. Other terrain features are more detailed, such as boulders, faults or meandering channels of rivers. The overall quality is enhanced, although the accuracy is equally enhanced since the point cloud gives a xyz value for a given point whereas a grid is an interpolation of the altitude onto a given xy point.

The height difference between correlations carried out with SocetSet NGATE 5.6.0 and XPro-DSM 6.3.1 with the same input images range from -0.2 to 0.2 meters.

5 Discussion

Due to clearly poor results, the three PAN combination is excluded for further production. Although better, DSMs calculated with NIR N0° - NIR B16°, NIR N0° - NIR B16°- PAN F27° still present noise and artefacts. The two most accurate DSMs are obtained by using the Green N0°- Green B16° or Green N0°- Green B16°- PAN F27° viewing angle combination. Although the Green N0°- Green B16° combination is 2.4 times faster than the Green N0°- Green B16°- PAN F27° combination, the accuracy and overall visual quality of the latter make it really interesting for further production. Less areas are void, especially in mountainous areas.

Using the terrain option improves the visual quality, even by using a 1'000 meters grid GTOPO30 seed DTM. However, the terrain option does not enhance the precision of the correlation in comparison to the "no terrain option". The process is moreover very time consuming. The use of a seed DTM could be recommended for mountainous areas because of their topographical disparities, but is not necessary in other areas. The use of the terrain option can therefore not be recommended as a systematical process, but could be locally used on specific, problematic mountainous areas.

The dark setting definitely enhances the quality of the DSM in shadow areas without deteriorating the quality on bright areas. Moreover, the time costs for this process can be neglected. The dark setting can be integrated as a fix parameter.

Finally, a smaller DSM resolution results in a more detailed DSM and accuracy in comparison to the actual SocetSet NGATE 1 meter production strategy, since the output is a point cloud and not a grid.

6 Conclusion

The replacement of the actual SocetSet NGATE 5.6.0 software is a necessity. The tests carried out up to this date have been mostly focused on DSM-XPro 6.3.1, further results with SocetSet GXP 4.1.0 will be available during summer 2016.

An enhanced resolution comes along with logical longer processing time.

Another concern is the merging of the data. Since XPro-DSM 6.3.1 does not offer merging possibilities of several DSMs, this process must be manually done, which is not time and cost-efficient. Therefore, the future tests with Socet GXP 4.1.0 might bring satisfying automated merging opportunities which are not available with XPro-DSM 6.3.1.

A way of optimizing the production would be to eventually use both DSM-XPro 6.3.1 for local areas and glaciers entirely covered by a single flight line. Socet GXP 4.1.0 could be used for merging multiple flights lines as well as for change detection. However, using two methods is not the most efficient strategy for a large scale production and involves undesirable further infrastructure and licenses costs.

7 References

- BÜHLER, Y., GINZLER, CH. & MARTY, M., 2012: High Resolution DEM Generation in High-Alpine Terrain Using Airborne Remote Sensing Techniques. *Transactions in GIS* **16** (5), 635-647.
- BÜHLER, Y., GINZLER, CH., MARTY, M. & BEITINGER, J., 2015: Snow depth mapping in high alpine catchments using digital photogrammetry. *The Cryosphere* **9**, 229-243.
- DEVENEZIA, K.; WALKER, S. & ZHANG B., 2007: New approaches to generating and processing high resolution elevation data with imagery. *Photogrammetric Week* **7**, 297-308.
- FRANK, P., BAUDER, A., MARTY, M. & NÖTZLI, J., 2011: Neige, glaciers et pergélisol 2009/10, Les Alpes, juillet 2007, 47-52.
- GINZLER, CH., 2015: Countrywide Stereo-Image Matching for Updating Digital Surface Models in the Framework of the Swiss National Forest Inventory. *Remote Sensing* **7**, 4343-4370.
- GINZLER, CH. & HOBI, M., 2015: Erstellung eines landesweiten digitalen Oberflächenmodelles aus operationellen Stereo-Luftbildern. *Geomatik Schweiz* **9**, 332-335.
- HAUCK, CH., COLLET, C., DELALOYE, R., HILBICH, CH., HOELZLE, M., HUSS, M. & SALZMANN, N., 2012: The potential of new measurement and modelling techniques in alpine cryosphere and geomorphology. *Geographica Helvetica* **67**, 26-37.
- HIRSCHMÜLLER, H., 2005: Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* **2**, 807-814.
- HIRSCHMÜLLER, H., 2011: Semi-Global Matching – Motivation, Developments and Applications. *Photogrammetric Week* **11**, 173-184.
- HOBI, M. & GINZLER, CH., 2012: Accuracy assessment of Digital Surface Models Based on WorldView-2 and ADS80 Stereo Remote Sensing Data. *Sensors* **12** (5), 6347-6368.
- WAGNER, R., 2011: Leica ADS80 and Leica XPro: A total Solution for Photogrammetric Mapping. *Photogrammetric Week* **11**, 79-87.
- ZHANG, B., 2014: Automatic Spatial Modeler (ASM): Elevation by innovation. BAE Systems, Geospatial eXploitation Products, 1-26.

Ersatz des SRTM Oberflächenmodells durch AW3D30?

KARSTEN JACOBSEN¹

Zusammenfassung: Das digitale Oberflächenmodell (DOM) der Shuttle Radar Topography Mission (SRTM), das auf Aufnahmen im Jahr 2000 basiert und seit September 2014, mit Ausnahme des Nahen Ostens, mit einem Punktabstand von einer Bogensekunde (arcsec) (~30m am Äquator) frei zur Verfügung steht, hat sich als Standardhöhenmodell für viele Anwendungen etabliert. Basierend auf allen verfügbaren ALOS PRISM tri-stereo Bildern, die vom Januar 2006 bis zum Mai 2011 aufgenommen wurden, hat die japanische Weltraumorganisation JAXA für den Bereich von 82° südlicher bis 82° nördlicher Breite ein DSM unter dem Namen ALOS World 3D (AW3D) erstellt. Das AW3D DSM wird kommerziell mit einem Punktabstand von 5m vertrieben. Hiervon wird mit einem Punktabstand von 1 Bogensekunde als AW3D30 eine Version erstellt, die bereits für große Teile der Erde kostenlos verfügbar ist. Unter anderem ist damit der Nahe Osten mit einer Auflösung von 1 arcsec abgedeckt.

Ein Vergleich von AW3D30 DOM mit ausreichend genauen Referenzhöhenmodellen ergab höhere Genauigkeiten als mit dem SRTM DOM. In gebirgigen Gebieten ist der Vorteil der AW3D30 DOM wegen der Probleme des synthetischen Apertur Radar mit der Signalüberlagerung (lay over) noch größer. Die Vorteile optischer Bilder im Gebirge sind allerdings mit Nachteilen in Gebieten mit extremer Wolkenbedeckung verbunden. In tropischen Gebieten können die Kriterien von AW3D30 gegenüber SRTM anders ausfallen.

1 Einleitung

Die japanische Weltraumbehörde JAXA hat die Erstellung des kommerziellen AW3D Oberflächenmodells aus allen verfügbaren ALOS PRISM Aufnahmen für den Bereich zwischen 82° südlicher Breite bis 82° nördlicher Breite fast abgeschlossen (TADONO et al. 2014; TAKAKU et al. 2014) und erstellt zurzeit die kostenlose Version AW3D30 mit einem Punktabstand von 1 arcsec, entsprechend etwa 30m Abstand am Äquator (Abb. 1). Probleme mit Wolkenbedeckungen deuten sich in Abbildung 1 an den schwarzen Streifen in den bereits erstellten Gebieten an. Die Höhenmodelle werden zusammen mit Qualitätsinformationen geliefert, die die Anzahl der für jeden einzelnen Objektpunkt benutzten Bilder angeben sowie die Standardabweichung gegenüber den Profilverpunkten von ICESAT und den nahezu weltweiten DOM von SRTM und ASTER GDEM2 (JACOBSEN 2014).

Die optischen ALOS PRISM Aufnahmen haben eine Objektpixelgröße von 2,5 m, das Aufnahmesystem hatte eine vorausblickende und eine zurückblickende Kamera mit jeweils 24° Nadirwinkel, sowie eine senkrecht schauende Kamera. Mit den geneigten Aufnahmen wird ein Basis-Höhen-Verhältnis von 1:1 erreicht. Die Aufnahmen haben nicht alle eine optimale Bildqualität, was sich jedoch nicht auf die Erstellung der Höhenmodelle ausgewirkt hat.

JAXA gibt für AW3D eine Standardabweichung der Lage und Höhe von 5 m an was in den bislang untersuchten Gebieten deutlich unterschritten wurde.

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

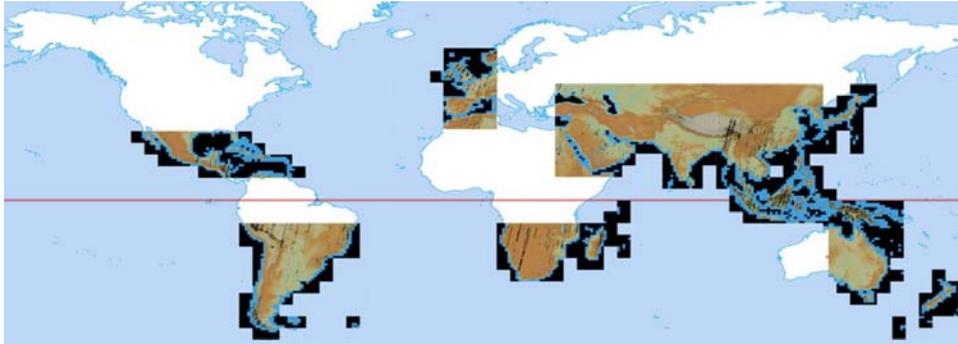


Abb. 1:
Verfügbarkeit von
AW3D30 DOM im
Januar 2016

© JAXA

Es wurden für drei Testgebiete die Standardabweichungen, die Normalized Median Absolute Deviation (NMAD) (HÖHLE & HÖHLE 2009) sowohl absolut, als auch als Funktion der Geländeneigung und für Gebiete mit Geländeneigungen kleiner als 10% ermittelt. Es zeigte sich für AW3D30 eine Abhängigkeit der Höhengenaugigkeit von der Anzahl der Bilder/Punkt.

Zum Vergleich mit den AW3D30 Oberflächenmodellen wurden ebenfalls SRTM-DOM und teilweise ASTER GDEM-2 DOM in den gleichen Gebieten untersucht. Wegen der Abhängigkeit der Höhengenaugigkeit vom Charakter des Geländes ist ein Vergleich nur im gleichen Gebiet möglich. Seit September 2014 sind, mit Ausnahme des Nahen Ostens, die SRTM DOM ebenfalls mit einem Punktabstand von einer Bogensekunde frei verfügbar (Abb. 2).

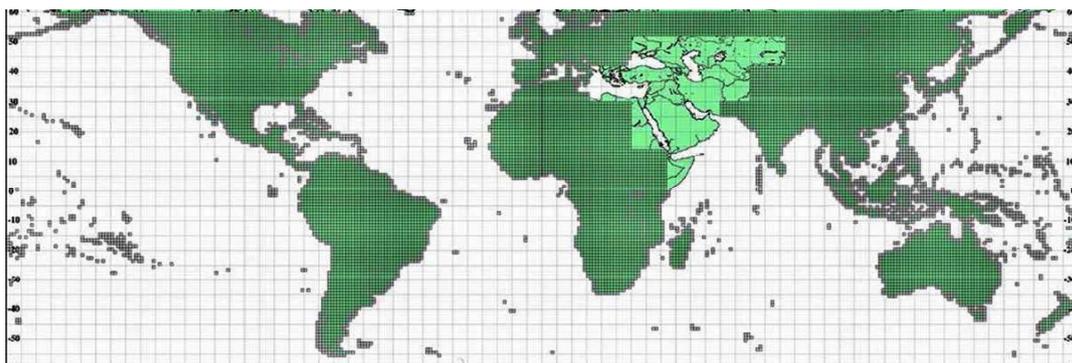


Abb. 2: dunkelgrün = am 22.9.2014 September freigegebene SRTM DOM mit 1 arcsec Punktabstand, hellgrün = Beschränkung auf 3 arcsec Punktabstand (USGS)

2 Untersuchte Datensätze

AW3D30 Oberflächenmodelle wurden in drei Testgebieten untersucht in denen ausreichend genaue Referenzhöhenmodelle vorhanden sind und die auch einen Vergleich mit dem SRTM DOM ermöglichen. Im Benchmark-Testgebiet der ISPRS für Ziyuan 3 in Sainte-Maxime, Frankreich, (<http://www.isprs.org/data/zy-3/data/referencedata/Default.aspx>) steht ein Referenz-DOM der IGN, erstellt aus hochauflösenden Luftbildaufnahmen, zur Verfügung. Das Gebiet weist ebene Flächen auf, ist aber dominierend gebirgig. Das Testgebiet der Bülent Ecevit Universität, Zonguldak, hat einen gebirgigen Charakter mit hoher Rauigkeit. Hier stehen zwei Referenzdatensätze, erstellt aus hochauflösenden Luftbildaufnahmen zur Verfügung. Das SRTM C-Band DOM hat hier zwar nur einen Punktabstand von 3 Bogensekunden, dafür liegen hier aber die etwa gleich genauen SRTM X-Band-Daten mit einer Bogensekunde Punktabstand vor.

Das Testgebiet Jordanien erstreckt sich vom Toten Meer bis nach Amman und ist durch Bergland ohne Bewaldung dominiert.

Zusammen mit den Höheninformationen wird eine Datei mit der Anzahl der Bilder je Objektpunkt und Kopfinformationen über die Höhenverschiebung und Standardabweichung von AW3D30 gegenüber SRTM, ASTER GDEM 2 und ICESAT, sowie die Häufigkeitsverteilung der ASTER PRISM Bildkorrelation geliefert. ICESAT war ein von 2003 bis 2010 aktiver Satelliten-Laserprofilierer, der eine Objektbestrahlungsfläche von 70m Durchmesser hatte und in ebenem Gelände eine absolute Höhengenaugkeit von 0,2 m erreichte. ICESAT-Profildaten wurden zur Höhenverbesserung der AW3D30 DOM und der SRTM DOM herangezogen.

Die Häufigkeitsverteilung der ALOS PRISM Bildkorrelationskoeffizienten (Abb. 3) ist für die Testgebiete unterschiedlich, den größten Prozentsatz hoher Korrelationskoeffizienten gibt es im nicht bewaldeten Testgebiet Jordanien, während die Korrelationskoeffizienten im Testgebiet Zonguldak am kleinsten sind. Ähnlich sieht es mit dem Quadratmittel der Höhenabweichungen von AW3D30 gegenüber SRTM, ASTER GDEM 2 und ICESAT aus, hier ist die Übereinstimmung im Testgebiet Jordanien, das kaum Bewuchs aufweist, am besten, während das Testgebiet Zonguldak die größten Abweichungen hat (Abb. 4). In den Testgebieten Jordanien und Saintes-Maxime zeigen die ICESAT-Daten die kleinsten Widersprüche gegenüber AW3D30. Im Gebiet Zonguldak gibt es offensichtlich zu wenig flache Gebiete, die gute ICESAT-Geländehöhen zulassen. Die größten Unterschiede hat AW3D30 gegenüber ASTER GDEM 2, das auch gegenüber Referenzhöhenmodellen größere Unterschiede als SRTM aufweist.

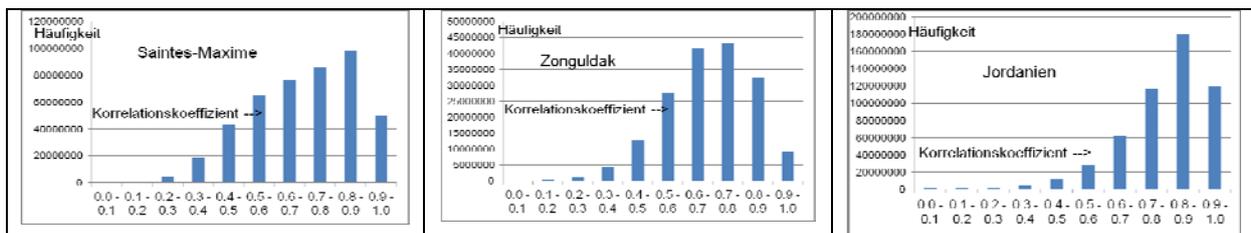


Abb. 3: Häufigkeitsverteilung der Korrelationskoeffizienten von ALOS PRISM (Informationen in den AW3D30-Datenlieferungen enthalten)

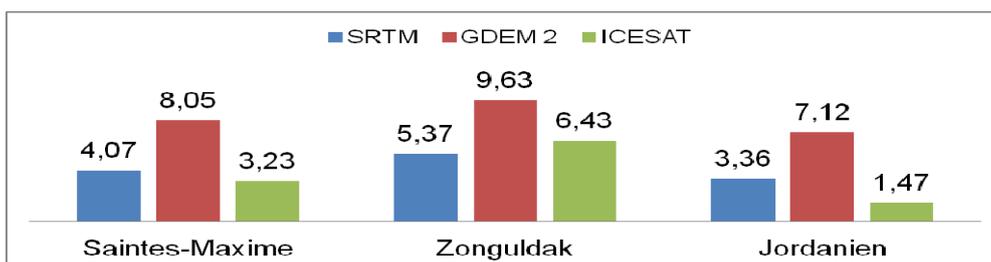


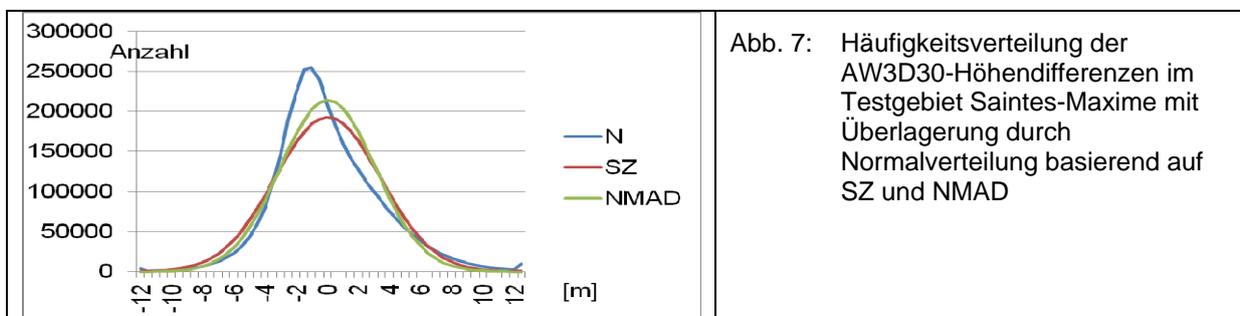
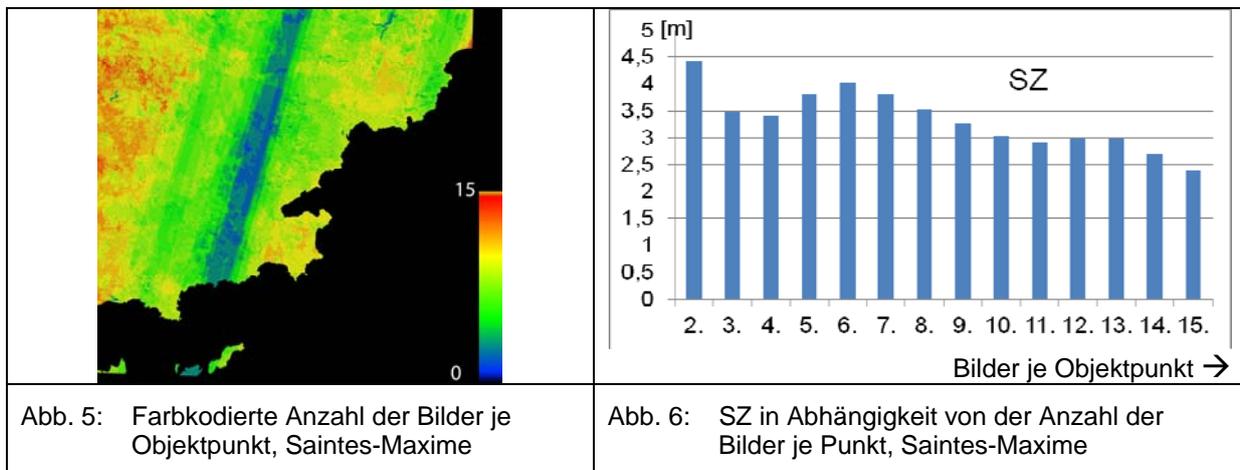
Abb. 4: Quadratmittel der Höhenabweichungen AW3D30 gegen SRTM, ASTER GDEM 2 und ICESAT [m] - entsprechend der AW3D30 Qualitätsinformation

3 Saintes-Maxime

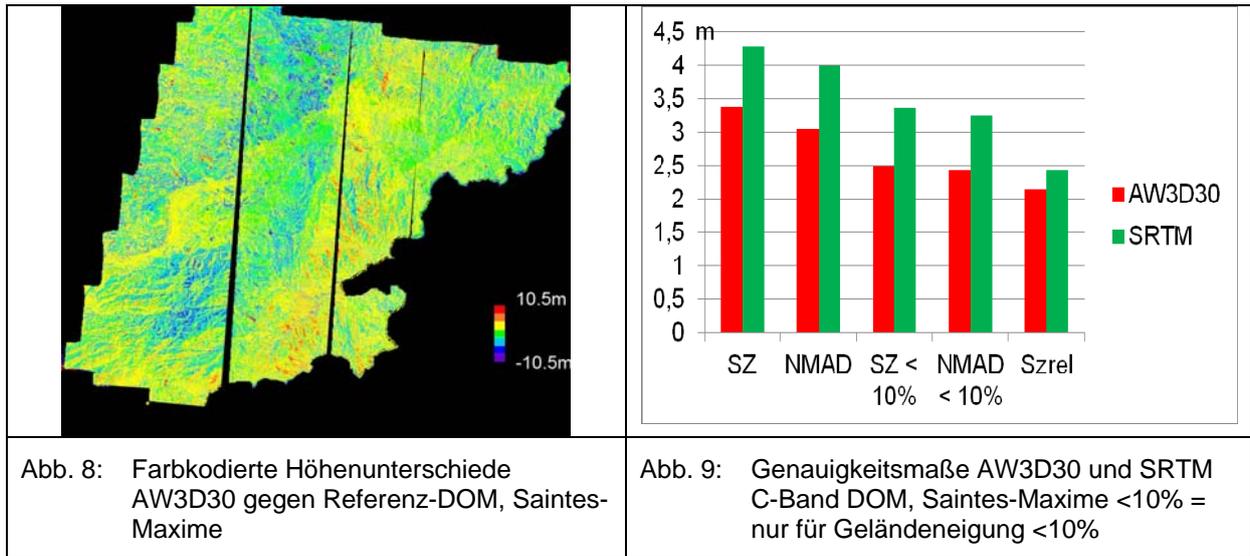
Entsprechend des Referenzhöhenmodells beträgt im Testgebiet Saintes-Maximes die mittlere Geländeneigung benachbarter Punkte 25% und das Quadratmittel der Rauigkeit über 20 m

(linear interpoliert und mit der Höhe in der Mitte verglichen) 2,62 m. Diese Parameter wurden aus dem Referenzhöhenmodell ermittelt. Die von dem Referenzhöhenmodell erfasste Fläche ist allerdings nicht mit dem Untersuchungsgebiet identisch, das durch eine Maske eingeschränkt wurde.

Wie auch in den anderen Testgebieten haben die untersuchten Höhenmodelle Neigungen von bis zu etwa 3m über den Bereich der Höhenmodelle. Das wirkt sich aber auf die Genauigkeitsmaße nur geringfügig aus (JACOBSEN 2016). Wegen Wolkenbedeckung variiert die Anzahl der Bilder je Objektpunkt erheblich. Abb. 5 zeigt die Anzahl der Bilder je Objektpunkt für das Testgebiet Saintes-Maxime, sie geht von 2 bis zu 15 Bildern (Abb. 6), mit im Mittel wurden 7,4 Bilder je Objektpunkt. Als ausgleichende Funktion ergibt sich die Standardabweichung der Höhe (SZ) zu $SZ=4.29 \text{ m} - 0.14 \text{ m} \times \text{Anzahl der Bilder}$.



Die Häufigkeitsverteilung der AW3D30 DOM Höhendifferenzen gegenüber dem Referenz-DOM (Abb. 7) zeigt eine Kurtosis von 1,03 und eine Schiefe von 0,50 obwohl ein Oberflächenmodell mit einem anderen verglichen wird. In den anderen Untersuchungsgebieten ist Kurtosis größer, während die Schiefe kleiner ausfällt. Die farbkodierten Höhendifferenzen (Abb. 8) deuten allerdings an, dass sich die Vegetation während der Aufnahmezeitpunkte verändert hat. Wie in den meisten Fällen stimmt die Normalverteilung basierend auf NMAD besser mit der tatsächlichen Häufigkeitsverteilung überein als die Normalverteilung basierend auf der Standardabweichung – NMAD ist somit eine bessere Beschreibung der tatsächlichen Normalverteilung (Abb. 7).



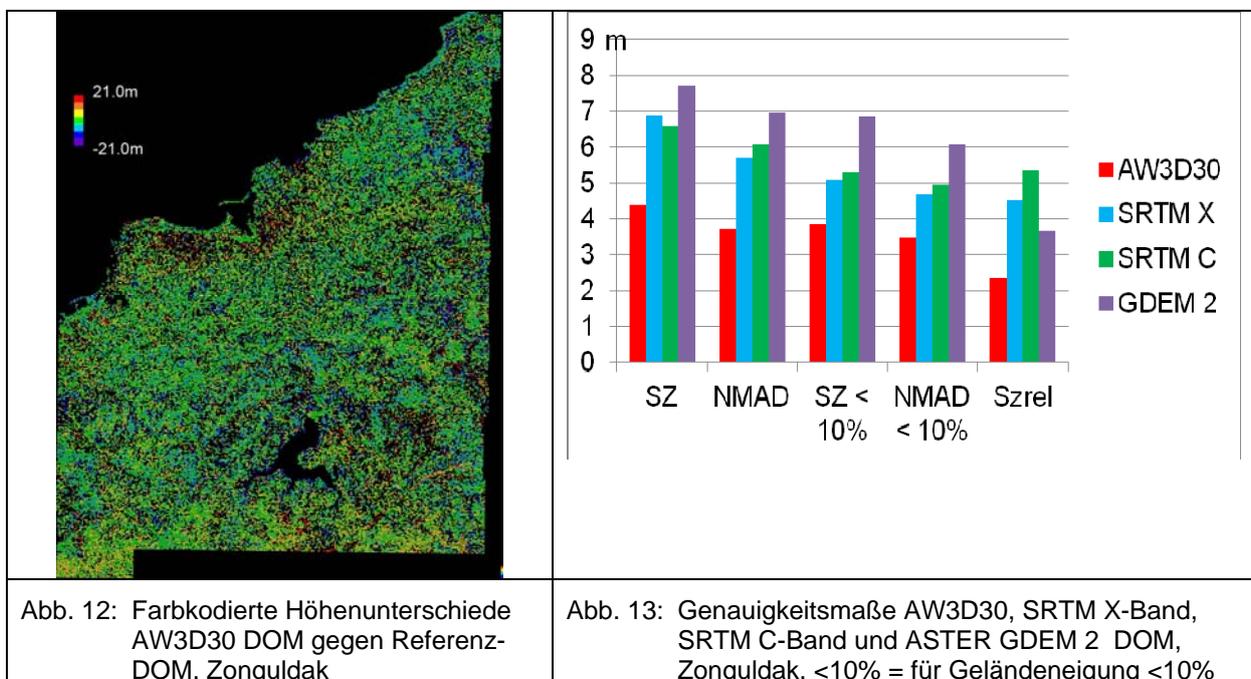
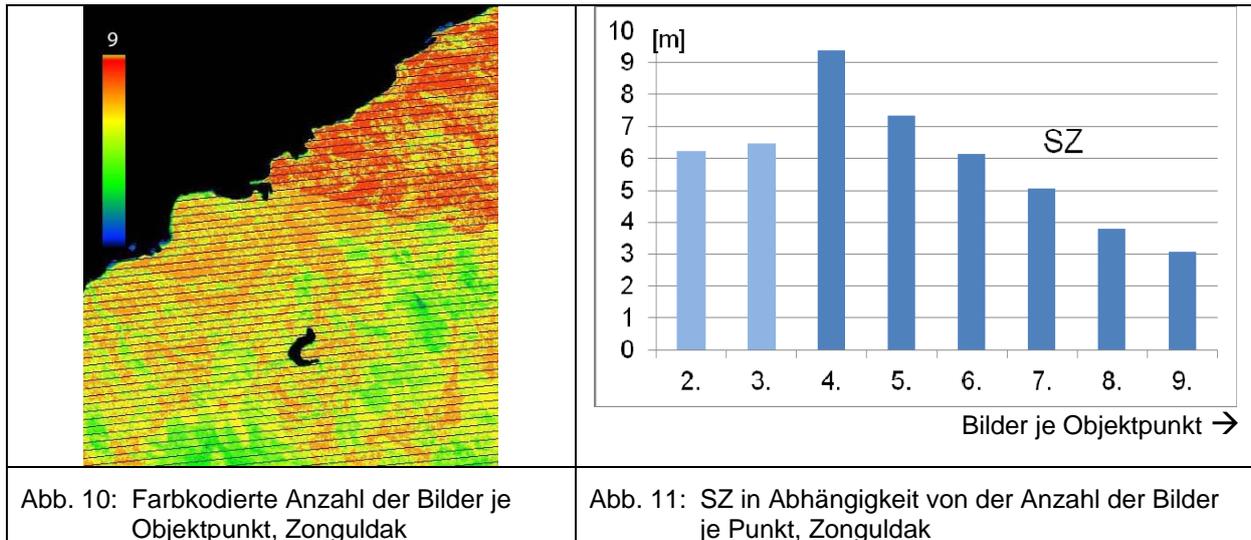
Tab. 1: Genauigkeitsmaße als Funktion der Geländeneigung, α = Geländeneigung

	SZ	NMAD
AW3D30	$SZ = 3.15 \text{ m} + 1.70 \text{ m} \times \tan \alpha$	$NMAD = 2.75 \text{ m} + 1.94 \text{ m} \times \tan \alpha$
SRTM C-Band	$SZ = 4.04 \text{ m} + 1.63 \text{ m} \times \tan \alpha$	$NMAD = 3.60 \text{ m} + 1.99 \text{ m} \times \tan \alpha$

Abbildung 9 zeigt die Genauigkeitsmaße des AW3D30 und des SRTM DOM. Wie üblich ist die Höhengenaugigkeit von der Geländeneigung abhängig (Tabelle 1), weswegen die Genauigkeit ebenfalls für die Bereiche mit einer Geländeneigung bis zu 0,1 (10%) dargestellt ist. Die Genauigkeitsabhängigkeit als Funktion der Geländeneigung kann nicht direkt mit der Genauigkeit für Bereiche mit Geländeneigungen < 10% verglichen werden, da die ausgleichende Funktion von der Verteilung der Genauigkeiten für alle Neigungsgruppen abhängt. Der neigungsabhängige Term sagt aber etwas über die Geländecharakteristik und die Bestimmungsmethode der Höhenmodelle aus. So ist das interferometrische syntaktische Radar von SRTM stark von nahezu senkrechten Geländeteilen abhängig. AW3D30 zeigt eine deutlich höhere Genauigkeit als SRTM. Die Relativgenauigkeit direkt benachbarter Punkte $\text{SQRT}((\sum(DZ_i - DZ_j)^2)/2n)$ gibt die Genauigkeit unabhängig von Modelldeformationen, Orientierungseinflüssen und auch großflächigen Veränderungen der Vegetation an. Sie stellt die Grenzgenauigkeit der Oberflächenmodelle dar.

4 Zonguldak

Das gebirgige Gebiet Zonguldak (Küste des Schwarzen Meeres nördlich von Ankara) enthält teilweise städtische Bereiche und offene Flächen, hauptsächlich aber leicht bewaldete Gebiete. Die mittlere Geländeneigung benachbarter Punkte beträgt 26% und das Quadratmittel der Rauigkeit über 24 m (linear interpoliert und mit der Höhe in der Mitte verglichen) 0.89 m, wobei zu berücksichtigen ist, dass die Rauigkeit in der Regel quadratisch von der Interpolationsweite abhängig ist. Eine Rauigkeit von 0.89 m über 24 m Distanz entspricht deswegen einer Rauigkeit von 0,62 m über 20 m Distanz.



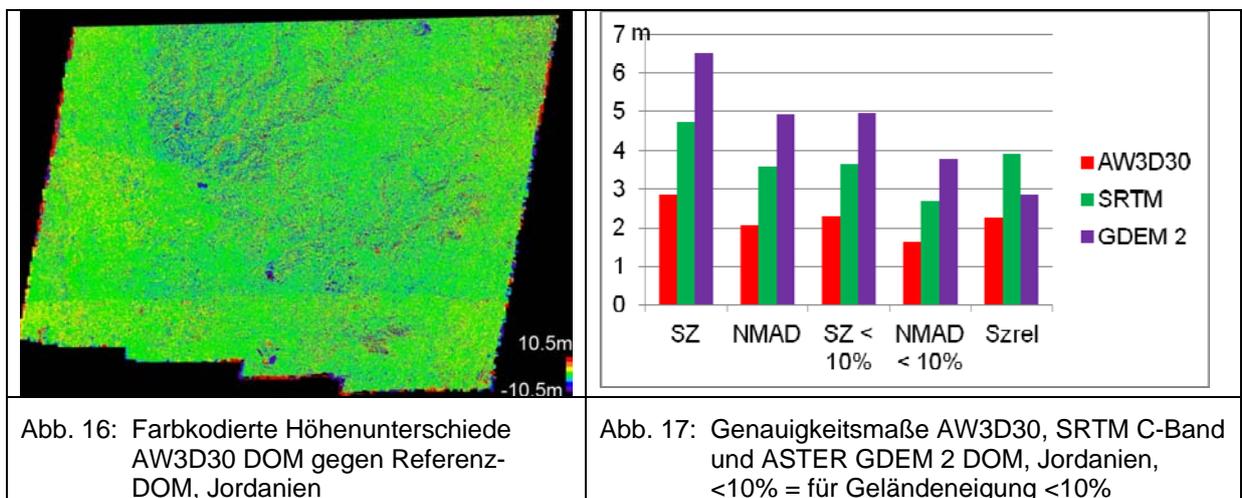
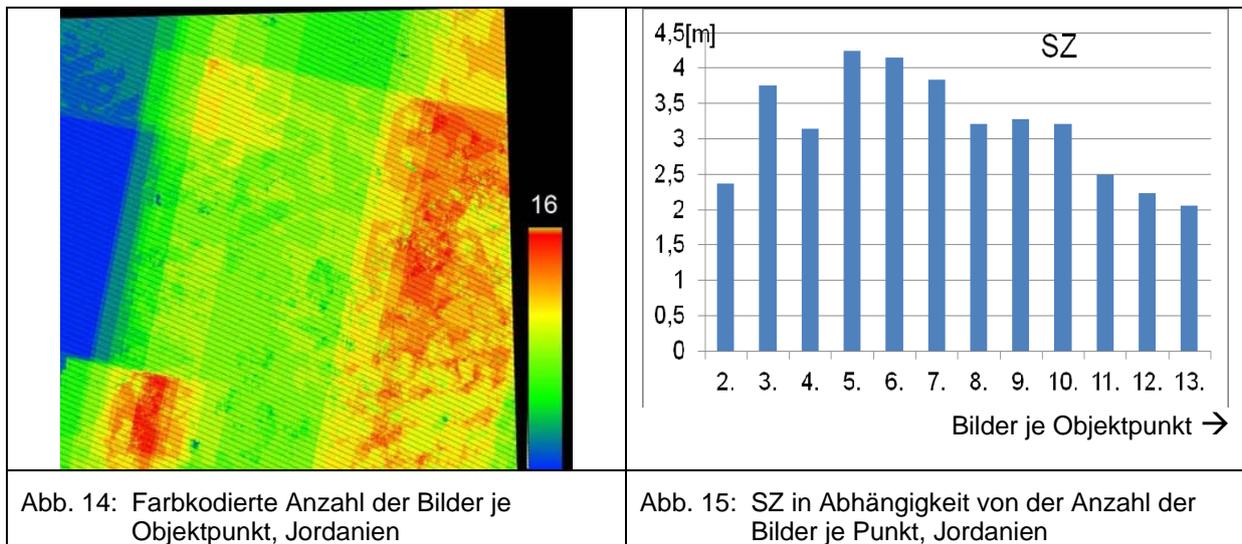
Im Testgebiet Zonguldak wurden mit im Durchschnitt 6.75 Bilder je Objektpunkt etwas weniger Bilder für die Erstellung von AW3D30 verwendet (Abb. 10) als in Saintes-Maxime. Die Abhängigkeit der Standardabweichung von der Anzahl der Bilder je Punkt (Abb. 11) zeigt eine Unstetigkeit für 2 bzw. 3 Bilder je Punkt, hier sind allerdings nur 21 bzw. 32 Höhenpunkte gegenüber im Durchschnitt mehr als 10 000 Punkte für die anderen Gruppen vorhanden, dementsprechend ist die gewichtete lineare Abhängigkeit: $SZ = 9,20 \text{ m} - 0,90 \text{ m} \times \text{Anzahl der Bilder je Punkt}$. Da die steileren Geländeabschnitte teilweise nahezu senkrechte Wände enthalten ist die Abhängigkeit von der Geländeneigung bei AW3D30 bei etwa $5,0 \times \tan \alpha$, während sie bei SRTM in der Größenordnung von $12,0 \times \tan \alpha$ liegt. Dieses erklärt teilweise die schlechteren Ergebnisse des Testgebiets Zonguldak.

Die Bereiche der Stadt Zonguldak (Mitte oben in Abb. 12) fallen durch größere Höhenabweichungen auf. Generell ist im Testgebiet Zonguldak wegen der hohen Geländerauigkeit und des teilweisen Einflusses von Bewaldung und Bebauung die Standardabweichung der Höhenmodelle (Abb. 13) größer als in den beiden anderen Testgebieten. Die hier vorliegenden SRTM C-Band und X-Band DOM unterscheiden sich nur geringfügig, wogegen das ASTER GDEM 2 DOM, mit Ausnahme der Relativgenauigkeit, deutlich schlechter ist. AW3D30 zeigt erheblich bessere Genauigkeiten als die anderen DOM.

5 Jordanien

Für das Testgebiet Jordanien beläuft sich die mittlere Geländeneigung benachbarter Punkte auf 19% und das Quadratmittel der Rauigkeit über 40 m auf 3,23 m, entsprechend einer Rauigkeit von 0,81 m über 20 m Distanz.

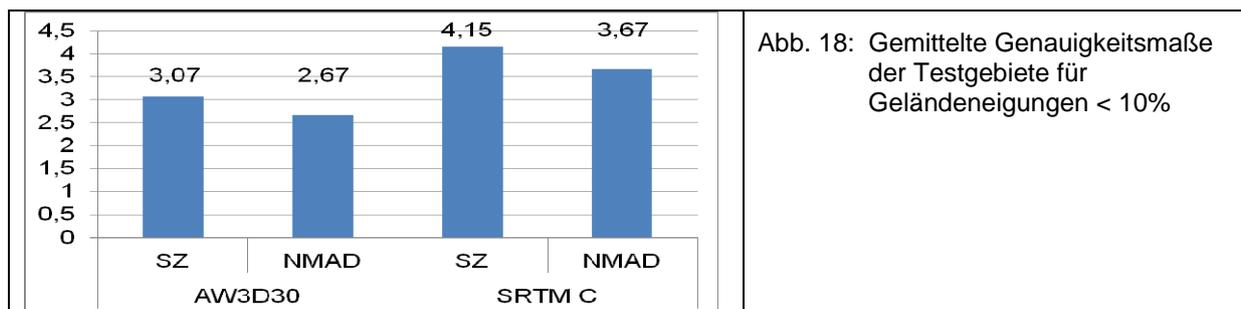
Mit im Durchschnitt 5,32 Bildern je Objektpunkt wurden weniger Bilder verwendet als in Saintes-Maxime und in Zonguldak. Als ausgeglichene Funktion ergibt sich: $SZ = 4,08 \text{ m} - 0,18 \text{ m} \times \text{Anzahl der Bilder je Punkt}$.



Die Genauigkeitsabhängigkeit von der Geländeneigung ist etwa so wie im Gebiet Saintes-Maxime. Im Testgebiet Jordanien bestätigen sich die Relationen der oben dargestellten Untersuchungen, AW3D30 zeigt eine deutlich bessere Genauigkeit als das SRTM DOM und ASTER GDEM 2 zeigt die größten Genauigkeitsmaße. Wie in Zonguldak ist allerdings die relative Genauigkeit von ASTER GDEM 2 besser als die für SRTM.

6 Zusammenfassung und Schlussfolgerung

Die Größen der Genauigkeitsmaße der drei untersuchten Testgebiete unterscheiden sich, wegen unterschiedlicher Vegetation, Geländeneigung, Bebauung und Geländecharakteristik, wie es auch in anderen Testgebieten der Fall ist (JACOBSSEN 2014). Bereits die Qualitätsinformationen von AW3D30 (Abb. 4) deuten auf schlechtere Ergebnisse für das Testgebiet Zonguldak hin. Dagegen stimmen die Ergebnisse der Gebiete Saintes-Maxime und Jordanien relativ gut überein. Sowohl der Mittelwert der Standardabweichungen als auch der von NMAD ist für das SRTM DOM um etwa 35% größer als für AW3D30 (Abb. 18). Die Relationen der Genauigkeitsmaße stimmen in den drei Gebieten überein.



Wie sich bereits früher zeigte (JACOBSSEN 2014) und auch der der Fall ist, erreicht das ASTER GDEM 2 nicht die Genauigkeit des SRTM DOM, wie es die in der Datenlieferung enthaltene Qualitätsinformation von AW3D30 auch angibt. Die Relativgenauigkeit und damit die morphologische Qualität, ist allerdings besser als bei SRTM, aber nicht so gut wie bei AW3D30. Auf der Grundlage dieser Untersuchungen kann das AW3D30 als Ersatz für das SRTM DOM empfohlen werden. Die untersuchten Testgebiete decken Problembereiche der Höhenmodelle, wie durch Vegetation und Gebirge, gut ab. Abbildung 1 zeigt aber auch Gebiete in denen AW3D30 wegen Wolkenbedeckung Lücken aufweist. Andererseits werden die Bereiche von 62° bis 82° nördlicher Breite erfasst, für die keine SRTM-Höhendaten vorliegen. Soweit AW3D30 für das zu bearbeitende Gebiet vorliegt, empfiehlt es sich, AW3D30 dem SRTM DOM vorzuziehen.

7 Literaturverzeichnis

- HÖHLE, J. & HÖHLE, M., 2009: Accuracy Assessment of Digital Elevation Models by means of Robust Statistical Methods. *ISPRS Journal of Photogrammetry and Remote Sensing* **64**, 398-406.
- JACOBSEN, K., 2014: Performance of large Area Covering Height Models. Proceeding of the 34th EARSeL Symposium, Warsaw.
- JACOBSEN, K., 2016: Analysis and Correction of Systematic Height Model Errors. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41** (I-B1).
- 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*, **2** (4), 71-76.
- TAKAKU, J., TADONO, T. & TSUTSUI, K., 2014: Generation of High Resolution Global DSM from ALOS PRISM. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (4), 243-248.

Ein adaptives und kompaktes Fernerkundungssystem für UL-Fluggeräte – Konzept und Anwendungen

CASPAR KNEER¹, ALEXANDER JENAL¹, IMMANUEL WEBER¹ & JENS BONGARTZ¹

Zusammenfassung: Systeme für die luftgestützte Fernerkundung werden häufig durch aufwendige Um- und Einbauten in bemannte Fluggeräte integriert und binden dieses damit an einen speziellen Anwendungszweck. Das hier vorgestellte modulare Fernerkundungssystem wurde speziell für die temporäre Verwendung in Ultraleichtflugzeugen (UL) entwickelt. In einem ersten Entwicklungsschritt (FlugKit) wurde ein Aufbau für das weit verbreitete Tragschraubermodell AutoGyro MTOsport entworfen. Alle Komponenten finden dabei in einem robusten und kompakten Aufbau Platz. Dieser besteht aus einem Trägersystem, einer Sensoraußenlast und Anzeigeinstrumenten für den Piloten. Durch eine reversible Haltevorrichtung kann das System binnen kürzester Zeit ein- und ausgebaut werden. Dabei sind keine dauerhaften Veränderungen am Fluggerät notwendig. Durch Miniaturisierung und Integration aller Systemkomponenten in einen Messgerätekofter (FlugCase) konnten weitere neue Möglichkeiten bezüglich Transportfähigkeit und Verwendung in anderen UL Modellen eröffnet werden.

1 Einleitung und Motivation

In früheren Arbeiten (JENAL et al. 2015; WEBER et al. 2015) zeigten die Autoren bereits die vorteilhaften Eigenschaften von Ultraleichtfluggeräten (UL) für die luftgestützte Fernerkundung am Beispiel von Tragschraubern. Dazu zählen im Vergleich zu größeren Flugzeugen und Hubschraubern die niedrigen Unterhalts- und Betriebskosten sowie die einfachere Abnahme von Einbauten und Modifikationen. Typspezifisch bieten Tragschrauber gute Langsamflugeigenschaften sowie hohe Wendigkeit und Stabilität. Im Vergleich zu UAVs verfügen ULs über deutlich größere Traglasten und Flugzeiten und dürfen sich als reguläre Teilnehmer frei im Luftraum bewegen.

Diese Aspekte aufgreifend, stellen die Autoren in dieser Arbeit den nächsten Schritt in der Entwicklung von Fernerkundungssystemen für UL-Flugzeuge vor. Dieser basiert auf zwei zusammenhängenden Erkenntnissen. Zum ersten ist das Verlegen eines Flugzeugs für eine Messkampagne mit einem erheblichen logistischen Aufwand verbunden. Zum anderen bindet die Modifikation und Ausrüstung eines einzelnen Flugzeugs, sowohl das Flugzeug an diesen Bestimmungszweck als auch den Betreiber an das jeweilige Flugzeug. Die Autoren adressieren den zweiten Punkt durch die Entwicklung eines Fernerkundungssystems, das ohne Modifikationen am Fluggerät in kurzer Zeit ein- und ausgebaut werden kann. Auf diese Weise bleibt das Fluggerät weiterhin flexibel einsetzbar. Damit bindet es den Betreiber nicht mehr an ein einzelnes Fluggerät, wodurch auch der erste Punkt in Angriff genommen werden kann: Durch die Wahl eines weit verbreiteten Fluggeräte-Modells als Basis für das modulare Fernerkundungssystem erübrigt sich in den meisten Fällen ein aufwändiges Verlegen des Flugzeugs. Das Fluggerät kann meistens in

¹ Fraunhofer-Institut für Hochfrequenzphysik und Radartechnik FHR, Anwendungszentrum für multimodale und luftgestützte Sensorik AMLS, Joseph-Rovan-Allee 2, D-53424 Remagen
E-Mail: [caspar.kneer, alexander.jenal, immanuel.weber, jens.bongartz]@fhr.fraunhofer.de

der Nähe des Befliegungsareals gechartert werden und die Ausrüstung wird beispielsweise per PKW zu einem Flugplatz in der Nähe transportiert. Der Ein- und Ausbau erfolgt vor Ort und die Befliegung wird durch einen ortsansässigen, erfahrenen Piloten durchgeführt.

Im Folgenden wird das Konzept und Design der Fernerkundungsausrüstung vorgestellt und abschließend anhand von zwei Beispielkampagnen die neuen Möglichkeiten des Systems erläutert.

2 Konzept und Design

Im folgenden Kapitel wird der prinzipielle Aufbau des Systems und die beiden Varianten - FlugKit und FlugCase - für die Verwendung in zwei unterschiedlichen UL-Fluggeräten (Abb. 2) beschrieben.

2.1 Bestandteile des Gesamtsystems

Das Grundelement der beiden Varianten ist ein mechanischer Träger, der für die Dauer der Befliegung mit dem Fluggerät verbunden wird. Dieser ist für das jeweilige Fluggeräte-Modell ausgelegt und die beiden bisher entwickelten Varianten werden in den folgenden Abschnitten näher vorgestellt. Hervorzuheben ist, dass für die Montage keine Änderungen am Fluggerät erforderlich sind. Der Träger nimmt sämtliche Komponenten auf und sorgt für die notwendige Stabilität des Systems. Ein Grundelement ist die integrierte, unabhängige Stromversorgung, die eine Einsatzdauer des Systems von ca. acht Stunden gewährleistet und einen Betrieb am Boden ohne laufenden Motor ermöglicht. Im Inneren des Fluggerätes befinden sich im Träger neben der Stromversorgung der Datenverarbeitungsrechner, der die Sensoren ansteuert, die anfallenden Daten verarbeitet und speichert, sowie weitere Steuersysteme kontrolliert. Die Sensoren werden als Außenlast montiert und in einem aktiven, elektromechanischen Gimbal betrieben, das auch bei unruhigerer Wetterlage und den damit verbundenen Kurskorrekturen des Piloten eine Nadir-Orientierung der Sensoren gewährleistet. Einen schematischen Überblick der Systemkomponenten zeigt Abb. 1. Das Gimbal ist mittels Schwingungsdämpfer mit dem Träger verbunden und wird dadurch von den Vibrationen des Motors und des Rotors entkoppelt. Ein Windabweiser vor dem Außenlast-Gimbal reduziert die Windlast auf dessen Motoren und schützt die Sensoren vor Verschmutzung und Beschädigung (siehe auch Abb. 3).

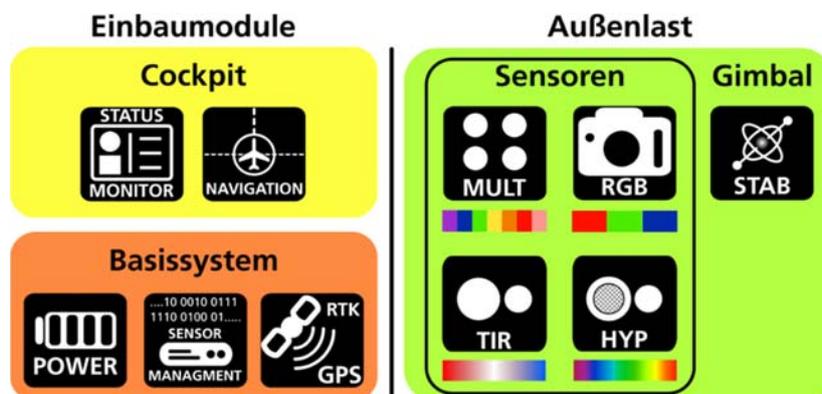


Abb. 1: Schema der Systemkomponenten: Im Cockpit befinden sich das Navigations- und Statusdisplay für den Piloten. Das Basissystem ist im Trägersystem integriert, an dem die Außenlast mit den Sensoren montiert ist.



Abb. 2: *Links:* Tragschrauber AutoGyro MTOsport mit integriertem FlugKit. *Rechts:* Trike P&M Aviation QuikR mit integriertem FlugCase

Neben dem Träger und seinen Komponenten werden im Cockpit ein Navigations- sowie ein Statusdisplay montiert. Letzteres informiert den Piloten über den Zustand des Systems. Das Navigationsdisplay, welches die grafischen Informationen des Navigationssystems darstellt, führt den Piloten intuitiv zum Befliegungsareal und dort zu den ortsgebundenen Mess- bzw. Aufnahmepunkten. Diese Punkte werden vorab per Software geplant und vor Abflug in das Navigationssystem eingespielt. Es sind sowohl Pfadbefliegungen als auch Flächenbefliegungen möglich.

Das gesamte System ist so ausgelegt, dass der Pilot sich auf den Luftraum und den Flug konzentrieren kann und bis auf die Navigation keine Berührungspunkte mit dem System hat. Vor dem Flug wird das System durch einen Operator gestartet, konfiguriert und die Sensoren werden kalibriert. Während des Fluges werden die Sensoren beim Erreichen der Aufnahmepunkte automatisch ausgelöst. Die anfallenden Messdaten der Kameras und der Umweltsensoren, dazu zählen u.a. Luftdruck, Lufttemperatur, Luftfeuchte und das Spektrum des einfallenden Sonnenlichtes, werden zusammengeführt und in Kombination mit Zeit- und Positionsdaten des GNS-Systems gespeichert.

2.2 FlugKit

Das Fernerkundungssystem FlugKit wurde speziell für Tragschrauber des Typs AutoGyro MTOsport entwickelt. Mit über 1500 weltweit zugelassenen Modellen handelt es sich hierbei um eines der meist verwendeten Ultraleichtfluggeräte. Das FlugKit wird über eine Haltevorrichtung reversibel mit dem Hauptrahmen des Tragschraubers verbunden. Dabei nimmt das System den Fußraum des Rücksitzes ein, sodass die Mitnahme einer zweiten Person nicht mehr möglich ist. Dies wiederum ermöglicht eine maximale Nutzlast von bis zu 100 kg. Das in Abb. 3 dargestellte System mit einseitiger Außenlast wiegt zurzeit 32 kg, was eine Flugzeit von bis zu fünf Stunden ermöglicht.



Abb. 3: FlugKit für AutoGyro MTOsport. Zu sehen ist auf dem linken Bild das eingebaute Trägersystem mit seinen zwei Containern für die Energieversorgung (links) und die Computersteuerung sowie GPS (rechts). Das rechte Bild zeigt eine Detailaufnahme der Außenlasthalterung für die Sensoren mit elektromechanischer Lageregelung.

Je nach Kamerasystem, Flughöhe und erforderlicher Bodenauflösung kann das System eine Fläche von bis zu 40 km² pro Stunde erfassen. Auf Grund der modularen Bauform kann das System innerhalb von einer halben Stunde eingebaut und in Betrieb genommen werden.

2.3 FlugCase

Die zweite Ausbaustufe des Systems, die sich im Wesentlichen durch die Integration der zuvor beschriebenen Komponenten in einen Messgerätekofter auszeichnet, wurde speziell für eine Messkampagne in Kambodscha entwickelt. Der in Abb. 4 dargestellte Koffer beinhaltet alle Basiskomponenten des Systems (Energieversorgung, Computer und GPS) und kann direkt auf dem Sitz eines Fluggerätes befestigt werden. Im konkreten Beispiel ist dies der Rücksitz eines Trikes des Typs QuikR von P&M Aviation.

Hinzu kommt ein kompakter, demontierbarer Aluminiumträger, der mit dem Rahmen der Sitzschalen-Aufnahme des Trikes verbunden ist. An diesem wird die Sensorik inklusive Gimbal befestigt. Das komplette für die Durchführung der Messkampagne benötigte Equipment, inklusive Werkzeug, konnte auf drei Koffer verteilt im Handgepäck zu einer Messkampagne nach Kambodscha transportiert werden (siehe Abs. 3.2).



Abb. 4: *Links:* Ein Trike des Modells QuikR von P&M Aviation mit eingebautem FlugCase und Halterahmen für das Kamerasystem. Im Cockpit befindet sich das Navigationsdisplay. *Rechts:* Das FlugCase ist ein umgebauter robuster Kunststoff-Koffer. Seitlich sind die verschiedenen Systemanschlüsse zu erkennen.

3 Anwendungen

Im Folgenden werden zwei Messkampagnen vorgestellt, die mit den zwei erwähnten Fluggerätetypen durchgeführt wurden.

3.1 Thermale Befliegung der Hahnhöfer Nebelnelbe

Zusammen mit der Bundesanstalt für Gewässerkunde (BfG) wurde im Sommer 2015 eine zweitägige Messkampagne zur flächigen Bestimmung der Oberflächentemperatur der Hahnhöfer Nebelnelbe und des Mühlenberger Lochs bei Hamburg, einer 24 km² großen Watt- und Uferfläche, durchgeführt. Zum Einsatz kam hierbei ein AutoGyro MTOsport, der mit einem FlugKit Trägersystem ausgestattet war. Zur Validierung vorhandener Temperatur-Modelle sollte das Untersuchungsgebiet an zwei aufeinanderfolgenden Tagen mehrfach mit einer Thermalkamera befliegen werden. Dies hatte den Zweck, die Temperaturentwicklung während eines Tidezyklus zu verschiedenen Zeitpunkten aufzuzeichnen. Für die Kampagne kam erschwerend hinzu, dass sich das Areal in unmittelbarer Nähe zum Airbus-Werksflughafen in Finkenwerder und in der Einflugschneise des Verkehrsflughafens Fuhlsbüttel befindet. Der Tragschrauber musste deswegen mit eingeschaltetem Transponder fliegen und der Pilot stand in permanenten Funkkontakt mit den Fluglotsen der Deutschen Flugsicherung. Die Flughöhe wurde mit 1400 m über Grund so gewählt, dass der Tragschrauber sich oberhalb der Einflugschneise befand. Die Kampagne konnte erfolgreich durchgeführt werden und an zwei Tagen wurden jeweils vier bzw. drei Flüge durchgeführt. Die Daten wurden zu Gesamtkarten prozessiert und befinden sich zurzeit in der themenspezifischen Auswertung durch die BfG (BONGARTZ et al. 2016).

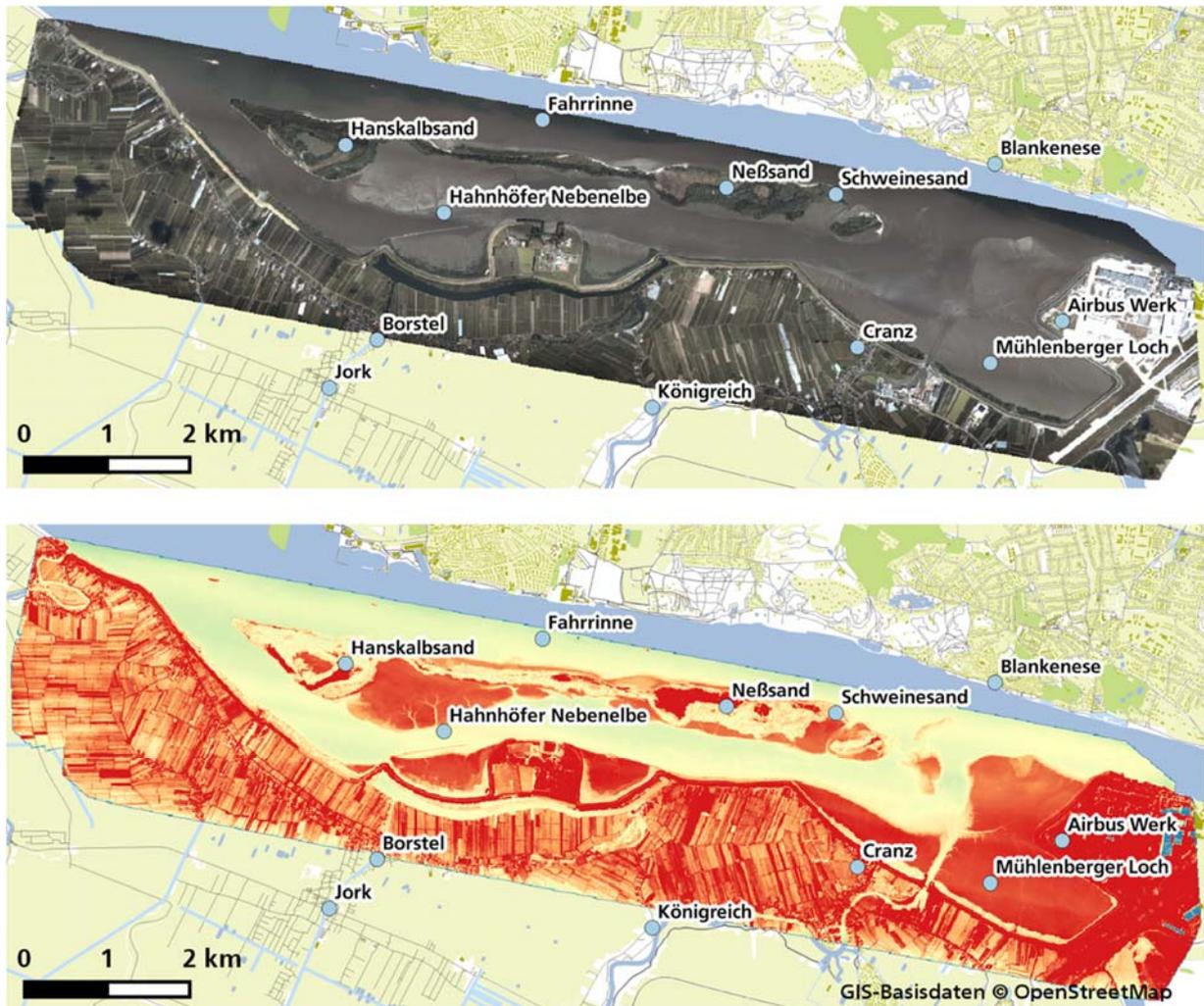


Abb. 5: Bei der Befliegung der Hahnöfer-Nebelbe erstellte Karten. Die obere zeigt ein Farbbild des Befliegungsgebietes und die untere die entsprechende Temperaturkarte. Die Daten stammen vom ersten Flug des Tages bei Niedrigwasser. Das Befliegungsgebiet umfasst eine Fläche von ca. 24 km².

3.2 Evaluierung von Möglichkeiten zur Landminendetektion aus der Luft

In Kambodscha ist die Minenproblematik noch allgegenwärtig. Bedingt durch den starken Vegetationsbewuchs vieler Flächen sowie die Regenzeit ist und bleibt die händische Minensuche und Entschärfung das Mittel der Wahl. Dies bedeutet jedoch auch, dass die Räumung von Flächen nur sehr langsam vorangeht. Da Minen zu einem überwiegenden Anteil wahllos und ohne System von der Landbevölkerung verlegt wurden, gibt es keinerlei belastbare Aufzeichnungen über verminte Flächen. Daher handelt es sich bei den meisten Flächen lediglich um Verdachtsgebiete. Das Projekt ist Teil einer ersten Konzeptstudie, die zeigen soll, ob es mit bildgebender multi- und hyperspektraler Sensorik möglich ist, verminte Flächen anhand von Pflanzensignaturen von minenfreien zu unterscheiden (NAUMANN et al. 2009; ZINNERT 2012). Dadurch soll eine Räumungspriorisierung von belasteten Flächen ermöglicht werden.

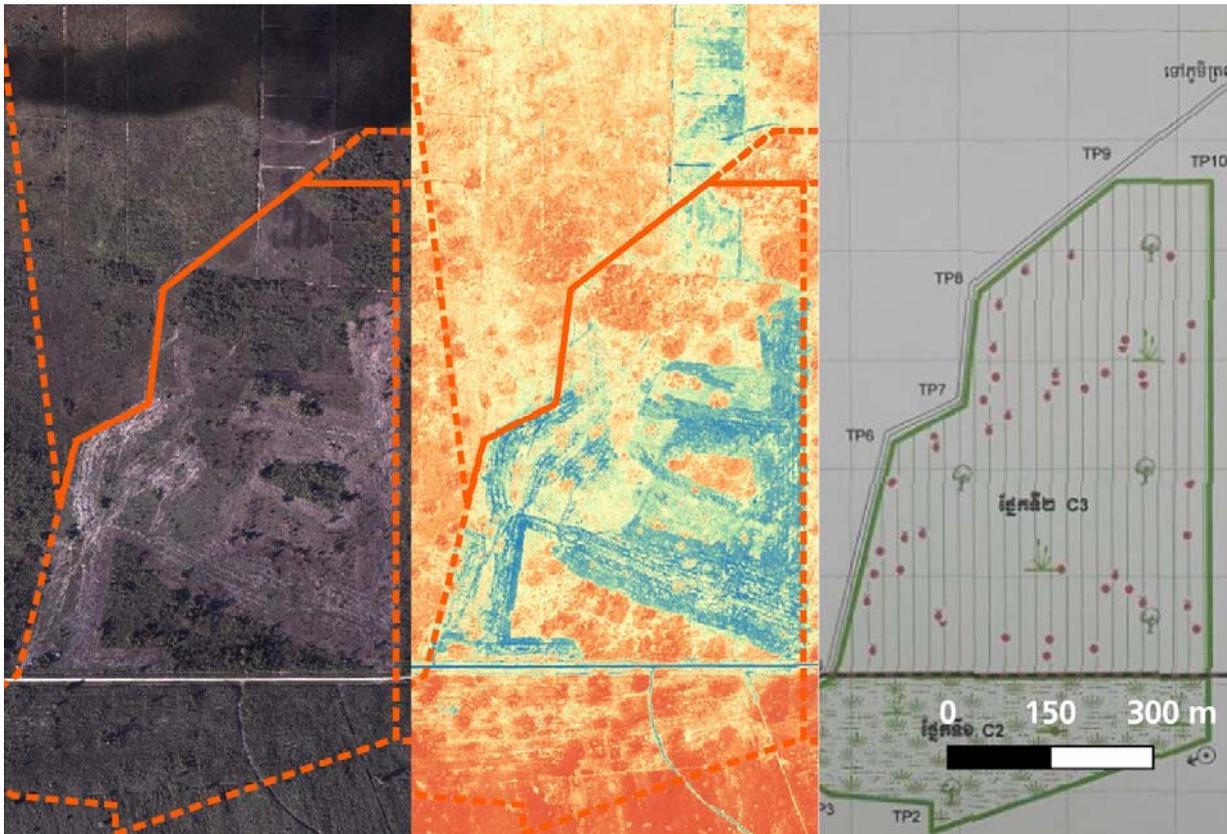


Abb. 6: *Links:* Luftbild einer Minenräumfläche *Mitte:* NDVI dieser Räumfläche *Rechts:* Karte der Minenfundorte

Durchgeführt wurde die Befliegungskampagne im November 2015 im Rahmen eines Fraunhofer-Förderprogramms. Die erfassten Luftbildkarten werden nachträglich mit den tatsächlich gefundenen Minenfundorten verglichen, um Merkmale in der hyperspektralen Signatur zu identifizieren. Neben den anspruchsvollen klimatischen und infrastrukturellen Bedingungen des Landes, stellte die Verfügbarkeit eines geeigneten Flugzeugs eine weitere Hürde dar. Nach intensiven Recherchen konnte ein UL-Trike der Firma P&M Aviation (QuikR) ausfindig gemacht werden. Da der für den MTOsport entwickelte Systemträger nicht mit diesem UL-Fluggerät kompatibel ist, musste das System angepasst werden. Hierbei wurde der in Absatz 2.3 erwähnte Messgerätekooffer (FlugCase) entwickelt. Dieser kann im Gegensatz zum FlugKit aufgrund seiner Abmaße universeller eingesetzt werden und ist nicht nur auf ein Flugzeugmodell beschränkt. Zusätzlich wurde ein Halterahmen für das UL-Trike konstruiert, der die einzelnen Sensorsysteme aufnehmen kann. Das optimierte System konnte per Linienflugzeug in das Zielland gebracht werden. Der Einbau vor Ort erfolgte problemlos, sodass nach einem Flugtag mit ausgiebigen Tests an den beiden darauffolgenden Tagen jeweils zwei Flüge über den ca. 50 km entfernten Minenfeldern durchgeführt werden konnten.

4 Fazit und Ausblick

Die von den Autoren entwickelte Fernerkundungsausrüstung in Form des FlugKits für den Tragschrauber AutoGyro MTOsport und des FlugCase für das Trike P&M Aviation QuikR ermöglicht es, mit niedrigem logistischen, zeitlichen und personellen Aufwand Fernerkundungskampagnen in vielen Teilen der Welt durchführen zu können. Ziel ist es, ein passendes Fluggerät in der Nähe des Zielgebietes zu chartern und für den Einsatzzweck umzurüsten. Die kompakte und leichte Bauform des Equipments ermöglicht z. B. den Transport mit einem PKW. In der Variante für das Trike sogar im Handgepäck während einer Flugreise. Die Entscheidung Ultraleichtflugzeuge zu verwenden, vereinfacht die regulatorischen Bedingungen erheblich und reduziert den finanziellen Aufwand einer Befliegung deutlich.

Die Autoren arbeiten zurzeit an einer zweiten Außenlastaufnahme des Systems für den gleichzeitigen Betrieb von zwei Sensorsystemen. Die Integration von Sensoren wie Laserscanner (LIDAR) oder synthetisches Apertur Radar (SAR) sind mittelfristig geplant. Weitere Trägerplattformen sollen durch eine Optimierung des Systems, hinsichtlich Gewicht und Größe, erschlossen werden. Auch der wachsende UAV-Markt liegt im Fokus dieser Arbeiten.

5 Literaturverzeichnis

- BONGARTZ, J., JENAL, A., KNEER, C., WEBER, I., BASCHEK, B., FRICKE, K., SCHÖL, A. & WYRWA, J., 2016: Ortsaufgelöste Messung der Wasseroberflächentemperatur an der Hahnöfer Nebelbe mit einem Tragschrauber. Tag der Hydrologie 2016, Koblenz, Volume: Forum für Hydrologie und Wasserbewirtschaftung, Heft 37, 16.
- JENAL, A., WEBER, I., KNEER, C. & BONGARTZ, J., 2015: Der Tragschrauber als Sensorplattform für die Fernerkundung. Publikationen der Deut. Gesellschaft für Photogrammetrie, Fernerkundung u. Geoinformation e.V., Band 24, Kersten, T. (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 16.-18. März 2015 in Köln, Tagungsband auf CD-ROM, 226-231.
- WEBER, I., JENAL, A., KNEER, C. & BONGARTZ, J., 2015: Gyrocopter-based Remote Sensing Platform. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **40** (7/W3), 1333-1337.
- NAUMANN, J.C., ANDERSON, J. E. & YOUNG, D.R., 2009: Remote detection of plant physiological responses to TNT soil contamination. *Plant and Soil* **329** (1-2), 239-248.
- ZINNERT, J.C., 2012: Plants as Phytosensors: Physiological Responses of a Woody Plant in Response to RDX Exposure and Potential for Remote Detection. *International Journal of Plant Sciences* **173** (9), 1005-1014.

Eignung von ALS, Aero- und UAS-Photogrammetrie zur Früherkennung und Erfassung von Tagesbrüchen

VOLKER SPRECKELS¹, ANDREAS SCHLIENKAMP¹, ANSGAR GREIWE² & LARS EBERHARDT³

Zusammenfassung: Im Verantwortungsbereich der RAG Aktiengesellschaft (RAG) liegen ca. 4.000 bekannte oberflächennahe Schächte und Tagesöffnungen (S&T) im südlichen Ruhrgebiet, die regelmäßig in Augenschein genommen werden. Der Steinkohlenabbau erfolgte zum Teil schon vor mehreren Jahrhunderten, so dass es zu Tagesbrüchen an der Erdoberfläche kommen kann. Tagesbrüche fallen nicht immer spontan, die Ereignisse kündigen sich gegebenenfalls durch Absacken der Füllsäule oder als Muldenbildung an. Die S&T haben Durchmesser um einen Meter und größer. Die Überwachung erfolgt derzeit durch regelmäßige Begehungen der bekannten Orte. Für eine großflächige Erfassung testet RAG seit Ende 2013 fernerkundliche Verfahren zur Detektion von Verdachtsbereichen. Die zentrale Frage ist: Welche Geländeänderungen (Mulden) können in welcher Genauigkeit mit welchem Sensor auf welchem Untergrund zu welcher Jahreszeit erfasst werden? Als Grundlage dienen jährliche großflächige Bildflüge mit Vexcel UltraCamX (UCX) und UltraCam Falcon (UCF) bei ca. 3 cm Ground Sample Distance (GSD), in 2015 erstmals simultan mit Airborne Laserscanning (ALS) bei ca. 20 Punkten/m² befliegen. Weiterhin werden spezielle Verdachts- und Überwachungsgebiete halbjährlich mit UAS bei 1 bis 3 cm GSD erfasst. Alle Befliegungsdaten werden mit denselben Passpunkten verknüpft, die mit festen Transformationsparametern über Echtzeit-GPS in das SAPOS-Netz der Landesvermessungsverwaltung NRW eingebunden werden. Entscheidend zur Erreichung vergleichbarer Genauigkeiten sind die Wahl der GNSS-Antenne und die Wahl des bildgebenden Sensors. Lagefehler der ALS-Steifen wurden über stereo-photogrammetrische Messungen kompensiert. Zur Untersuchung der Sensoren über die unterschiedlichen Befliegungszeiten seit 2013 wurden spezielle Referenzflächen auf unterschiedlichen Untergründen festgelegt (Asphalt, Pflaster, Dachflächen, Wiese). Der in einem FME-Workflow automatisierte multitemporale Ansatz ermittelt in der 1. Stufe über jede Referenzfläche für jeden Zeitstempel in der 3D-Punktwolke die aktuelle Rauigkeit und Bewegtheit. In der 2. Stufe erfolgt die Differenzenbildung für einen fest gebufferten Bereich um jeden bekannten S&T aus dem GIS der RAG herum, unter Berücksichtigung der in Stufe 1 ermittelten Bewertungsparameter. Es wird für jeden S&T ein Report für Außendienstmitarbeiter generiert, die anhand dieser Informationen Begehungen priorisieren und optimieren können. Zukünftig soll dieser Workflow um eine Zeitreihenanalyse für bekannte Schächte und um eine grid-basierte Suche zur Detektion von undokumentierten Verdachtsbereichen erweitert werden.

1 Einleitung

Im Schacht-Informationssystem der RAG sind derzeit über 5.000 Schächte und Tagesöffnungen (S&T) für das Ruhrgebiet enthalten (HAGER et al. 2014). Davon liegt die Mehrheit von etwa 4.000

¹ RAG Aktiengesellschaft, D-44623 Herne, E-Mail [volker.spreckels, andreas.schlienkamp]@rag.de

² Frankfurt University of Applied Sciences, Labor für Photogrammetrie und Fernerkundung, Nibelungenplatz 1, D-60318 Frankfurt am Main, E-Mail ansgar.greiwe@fb1.fra-uas.de

³ Technische Universität Bergakademie Freiberg, Fakultät Geowissenschaften, Geotechnik und Bergbau, Akademiestraße 6, D- 09599 Freiberg, Lars_Eberhardt2066@web.de

S&T im südlichen Ruhrgebiet, wo vor mehr als 300 Jahren der Steinkohlenbergbau oberflächennah begann. Die S&T liegen über das gesamte Gebiet verteilt und haben je nach Art und Weise des technischen Stands der Förderung unterschiedlichste Ausprägungen: von den im Süden an der Tagesoberfläche austreichenden Flözen mit Pingen und Stollenmundlöchern bis zum Tiefbauschacht mit Teufen von 1.500 m im nördlichen Ruhrgebiet. Die nicht mehr benötigten S&T wurden nach den jeweils zur Zeit der Ausführung vorherrschenden Möglichkeiten verfüllt. Bedingt z.B. durch die mit der Zeit einhergehende Kompaktion des Füllmaterials, Auswaschung, Erosion etc. kann es zu Sackungen der Füllsäule oder dem Zusammenbruch von Hohlräumen kommen, was sich an der Tagesoberfläche in Form von Mulden bis hin zu Tagesbrüchen zeigen kann (Abb. 1).



Abb. 1: Tagesbruch auf einer Weide (links) und einer Hauszufahrt (rechts)

Abbildung 2 zeigt eine Zeitfolge von Luftbildern für die Kleinzeche Meckel, deren Schächte 1926 verfüllt wurden, und für die in den 1950er Jahren eine möglicherweise erneut aufgenommene Steinkohlenförderung sowie im Jahr 2009 ein Tagesbruch erkennbar ist:

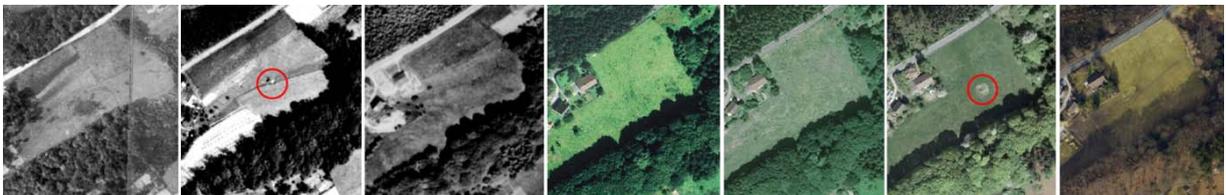


Abb. 2: Zeitfolge von Orthophotos 1926, 1952, 1969, 1998, 2006, 2009 und 2015, mit einem 2009 gefallenen und sanierten Tagesbruch. Quelle: <http://luftbilder.geoportal.ruhr/>

Aus Gründen der Gefahrenabwehr werden bei RAG die bekannten Tagesöffnungen nach einem zertifizierten Monitoringsystem regelmäßig überwacht, in der Regel mindestens ein Mal jährlich. Das geschieht bislang durch in Augenscheinnahme vor Ort. Daher ist das derzeitige Verfahren sehr zeit-, personal- und folglich kostenintensiv. Eine Reduzierung dieses Aufwandes ist aus betrieblichen Gründen umzusetzen. Die S&T liegen in Wäldern, auf Wiesen, Weiden, Äckern, Privatgrundstücken, zu denen Zugang möglich sein und gewährt werden muss. Oftmals ist es durch die Topographie und dem Bewuchs nicht erkennbar, ob es im Laufe der Zeit zu Setzungen an S&T von wenigen Dezimetern gekommen ist. Hier können die fernerkundlichen Verfahren der Aero- und UAS-Photogrammetrie in einseharen Bereichen wichtige Informationen liefern. In Waldbereichen finden photogrammetrische Verfahren ihre Grenzen, so dass hier seit 2015 das Airborne Laserscanning (ALS) getestet wird, wobei der Laserscanner IGI Litemapper LM 7800 zusammen mit der Digitalen Luftbildkamera Vexcel UltraCam Falcon (UCF) in einem

Bildflugzeug eingesetzt wird. Ein als signifikant zu bestimmendes Maß der Veränderung der Tagesoberfläche zur Frühdetektion von Tagesbrüchen wurde auf ± 10 cm festgelegt, so dass für das Endprodukt und damit über die Messverfahren und dem Verarbeitungsprozess hinweg eine Genauigkeit von ± 3 cm gegeben sein muss.

Für eine großflächige Erfassung testet RAG seit 2013 UAS-Verfahren in Forschungsstudien sowie im Tagesgeschäft zur Detektion von S&T-Veränderungsbereichen (STUDIE 193 2013; STUDIE 196 2013; STUDIE 200 2014). Im Rahmen einer Diplomarbeit an der TU Bergakademie Freiberg wurden im Jahr 2015 sehr umfangreiche Untersuchungen zur Erstellung eines multitemporalen sowie multisensoralen Workflows zur Früherkennung bzw. Detektion von Tagesbrüchen durchgeführt (EBERHARDT 2015). Abbildung 3 zeigt eine Übersicht mit einer Ausdehnung von ca. 12 km x 6 km mit den Bereichen, die ab 2013 mit fernerkundlichen Verfahren überwacht werden, südlich von Bochum und der Ruhr gelegen.

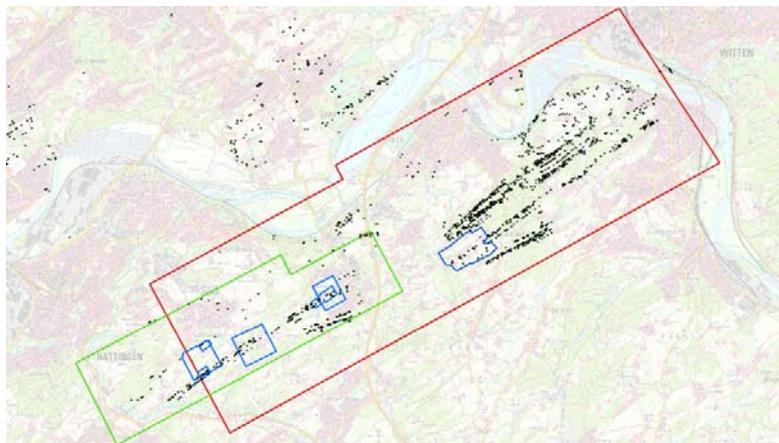


Abb. 3: Übersicht des Untersuchungsgebietes: Lage der S&T (schwarz), UAS-Befliegungsbereiche ab 2013 (blau), UCF Befliegung 2014 (grün), kombinierte Befliegungen UCF mit ALS 2015 & 2016 (rot). Abgebildeter Bereich ca. 12 km x 6 km

2 Grundlagen, Genauigkeiten, Referenzdaten

2.1 Untersuchungsgebiet

Das Projektgebiet liegt im südlichen Ruhrgebiet zwischen Witten und Hattingen. Das Gebiet des UCF-Bildfluges 2014 weist eine Ausdehnung von 6 km x 2 km auf. Die UCF Befliegung aus dem Jahr 2015 umfasst eine Fläche von 10 km x 3 km. Diese Flächen über Altbergbaugebieten der RAG werden nun jährlich per Bildflug erfasst. Charakteristisch für diesen Bereich ist eine sehr hügelige Topografie mit Höhenunterschieden von bis zu 130 m. Es liegen landwirtschaftlich genutzte Flächen, verstreute Waldgebiete und verstreute Bebauung abwechselnd nebeneinander. In Abbildung 3 sind neben den zwei Bildflugbereichen (grün und rot), jene Bereiche in blau dargestellt, die halbjährlich per UAS befliegen werden. Die per ALS abgetastete Fläche entspricht der Ausdehnung des UCF-Bildfluges (rot), da beide Sensoren simultan im Flugzeug eingesetzt wurden.

2.2 Befliegungen

Für die Detektion von Höhenänderungen im Bereich von wenigen Zentimetern auf der Grundlage bildgestützter Punktwolken müssen viele Faktoren sinnvoll ineinandergreifen. Zunächst gilt es,

den Bild- bzw. UAS- und ALS-Flug sorgfältig vorzubereiten und die für den Aufnahmepurpose geeignete Sensorik und Plattform auszuwählen. Nach erfolgreicher Durchführung stellt schließlich die Wahl geeigneter Auswertesoftware und leistungsfähiger Hardware den wichtigen Schlusspunkt einer photogrammetrischen Auswertung dar. Diese Eckpunkte werden im Folgenden eingehend behandelt. Tabelle 1 zeigt eine Übersicht zu den Bild- und UAS-Flügen sowie den verwendeten Kameras:

Tab. 1: Übersicht Flüge.

Befliegung	Kamera	h_g [m]	GSD [m/pix]
0114	DP2 Merrill	85	0,014
0214	DP1 Merrill	119	0,029
1114	DP2 Merrill	71	0,011
0315	DP2 Merrill	75	0,025
UCX	UCX	775	0,053
UCF	UCF	400	0,029

2.2.1 Kombiniertes Aero-/ALS Flug

In diesem kombinierten UCF und ALS Flug zeigte sich im Vergleich zum UCF-Datensatz eine unzureichend genaue Verknüpfung der ALS Flugstreifen, die über einen Lageversatz von bis zu 35 cm zu Höhenfehlern von einigen Dezimetern führte. Am Beispiel einer Dachfläche, die mehr als 30 cm Höhendifferenz zwischen zwei ALS Streifen zeigt, ist ein Lageversatz von ca. 30 cm (s. Abb. 4) erkennbar. Ein Versatz in dieser Größenordnung resultiert aus der direkten on-board-Referenzierung ohne Passpunkte am Boden, was zu einer ungenügend präzisen Bestimmung des Flugweges führt. Die Passgenauigkeit konnte im Nachhinein über stereo-photogrammetrische Messungen von Dachflächen als Referenzflächen von RAG verbessert werden.

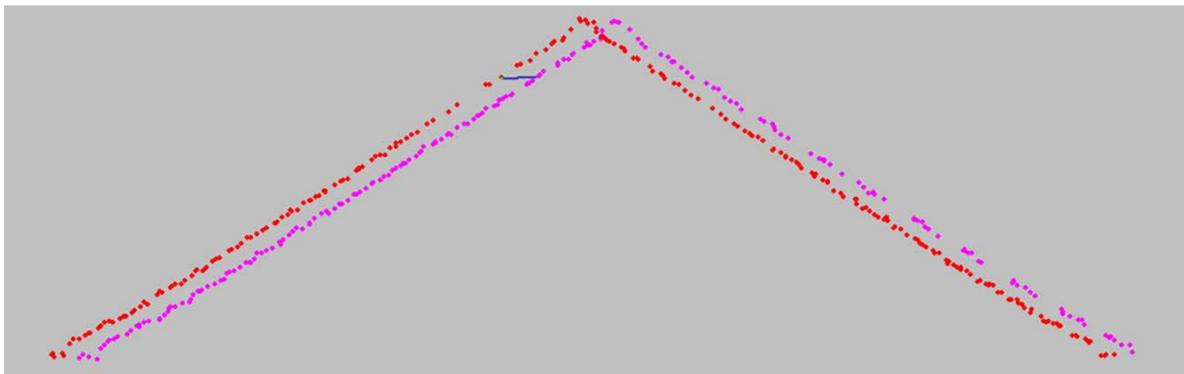


Abb. 4: Lageversatz und daraus resultierenden Höhenversatzes benachbarter ALS-Flugstreifen an einer Dachfläche, hier ca. 30 cm Abweichung in Lage- und Höhe.

2.2.2 UAS-Sensor

Für den Einsatz am UAS werden mittlerweile zahlreiche leichtgewichtige Kameras angeboten, so dass eine Bandbreite von nicht-modifizierten Consumer-Kameras bis hin zu speziell entwickelten

Mittelformatkameras im Einsatz ist. Die Abmessungen der Sensoren und somit die Detektorfläche des einzelnen Pixels am Sensor variieren hierbei stark. Allerdings beruhen fast alle diese Systeme auf dem sogenannten „Bayer-Pattern“ zur Aufnahme der Farbinformationen. Jedem Detektorelement wird ein einziger Farbfilter (Rot oder Grün oder Blau) vorgeschaltet. Die volle Farbinformation über den Rot- Grün- und Blauanteil einer reflektierenden Oberfläche kann somit nicht simultan pro Pixel aufgezeichnet, sondern allein durch Interpolation mit Farbwerten der Nachbarelemente erreicht werden (VERHOEVEN 2010).

Die Aufzeichnung der vollen Farbinformationen des Rot- Grün- und Blauanteils für ein einzelnes Detektorelement (Pixel) erfolgt hingegen bei den sogenannten Foveon-Sensoren nach folgendem Prinzip: Anstelle eines vorgeschalteten Farbfilters wird die Trennung der Farbinformationen hier durch die wellenlängenabhängige Eindringtiefe der elektro-magnetischen Welle in den CMOS Chip genutzt (GILBLOM et al. 2003). Daher entspricht diese Aufnahmeweise eher einem Luftbildfilm für Reihenmesskammern. Diese unterschiedliche Art der Aufnahme von Farbinformationen führt zu einem erhöhten Mikrokontrast bei Foveon-Sensoren, verglichen mit Aufnahmen, die durch ein Bayer-Pattern aufgezeichnet werden (Abb. 5).

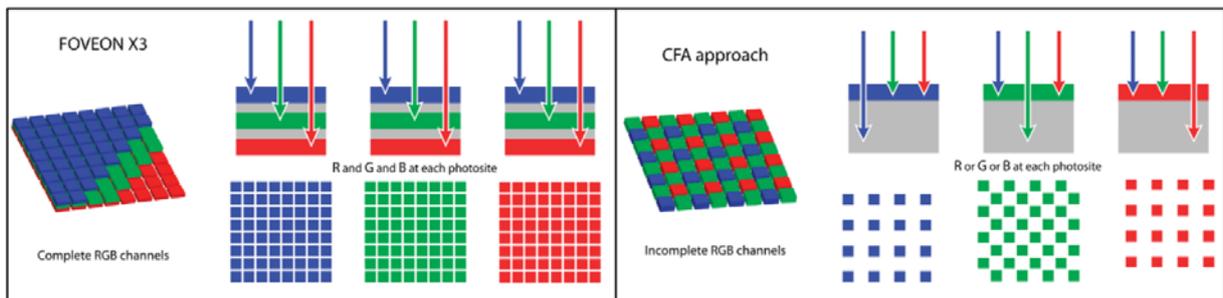


Abb. 5: Foveon (links) und Bayer/CFA – Color Filter Array (rechts) (VERHOEVEN 2010)

Abbildung 6 zeigt die MTF-Kurve eines Bayer-Patterns im Vergleich zu einem Foveon-Chip. Das Kontrastaufklärungsvermögen bei sehr feinen Strukturen – hier liegt die Ortsfrequenz im Bereich der Nyquist-Frequenz – beträgt bei Bayer-Pattern weniger als 10%. Foveon-basierte Sensoren weisen in diesem Bereich 30% auf (GREIWE & GEHRKE 2013).

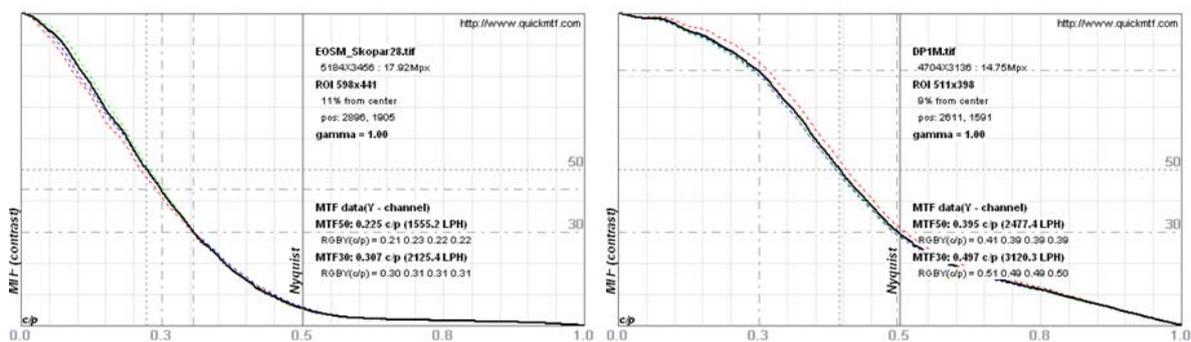


Abb. 6: MTF Kurve eines Bayer-Pattern (links) und eines Foveon-Chips (rechts) (GREIWE & GEHRKE 2013)

Die UAS-Befliegungen werden im Gegensatz zu den anderen Befliegungsmissionen in kleineren Gebieten durchgeführt. Bei der hohen Abtastrate von 1 bis 3 cm bestehen die Bildverbände aus bis zu 2.500 Aufnahmen. Als Auswerteansatz wurde hier Structure from Motion gewählt. Die

Ergebnisse dieses Auswertansatzes sind hochgradig von der Bildqualität der auszuwertenden Aufnahmen abhängig. Dieses gilt insbesondere für die 3D-Punktwolke, welche als Endergebnis des Dense Matching die Grundlage für die Ableitung des Oberflächenmodells bildet. Hier können je nach eingesetztem Sensor – Bayer-Pattern oder Foveon-Chip – qualitative Effekte beobachtet werden, die durch die MTF-Messungen bestätigt werden. Zur Bewertung des Einflusses der Bildqualität auf das Dense Matching wurden drei Sensoren ausgewählt, um eine 3D-Punktwolke von einem Testkörper zu erstellen. Zur Verfügung standen das iQ180 Back (Mittelformat Phase One), 80 MPix und ein Sony ILCE7s (Vollformat Sony Alpha 7s) mit 12 MPix, welche beide auf dem Bayer-Pattern beruhen. Zum Vergleich wurde ein Foveon-Chip (Merrill) mit 12 MPix im APS-C Format gewählt. Die Detektorgröße auf dem Sensor beträgt zwischen $5\ \mu\text{m}$ und $6\ \mu\text{m}$. Es wurden in drei unterschiedlichen Ebenen jeweils 36 Aufnahmen erzeugt. Die Aufnahmeentfernung (Bildmaßstab) wurde für jeden Sensor individuell gewählt, um die Abtastrate konstant zu halten. Bereits die qualitative Betrachtung der Ergebnisse zeigt wesentliche Unterschiede in den jeweiligen 3D-Punktwolken, siehe Abbildung 7. Während der linke und mittlere Testkörper (Bayer-Pattern) ein hohes Rauschen aufweisen, ist das rechts dargestellte Ergebnis des Foveon Sensors deutlich homogener.



Abb. 7: 3D-Punktwolke eines Testkörpers: Sensoren mit Bayer-Pattern (links und Mitte). Rechts das Ergebnis aus Bilddaten eines Foveon-Chips.

Im Praxiseinsatz der UAS-Befliegungen wird das Ergebnis der Laborversuche bestätigt. Hierzu wurden zwei Befliegungen mit unterschiedlichen Sensoren (Sony ILCE, Foveon Merrill) analysiert, deren Aufnahmeparameter GSD, Überlappungsverhältnisse und Bildmaßstab vergleichbar waren. Der Fokus des Vergleiches lag auf der Rekonstruktionsgüte der 3D-Punktwolke. Ebene Flächen sollten in der Punktwolke im Profil auch als ebene, gleichmäßige Anordnung der Punkte im Raum erscheinen. Verglichen wurden natürliche Objektoberflächen (Bordsteine). Die Punktwolke des Bayer-Pattern - Sensors weist eine Streuung („Rauschen“) im Bereich von 4 cm auf, in dem Datensatz mit einem Foveon-Chip ist es lediglich 1 cm, wie in Abbildung 8 ersichtlich.

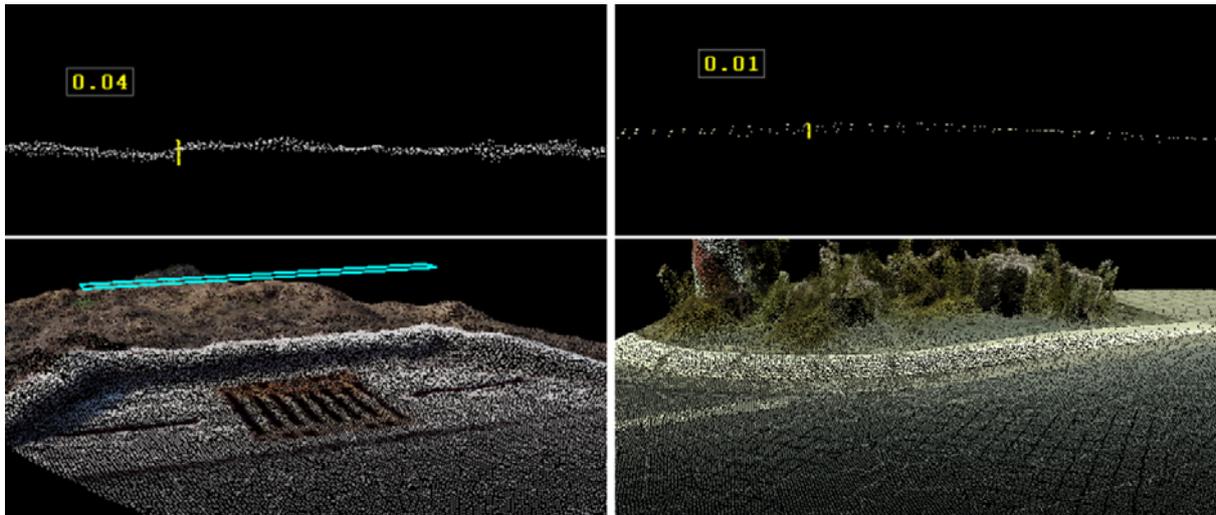


Abb. 8: Rauschen der 3D-Punktwolke an einem Bordstein: 4 cm bei Bayer-Pattern (links), 1 cm bei Foveon-Chip (rechts)

Daraus folgt, dass die Bildqualität des Sensors die entscheidende Größe im Workflow zum Erreichen hoher bis höchster Genauigkeiten zur Ermittlung signifikanter Bodenbewegungen aus Zeitreihenvergleichen von Wiederholungsbefliegungen ist.

Es ist zudem streng darauf zu achten, dass auch im weiteren Bearbeitungsprozess mit der Originalauflösung der Bilder gearbeitet wird. Ein Weiterarbeiten mit verringerter Bodenauflösung in den Bildpyramiden führt zu einer Glättung der Punktwolke und darüber zu Detailverlusten. In der Regel weisen diese Punktwolken auch einen Höhenoffset gegenüber Punktwolken aus Originalbilddaten auf. Abbildung 9 zeigt, dass ein Höhenprofil aus der UCX-Befliegung bei vierfach verkleinertem Bildmaterial (Einstellung in Agisoft PhotoScan: medium, blau) zu jenem aus Originalbilddaten (ultra high, gelb) einen Höhenoffset von 4 cm erfährt.



Abb. 9: Höhenoffset der Punktwolkenberechnung von ca. 4 cm zwischen Originalauflösung (gelb) und vierfach verkleinerter Bildpyramide (cyan)

2.3 Terrestrische Photogrammetrie – Referenz Wasserbehälter

Im Untersuchungsgebiet befindet sich der in Abbildung 10 dargestellte Hochwasserbehälter. Er eignet sich aufgrund seiner ebenen Oberfläche (Waschbeton) sehr gut als Referenzobjekt zur Untersuchung der Genauigkeit der einzelnen Punktwolken. Um als Referenz eine Oberfläche höherer Genauigkeit zu erhalten wurde der Hochwasserbehälter mittels terrestrischer Photogrammetrie erfasst. Zum Einsatz kam hier die Nikon D800 mit einem 14mm Objektiv. Aus 106 Aufnahmen wurde mittels Structure from Motion eine Punktwolke mit einem GSD von unter 1 mm abgeleitet. Georeferenziert wurde das Objekt über Passpunkte, die per Echtzeit-GPS bestimmt wurden. Zusätzlich wurden viele Strecken am Behälter gemessen, um als Maßstabsvergleich zu dienen. Zwei Gliedermaßstäbe wurden auf den Behälter gelegt, um eine weitere Kontrollmöglichkeit zur Bestimmung der Referenzfläche zu haben.



Abb. 10: Ansicht (links) und Aufsicht der terrestrischen Referenzfläche „Hochwasserbehälter“ mit Referenzpunkten (rechts)

2.4 GPS-Messungen zur Passpunktbestimmung

Die Georeferenzierung der Flugzeug- und Kopter-gestützten Bildflüge erfolgt über am Boden signalisierte Passpunkte. Da RAG seit 2003 für das Ruhrrevier einen festen, einheitlichen Transformationsparametersatz zur Überführung von geozentrisch-kartesischen ETRS89 Koordinaten in das Gauß-Krüger Koordinatensystem verwendet, kommt als Unschärfe in diesem Verfahren nur die Güte der Passpunktmessung zum Tragen (SPRECKELS 2003). Bei den verwendeten GSD von 1 cm bis 3 cm muss nunmehr darauf geachtet werden, dass der Messnagel auf jeden Fall mittig im Signal liegt. Die Koordinatenbestimmung erfolgt althergebracht über Echtzeit-GNSS-Verfahren im Anschluss an SAPOS-Stationen, gemäß den GPS-Richtlinien NRW mit einer Genauigkeit von ± 3 cm in Lage und Höhe (GPS-RICHTLINIEN). Diese Genauigkeit reicht gerade noch für die UCX- und UCF Blöcke aus, für die UAS-Projekte mit GSD von 1 cm zeigte sich in den Projekten, dass die Passpunktbestimmung mit dem Standard-GPS-Antennen zu Blockverbiegungen führte. RAG ist im Besitz von sechs hochpräzisen, ca. 7 kg schweren Leica AR 25 Choking Antennen und von sieben 3G+C Antennen. Vergleichende Untersuchungen der RAG im Forschungsprojekt FE-0760 ABSMon (ABSMON) zeigten, dass die schwere Leica AR 25 und die nur wenige 100g leichte 3G+C Antenne im Postprocessing vergleichbare relative Lage- und Höhengenaugigkeiten im Bereich von ± 3 mm erreichen und somit für diese Antennen ein stabiles Antennenphasenzentrum vorliegt. Im Außendienstesinsatz wurde die 3G+C Antenne zur Passpunktbestimmung eingesetzt und lieferte mit ± 1 cm Lage- und Höhengenaugigkeit Passpunktkoordinaten, die auch für UAS-Befliegungen verwendet werden können. Im Untersuchungsgebiet liegen Passpunkte allerdings auch in Gebieten, die keinen GPS-Empfang ermöglichen. Hier müssen die Passpunkte mit Tachymetern eingemessen werden.

3 Untersuchungen

3.1 Bedingungen

Für die Untersuchungen kam eine große Palette an Software und Hardware zum Einsatz. Für die Orientierungen und Ableitung der Punktwolken wurde Agisoft PhotoScan und Erdas Imagine – ORIMA eingesetzt. Die Weiterverarbeitung erfolgte mit FME (Feature Manipulation Engine – Safe Software) und CloudCompare. Für ein späteres Geomonitoring ist es entscheidend, die eingesetzte Software richtig und für den Anwendungsfall angepasst einzusetzen. Auch wenn viele

der Structure from Motion Programme auf den ersten Blick wie eine Blackbox wirken, gibt es eine Vielzahl an Parameter und Analysefunktionen, um eine saubere Orientierung zu rechnen, und damit auch eine ordentliche Punktwolke abzuleiten. In intensiven Vergleichen mit klassischen Orientierungsprogrammen wie ORIMA konnte RAG die Stärken und Schwächen der einzelnen Produkte herausstellen und so für den jeweiligen Anwendungsfall richtig einsetzen. Gerade für die großflächigen Befliegungen mit der UCX oder UCF ist es beim Einsatz von Agisoft PhotoScan entscheidend, eine dichte, dem Gebiet angepasste Passpunktverteilung und ~messung gleichmäßig über den Block verteilt und nicht nur am Blockrand zu haben. Ist dies nicht der Fall, kommt es zu Blockdurchbiegungen von 1 - 2 dm.

Durch die Vielzahl an UAS Befliegungen der RAG in unterschiedlichen Bildflugkonfigurationen (Quer- und Längsüberdeckung) konnte in den vergangenen Jahren neben dem Einsatz des passenden Sensors auch die Wichtigkeit der Längs- und Querüberdeckung festgestellt werden. Anders als bei einem klassischen flugzeuggestützten Bildflug genügt bei UAS-Kampagnen nicht die Standardüberdeckung von 60% Längs- und 30 % Querüberdeckung. Oft wird aufgrund von zu geringer Akkukapazität, Zeit, Wetterbedingungen – aber auch durch fehlendes Knowhow – das Gebiet nicht mit ausreichender Überdeckung befliegen, was dann zu schlechteren Ergebnissen führt. Werden diese genannten Bedingungen bei allen Bildflügen eingehalten, können optimale Ergebnisse erzielt werden und dadurch belastbare Vergleiche ermöglichen.

3.2 Multisensoraler Vergleich

3.2.1 Multisensoraler Vergleich am Referenzobjekt Hochwasserbehälter

Das in Kapitel 2.4 beschriebene Objekt, ein Wasserbehälter mit einer Waschbetonoberfläche, dient im Folgenden als Referenz für den Vergleich mit den Punktwolken aus unterschiedlichen Sensoren, Plattformen, bzw. unterschiedlichen GSD und Jahreszeiten der Aufnahmen.

Um ein Genauigkeitsmaß für jede einzelne Punktwolke zu ermitteln, wird eine Ebene oder Flächenkurve durch jede einzelne Punktwolke gelegt. Bei der Erstellung dieser Ebene wird als Ergebnis der Ausgleichung der RMS-Wert ermittelt. Dieser beschreibt die mittlere Standardabweichung aller Objektpunkte (gesamte Punktwolke) und gibt somit Auskunft über das Genauigkeitsniveau der Ausgleichung (LUHMANN 2010). Der Wert kann auch als Punktrauschen verstanden werden. Mit der unabhängigen Berechnung der Genauigkeitsmaße wird sichergestellt, dass die Passpunktdifferenzen keinen Einfluss haben.

Fehler! Verweisquelle konnte nicht gefunden werden. zeigt eine Übersicht der ermittelten RMS-Werte für jeden Datensatz:

Tab. 2: Übersicht der RMS-Werte am Hochwasserbehälter

Punktwolke	terrestrisch	0114	0314	1114	0315	UCF	UCX	ALS
GSD in cm	0,1	1,4	2,9	1,1	2,5	2,8	5,3	-
RMS in cm	0,3	0,4	-	0,3	0,4	0,6	2,0	0,9

Die terrestrische Punktwolke bildet die Vergleichsgrundlage und weist entgegen den Erwartungen kein wesentlich geringeres Punktrauschen als die beiden hochauflösenden UAS-Punktwolken 0114 und 1114 auf. Jedoch erfasst die terrestrische Punktwolke mit einer GSD von 1 mm auch

jeden Kieselstein, so dass in diesem Fall der RMS die echte Oberflächenrauigkeit darstellt, und nicht das Punktrauschen. Die UAS-Flüge 0114 und 1114 erzeugten die höchst aufgelösten GSD. Die Erwartung, dass deren RMS-Werte am geringsten ausfallen, ist erfüllt worden. Beide weichen nur wenige Millimeter von ihrer ausgleichenden Ebene ab. Eine Darstellung der unterschiedlich gelagerten Punktwolken in Form eines Profils ist sehr anschaulich. Hier fällt auf, dass bei zunehmender GSD das Punktrauschen größer wird und der Höhenoffset zur Referenzfläche zunimmt (Abb. 11).

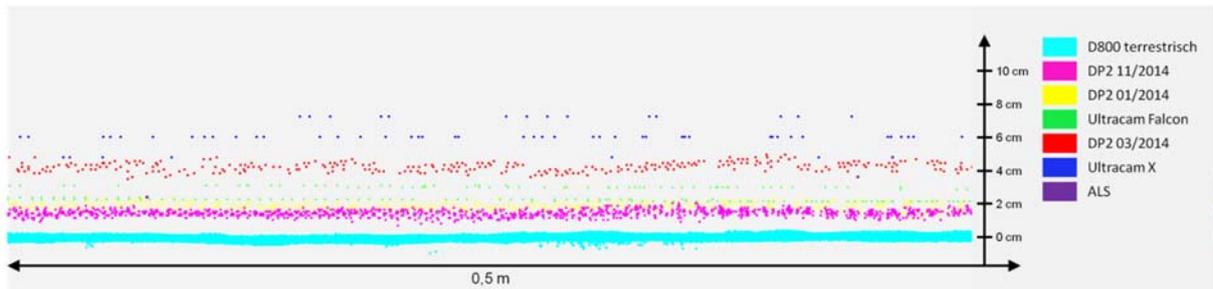


Abb. 11: Profilschnitt durch die Oberfläche des Hochwasserbehälters – Punktwolkenvergleich

3.2.2 Testflächen als Referenz

Ein Tagesbruch kann in einem Altbergbauggebiet nahezu überall auftreten. Vielfach sind ehemalige Grubenbaue heute durch Straßen oder Parkplätze überbaut. Aus diesem Grund ist es wichtig, das Verhalten der Punktwolken auf diversen Oberflächenmaterialien an kleinen Testflächen zu untersuchen. Für die Untersuchungen sind Testflächen auf den Materialien Wiese, Asphalt, Pflaster, und Dachflächen ausgesucht worden. Ganz bewusst sind die meisten Testflächen für das Material Wiese generiert worden, da derartige Flächen sehr häufig in den RAG-Altbergbaugebieten vorkommen. Asphalt und Pflaster treten vorrangig bei öffentlichen Flächen und Verkehrswegen auf. Zusätzlich sind Dachflächen digitalisiert worden, um steiles Gelände zu repräsentieren. Wie schon im Kapitel zuvor, werden zur Bestimmung des Genauigkeitsmaßes ausgleichende Flächenkurven für jede Punktwolke erstellt und der RMS Wert bestimmt. Im Folgenden sind die Ergebnisse für die Testflächen Wiese und Asphalt dargestellt.

Testflächen Wiese

Das große Problem bei der Bestimmung eines Genauigkeitsmaßes für Wiesenflächen ist deren je nach Jahreszeit mehr oder weniger unregelmäßige Oberfläche im Hinblick auf das Anlegen eines ausgleichenden Objekts. Die hier anzulegenden Flächenkurven stellen tatsächlich nur eine grobe Annäherung an die durch unterschiedliche Vegetationsstände beeinflusste Oberfläche dar. Die resultierenden Genauigkeitsmaße sind daher als sehr grob einzustufen, geben sie doch mehr ein Maß für die (eigentlich unbekannt) Oberflächenrauigkeit als für das Rauschen der untersuchten Punktwolke an. Dennoch sind die Ergebnisse für die weitere Bearbeitung geeignet, da eine großzügigere, konservative Annahme bei der Erfassungsgenauigkeit der Topografie praxistauglicher ist, als eine unter annähernd idealen (Labor-) Bedingungen abgeleitete. Dies erschwert zwar das Erkennen kleinster Bodenabsenkungen, führt aber zu weniger Fehlinterpretationen.

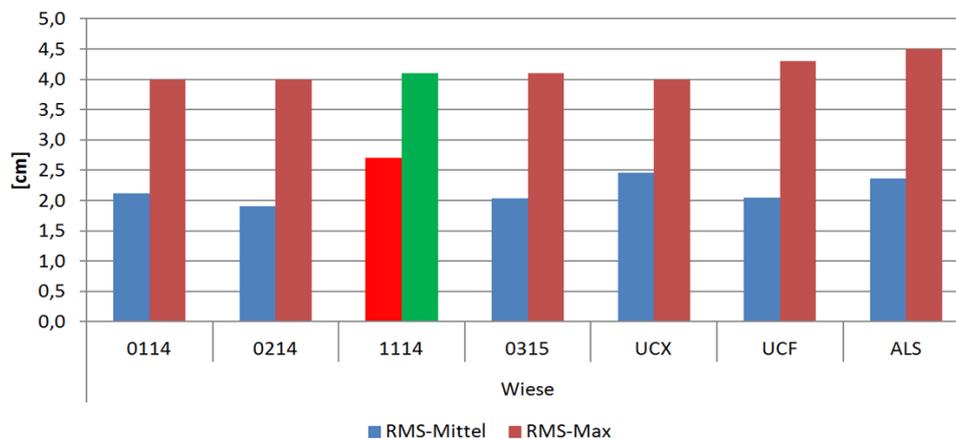


Abb. 12: Übersicht RMS-Werte Testfläche Wiese

Abbildung 12 zeigt, dass alle Punktwolken im Mittel von $\pm 1,9$ cm bis $\pm 2,7$ cm streuen. Die mitunter doppelt so großen maximalen Werte resultieren zum einen aus sich ändernder Oberflächenrauigkeit, zum anderen aus unterschiedlicher Geländeneigung. Bemerkenswert ist das höhere Rauschen der Referenzpunktwolke 1114 gegenüber allen anderen Datensätzen, besonders aber im Vergleich zu 0214, obwohl diese aus geringerer Bildüberdeckung prozessiert wurde und eine geringere GSD aufweist. Eine nähere Betrachtung zeigt, dass die Vegetation zwischen beiden Flügen stark unterschiedlich ist. Aufgrund der sehr hohen GSD von 1114 und der dichten Vegetation ist die Oberfläche für einen automatisierten Rekonstruktionsalgorithmus zu unruhig. Das führt zu vielen Ausreißern, die das höhere Rauschen verursachen. Dieses Beispiel zeigt, dass es sinnvoll ist, die Genauigkeit bei der Erfassung der Oberfläche durch den jeweiligen Sensor eher konservativ anzusetzen.

Testflächen Asphalt

Diese Testfelder sind in Gebiete gelegt worden, die nicht durch Vegetation verdeckt sind, um die Rekonstruktionsqualität auf Asphalt ohne störende Einflüsse zu überprüfen.

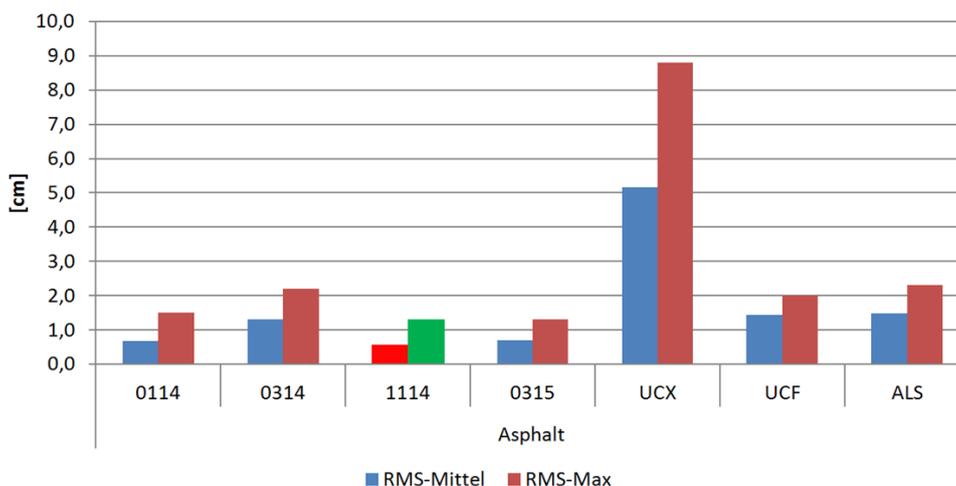


Abb. 13: Übersicht RMS-Werte Testfläche Asphalt

Alle UAS-Punktwolken weisen in Abbildung 13 ein signifikant geringeres Rauschen als die niedriger aufgelöste UCX-Punktwolke auf. Der UCF-Datensatz mit 2,8 cm GSD rauscht nur geringfügig stärker als die UAS-Datensätze, aber beinahe vier Mal weniger als UCX. Da die beiden Luftbildkameras bauartgleich sind, üben vermutlich die große Flughöhe bzw. die geringere GSD der UCX sowie der unterschiedliche Sonnenstand einen wesentlichen Einfluss auf die Rekonstruktion aus. Innerhalb der UAS-Punktwolken weist 0214 den größten Wert auf, was möglicherweise auf die geringe GSD von 2,9 cm zurückzuführen ist. Die Referenzpunktwolke 1114 streut im Mittel mit $\pm 0,6$ cm erwartungsgemäß am wenigsten. Besonders hervorzuheben ist die hohe innere Genauigkeit von 0315. Trotz einer wesentlich geringeren Auflösung von 2,5 cm GSD streut dieser Datensatz nahezu gleichstark wie der Referenzdatensatz 1114. Es zeigt sich hier, dass die Wahl der Kamera mit dem Foveon Sensor ein entscheidendes Kriterium ist, um eine hohe Rekonstruktionsgenauigkeit zu erreichen.

Der ALS rauscht vergleichbar stark wie UCF, jedoch zum Teil erheblich stärker als der Referenzdatensatz 1114 (siehe Abb. 14). Im Vergleich zu 1114 zeigt die ALS Punktwolke ein für ALS typisches Wellenmuster.

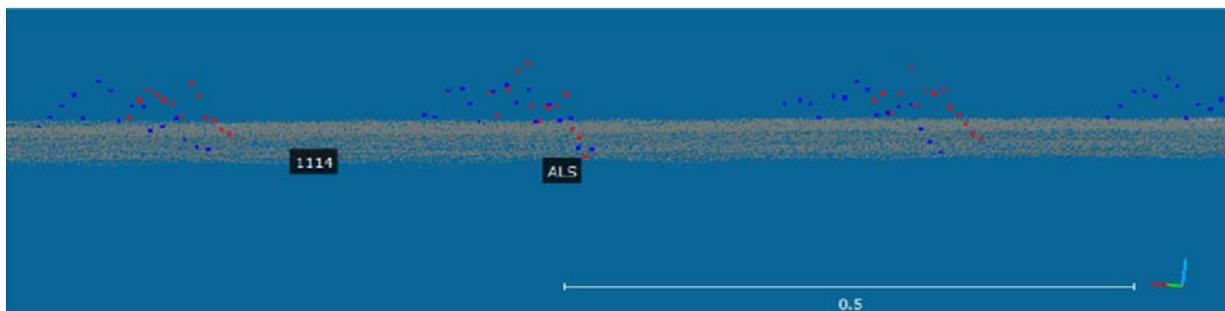


Abb. 14: Vergleich der Punktwolken aus UAS 1114 (beige) zu ALS (zwei Flugstreifen, rot und blau)

Die Punktwolken für wiesenähnliche Bereiche rauschen etwa gleich stark, mit Ausnahme von 1114, welche aufgrund hoher Vegetation und der vorliegenden GSD die Oberflächenrauigkeit besonders gut wiedergibt. Insgesamt liefern die UAS-Punktwolken die geringsten RMS-Fehler, mit Ausnahme der wiesenähnlichen Testflächen. Besonders hervorzuheben sind die Referenzpunktwolke 1114 und 0315, die fast immer innerhalb eines Zentimeters streuen. Die multisensoralen Untersuchungen zeigen das hohe Potential der photogrammetrischen Vermessung mittels Foveon Sensor. Ein Punktrauschen im Bereich von 1 bis 3 cm zeigt die Praxisfähigkeit des routinemäßigen Einsatzes von Kameras zur Detektion vertikaler Bodenbewegungen. Es bleibt zu untersuchen, mit welcher Genauigkeit Veränderungen in Abhängigkeit der Jahreszeit und des Vegetationsstandes ableitbar sind. Die in diesem Abschnitt ermittelten inneren Genauigkeitsmaße sind daher beim multitemporalen Vergleich im nachfolgenden Abschnitt zu berücksichtigen.

3.3 Multitemporaler Vergleich

Für den multitemporalen Vergleich wurde mit Hilfe des Programms FME (Feature Manipulation Engine, Safe Software) ein Workflow erstellt, der aus den unterschiedlichen Punktwolken Höhendifferenzen ableitet und Statistiken zu diesen aufstellt. Die Höhendifferenzen wurden für dieselben Testflächen Wiese, Asphalt, Pflaster, und Dachflächen ermittelt. Wichtig war es, im Vorfeld sinnvolle Vergleichsdatsätze auszuwählen. Da es für den ALS noch keinen zweiten Datensatz gibt, ist ein temporaler Vergleich zurzeit noch nicht möglich. Aufgrund der von der Jahreszeit abhängigen Vegetationsstände ist ein Vergleich zwischen den UAS Befliegungen getrennt nach Frühjahr (0114, 0214, 0315) und Herbst (1114) sinnvoll. Es sind zunächst die

Differenzen zwischen den Mittel- und Median-Werten aller Testflächen gebildet worden. Der robuste Median glättet bspw. an jenen Stellen das Modell, wo ein Tagesbruch entstanden ist. Der Mittelwert hingegen erfasst diese großen Abweichungen und wird verzerrt. Die Größe der Differenz kann somit auf eine mögliche Verdachtsfläche hinweisen. Abbildung 15 zeigt am Beispiel eines Tagesbruchs innerhalb der Testfläche Wiese_11 die Unterschiede zwischen Mittelwert und Median für die Epochen 1114-0114 und 0315-0214.

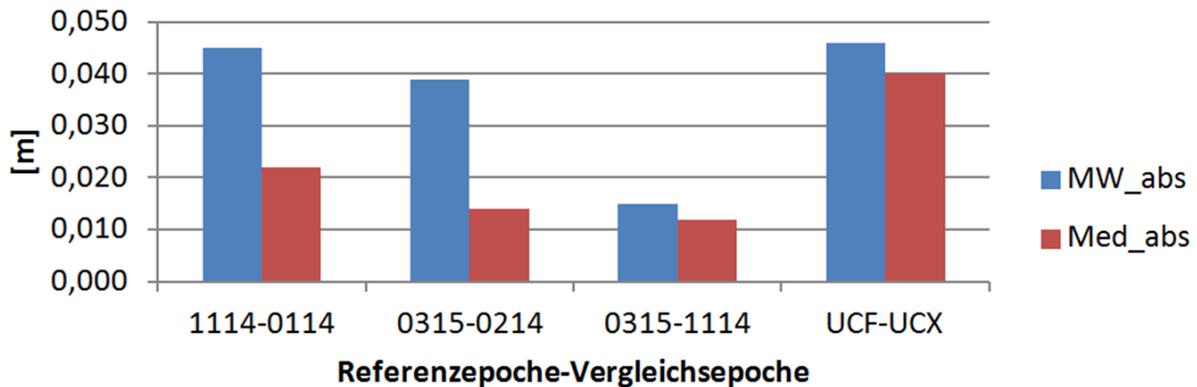


Abb. 15: Vergleich der Höhendifferenzen im Bereich eines Tagesbruchs: UAS zu UAS, UCF zu UCX)

0315-1114 und UCF-UCX zeigen keinen großen Unterschied zwischen Mittelwert und Median. Die Differenz der UAS Flüge 0315-1114 sollte keinen Tagesbruch mehr aufweisen, da die Flüge nach der Sanierung der Fläche durchgeführt wurden. Die Differenz der UltraCam Bildflüge könnte aber einen Unterschied für den Bereich um den gefallenen Tagesbruch zeigen, jedoch ist hier aufgrund der unterschiedlichen GSD und dem großen Rauschen der UCX Punktwolke eine signifikante Detektion des Tagesbruchs nicht möglich (siehe auch innere Genauigkeit UCX).

Doch nicht nur ein Tagesbruch, auch Vegetationswechsel oder verschobene Gegenstände können den Mittelwert verzerrten, was eine Lokalisierung eines Tagesbruchs alleine über die Differenzen erschwert. Die Min- und Max-Werte sind beim Material Wiese für die Detektion von Tagesbrüchen ungünstig, da diese ebenso leicht durch vereinzelte, hohe Vegetation verfälscht werden können. Die hohen Max-Werte beim UCF-UCX-Vergleich für das Material Asphalt resultieren aus dem starken Rauschen der UCX. Der Abbildung 16 kann entnommen werden, dass, unabhängig vom Sensor, gemittelte Höhenunterschiede von weniger als 6 cm detektiert wurden und 1,5 bis 3 cm den Massendurchschnitt bilden. Höhere Werte treten tendenziell beim Material Wiese auf.

Absolute Genauigkeiten von Höhenunterschieden können nicht für die Evaluierung der Ergebnisse einer photogrammetrischen Rekonstruktion gegenüber anderen Messverfahren angegeben werden. Gerade für wiesenähnliche Flächen stellt die (unbekannte) und je nach Jahreszeit stark unterschiedliche Oberflächenrauigkeit ein noch ungelöstes Problem dar. Generell scheint ein Vergleich zwischen zwei DHM sinnvoll, die beide aus Befliegungen im Frühjahr hervorgegangen sind. Der Vorteil ist, dass noch keine "neue" Vegetation die "ältere", flach gedrückte, überragt und die Oberflächenrauigkeit somit geringer ist (Epochenvergleich 0315-0214).

Im Zuge der Umsetzung, die Passpunkte für die Georeferenzierung der Flüge zukünftig ausschließlich mit 3G+C-Antennen zu messen und der Annahme, dass damit Genauigkeiten von ca. 1 cm erreicht werden, sollten für einen Epochenvergleich (0315-0214) Höhenunterschiede als signifikant ab ca. 4 cm (versiegelte Oberfläche) und ab ca. 7 cm (Wiese) gelten.

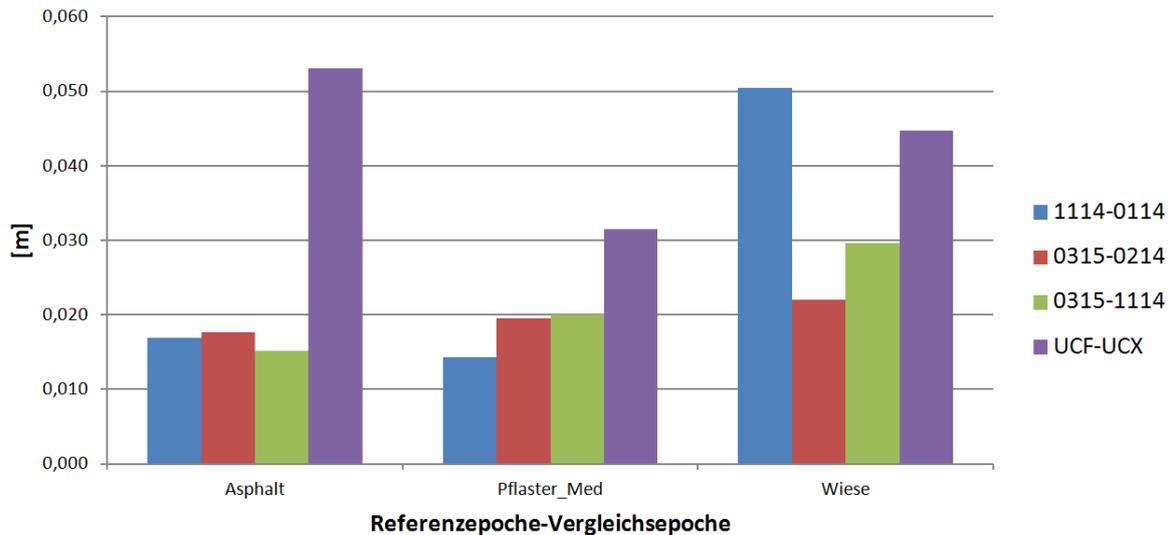


Abb. 16: Zusammenstellung der Höhendifferenzvergleiche für Testflächen und Sensoren

Es hat sich gezeigt, dass die Differenz zwischen Median und Mittelwert der Höhendifferenzen ein gutes Mittel zur Detektion von Verdachtsbereichen darstellt.

4 Multitemporales-multisensorales Monitoringkonzept

Die multisensoralen und multitemporalen Untersuchungen bilden die Basis für einen Workflow zur Erfassung von Bodenbewegungen im Altbergbaubereich an Schächten der RAG. Es zeigt sich, wie wichtig grundlegende Koordinatenreferenzsysteme und hochgenaue Passpunktmessungen sowie ein mehrstufiger Ansatz zur Ableitung der Höhendifferenzen sind. Im ersten Auswerteschritt wird eine aus dem aktuellen Bildflug abgeleitete Punktwolke über Testflächen dahingehend analysiert, wie gut die Punktwolke die erfasste Oberfläche repräsentiert und mit welchem Punktrauschen bzw. welchen Messungenauigkeiten die unterschiedlichen Oberflächen (Asphalt, Wiese, etc.) erfasst wurden. Mit diesen Parametern können im zweiten Schritt die absoluten Differenzen an den Schächten ermittelt und bewertet werden. In Form eines Protokolls als PDF Datei können so die Differenzen und daraus abgeleiteten Ergebnisse dargestellt werden (siehe Abb. 17). Die zusätzliche Darstellung von Orthophotos und Reliefdarstellungen der beiden Zeitschnitte kann bei der Bewertung von Veränderungen helfen. Eine Integration der Ergebnisse in die Schachtdatenbank der RAG und damit auch in die Digitale Service-Akte (D S A) der RAG (VOSEN ET AL. 2014) ermöglicht so einen firmenweiten Zugang und die schnelle Bearbeitung im Falle einer Veränderung, bzw. eines Tagesbruchs.

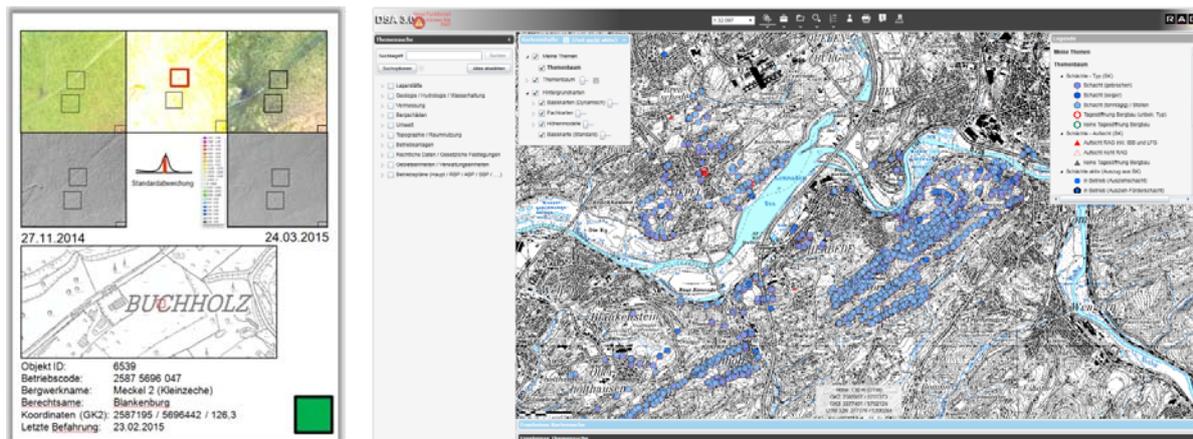


Abb. 17: Protokoll als PDF (links), Digitale Service-Akte der RAG mit dargestellten Schächten (rechts)

5 Ausblick

In 2015 und 2016 wurden und werden anhand einer Fläche von 33 km² erste Erfahrungen mit dem vorgestellten Workflow gesammelt. Ziel ist es, die jährlichen Begehungen von Tagesbrüchen auf ein Minimum zu reduzieren und durch einen photogrammetrischen Workflow weitgehend zu ersetzen. Die Untersuchungsfläche wird ab 2017 auf 230 km² steigen. Bei dieser Flächengröße wird von einem jährlichen Datenvolumen von 100 Terabyte ausgegangen. Das stellt die RAG neben den inhaltlichen Herausforderungen vor ganz neue Herausforderungen, wie z.B. an die Hard- und Software, die Datenspeicherung, das Datenhandling und die Netzwerkumgebung.

6 Literaturverzeichnis

- ABSMON: RAG Forschungsprojekt FE 0760 „Monitoring des Altbergbaubereichs über a) hochpräzise GPS- und satellitengestützte Referenznetze unter Berücksichtigung der Risikoanalyse des untertägigen Grubenwasseranstieges und b) der Entwicklung eines Echtzeit-Schachtinspektionssystems (ABSMon)“. Forschungsprojekt der RAG, 06-2012 bis 07-2015, unveröffentlicht.
- EBERHARDT, L., 2015: Untersuchungen multitemporaler und multisensoraler 3D-Punktwolken für Altbergbaubereiche der RAG. Diplomarbeit an der Fakultät Geowissenschaften, Geotechnik und Bergbau der Technischen Universität Bergakademie Freiberg.
- HAGER, S. & WOLLNIK, F., 2014: Monitoring im Altbergbau der RAG Aktiengesellschaft. 14. Altbergbau-Kolloquium, G. Meier (Hrsg.), 06. bis 08. November 2014, Gelsenkirchen, Technische Fachhochschule Georg Agricola (TFH) Bochum, 349-357.
- GILBLOM, D. L., YOO, S., K. & VENTURA, P., 2003: Operation and performance of a color image sensor with layered photodiodes. SPIE AeroSense 2003 - April 22, 2003 - Orlando, Florida, USA.
- GREIWE, A. & GEHRKE, R., 2013: Foveon Chip oder Bayer Pattern - geeignete Sensoren zur Aerophotogrammetrie mit UAS. Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage 2013, Luhmann, T. & Müller, C. (Hrsg.), Wichmann Verlag, 334-343.

- GPS-RICHTLINIEN, 2002: Richtlinien zum Einsatz von satellitengeodätischen Verfahren im Vermessungspunktfeld - GPS - Richtlinien - vom 02.09.2002, in der Fassung vom 15.04.2003 36.3 – 7132, Stand v. 20.09.2007 37-51.12.01 - 7132. Herausgeber: Innenministerium des Landes Nordrhein-Westfalen, Herstellung und Vertrieb: Landesvermessungsamt Nordrhein-Westfalen 2002. http://www.katastermodernisierung.nrw.de/documente/broschuerenerlasse/GPS_Richtlinie_n_%2020_09_07.pdf (abgerufen am 29.04.2016).
- LUHMANN, T., 2010: Nahbereichsphotogrammetrie. Grundlagen, Methoden und Anwendungen. 3., völlig neu bearb. und erw. Aufl. Berlin: Wichmann. ISBN-13: 978-3879074792
- SPRECKELS, V., 2003: Einführung fester Transformationsparameter vom ETRS'89 zum Gauß-Krüger-Koordinatensystem über NWREF-Punkte zur einheitlichen Führung der Koordinaten und GIS-Datenbestände der Deutschen Steinkohle AG (DSK AG). Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 12, 23. Wissenschaftlich-Technische Jahrestagung der DGPF, FH Bochum, 9.-11. September 2003, 145-156.
- STUDIE 193, 2013: RAG-Studie 193. Einsatz von UAV und Sensoren zur Geländeerfassung, Bericht. Hg. v. Ansgar Greiwe, Fachhochschule Frankfurt/Main. 23.08.2013. unveröffentlicht.
- STUDIE 196, 2013: RAG-Studie 196. CIR-UAS Datenerhebung mit UAV und CIR-Kameras, Bericht Hg. v. Ansgar Greiwe, Fachhochschule Frankfurt/Main. 29.11.2013, unveröffentlicht.
- STUDIE 200, 2014: RAG-Studie 200. MultiCam - Geländeerfassung durch einen Mehrkamera-Sensorkopf. 4 Berichte. Hg. v. Ansgar Greiwe. Steinbeis Transferzentrum Geoinformatik. Nordkirchen. 05.12.2014, unveröffentlicht.
- VERHOEVEN, G.J.J., 2010: It's all about the format – unleashing the power of RAW aerial photography. *International Journal of Remote Sensing* **31** (8), 2009-2042.
- VOSEN, P., KOSLOWSKI, T., FISCHER, P. & MUSIEDLAK, J., 2014: DSA 2.0 – Erweiterung der Integrations- und Informationsplattform der RAG für Geodaten. Schriftenreihe des Institutes für Markscheidewesen und Geodäsie an der Technischen Universität Bergakademie Freiberg, Heft 2014-1. 15. Geokinematischer Tag, 15.-16. Mai 2014, Freiberg, 156-167.

Ein systematischer Vergleich zwischen direkter und bildbasierter Georeferenzierung von Mobile Mapping-Stereosequenzen in einem anspruchsvollen Stadtgebiet

STEFAN CAVEGN^{1,2}, STEPHAN NEBIKER¹ & NORBERT HAALA²

Mobile Mapping Systeme erlauben eine effiziente Erfassung von Multi-View Bildsequenzen. Um genaue Messungen in einem gegebenen Referenzrahmen durchführen zu können, ist aber eine Georeferenzierung höchster Qualität dieser Bildsequenzen gefordert. Ausserdem sind Sub-Pixel genaue relative Orientierungen dieser hochredundanten Sequenzen für eine dichte Bildkorrelation als Basis für die Generierung von 3D Punktwolken und Vermaschungen notwendig. Während die direkte Georeferenzierung von Mobile Mapping Daten in offenen Bereichen zu guten Resultaten führt, erschwert eine unzureichende GNSS Abdeckung in Strassenschluchten auch bei Verwendung von hochwertigen Inertial-Navigations-Sensoren die Erfüllung der hochgenauen Anforderungen. Somit wurden systematische Untersuchungen vorgenommen, um die Qualität der direkt bestimmten Sensororientierungen wie auch die erwartete Verbesserung durch bildbasierte Georeferenzierung in einem anspruchsvollen Stadtgebiet mit regelmässigen GNSS Ausfällen zu ermitteln. Die Untersuchungen lieferten wiederholt mittlere Trajektorienunterschiede von bis zu 80 cm. Durch bildbasierte Georeferenzierung mit Bündelausgleichung sowie einer limitierten Anzahl Kameras und Passpunkte konnten mittlere Kontrollpunkt-Residuen von ca. 40 cm auf 4 cm vermindert werden. Auch konnte aufgezeigt werden, dass die bildbasierte Georeferenzierung in der Lage ist, Unstetigkeiten von direkt georeferenzierten Trajektorien zu detektieren und zu kompensieren.

1 Einleitung

Bildbasiertes Mobile Mapping hat sich in den letzten Jahren zu einer sehr effizienten Kartierungstechnologie entwickelt. Im Gegensatz zu luftbildgestützten Anwendungen, wo die Bodenauflösung über das ganze Kartierungsgebiet konstant bleibt, treten bei fahrzeuggestütztem Mobile Mapping grosse Tiefenunterschiede auf. Im vorliegenden Fall entspricht ein Pixel 2-6 mm im Objektraum für einen gewöhnlichen Messbereich von 4-14 m und ist 1 cm bei einer Distanz von 23 m. Während typische Anwendungen der Infrastrukturverwaltung häufig 3D Messgenauigkeiten besser als 10 cm voraussetzen, werden für städtische Kartierungsarbeiten absolute Genauigkeiten im Zentimeterbereich gefordert. Diese Anforderungen konnten in offenen Gebieten bereits in früheren Arbeiten erfüllt werden. BURKHARD et al. (2012) erreichten mit ihrem stereobildbasierten Mobile Mapping System bei mittleren bis guten GNSS Bedingungen absolute 3D Punktmessgenauigkeiten von 4-5 cm. Dass das StreetMapper LiDAR Mobile Mapping System unter guten GNSS Bedingungen 3D Messungen im Bereich von 3 cm ermöglichen kann, wurde von HAALA et al. (2008) demonstriert. Eine andere Version dieses

¹ Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: [stefan.cavegn, stephan.nebiker]@fhnw.ch

² Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, E-Mail: norbert.haala@ifp.uni-stuttgart.de

Systems ist in PUENTE et al. (2013) vorgestellt, die einen Vergleich der Leistungsfähigkeit verschiedener mobiler Laserscanning Systeme präsentieren. Die GNSS Bedingungen der strassenbasierten Mobile Mapping Systeme werden aber häufig durch Signalabschattungen und Mehrwegeeffekte verschlechtert, was die Erfüllung der Genauigkeitsanforderungen durch direkte Georeferenzierung verunmöglicht. Auch betragen die Distanzen zwischen den Kameras und den zu messenden Objekten einige Meter, im Gegensatz zu Luftbildanwendungen mit einigen hundert Metern. Verglichen mit einer Orientierungsungenauigkeit wirken Ungenauigkeiten in der GNSS Positionierung dementsprechend weitaus stärker auf die Endgenauigkeit aus. Da Luftbildkampagnen viel weniger von GNSS Verschlechterungen betroffen sind, stellten NEBIKER et al. (2012) die Fusion von boden- und flugzeugbasierten Bilddaten vor. Erste Untersuchungen ergaben horizontale Genauigkeiten im Bereich von 5 cm, was der Bodenauflösung der Luftbilddaten entspricht, und Höhengenaugigkeiten von ca. 10 cm.

Eines der Hauptmerkmale unseres Mobile Mapping Systems ist das Vorliegen mehrerer Kameras, die durch Multi-View Stereo Matching für eine dichte 3D Datenerfassung verwendet werden können (CAVEGN et al. 2015). Hierfür werden aber genaue relative Orientierungen der Bildsequenzen vorausgesetzt, möglichst im Sub-Pixel Bereich. Die Ausnutzung der gut co-registrierten hochredundanten Multi-View Bildsequenzen kann dann zu einer Steigerung der Genauigkeit, Zuverlässigkeit und Vollständigkeit resultierender Produkte wie Tiefenkarten, 3D Punktwolken und Vermaschungen führen.

Eine integrierte Georeferenzierung ist in Stadtgebieten mit ausgedehnten Bereichen schlechter GNSS Abdeckung unerlässlich, besonders wenn Daten an unterschiedlichen Tagen und Tageszeiten, was typisch bei der Kartierung ganzer Städte ist, kombiniert werden sollen (NEBIKER et al. 2015). Im Gegensatz zur direkten Georeferenzierung, die ausschliesslich auf Positions- und Orientierungsdaten vom GNSS/INS basiert, verwendet die integrierte Georeferenzierung häufig vorberechnete Kamerapositionen in der Bündelausgleichung. ELLUM & EL-SHEIMY (2006) schlagen vor, Koordinatenupdates (CUPTs) aus einer photogrammetrischen Bündelausgleichung in einer zweiten Iteration im Loosely-Coupled GNSS Kalman Filter einzuführen. Dieser Ansatz wurde später von EUGSTER et al. (2012) verwendet, womit sie die absolute 3D Messgenauigkeit für strassenbasiertes Mobile Mapping von einigen Dezimetern auf 5-10 cm reduzieren konnten. HASSAN et al. (2006) führen eine Bündelausgleichung unter Verwendung von Kamerapositionen und -orientierungen aus einem Kalman Filter durch. In schlechten GNSS Bereichen reduzieren die Gewichte dieser äusseren Orientierungsparameter gegen Null, was dazu führt, dass nur noch Bildbeobachtungen berücksichtigt werden. Ähnliche Ansätze des Photogrammetrischen Bridging wurden auch von FORLANI et al. (2005) und von SILVA et al. (2014) entwickelt, um strassenbasierte Mobile Mapping Stereobildsequenzen in Gebieten ohne GNSS Empfang zu überbrücken.

2 Mobile Mapping System und Messfahrten

Um Genauigkeitsuntersuchungen zur Georeferenzierung vornehmen zu können, wurden zwei Mobile Mapping Kampagnen mit unterschiedlichen Sensorkonstellationen durchgeführt. Diese Messfahrten wie auch das Untersuchungsgebiet werden in den folgenden zwei Abschnitten vorgestellt.

2.1 Mobile Mapping System

Alle Daten für diesen Beitrag wurden mit dem Multi-Sensor Stereobild Mobile Mapping System vom Institut Vermessung und Geoinformation (IVGI) der Fachhochschule Nordwestschweiz (FHNW) erfasst. Obwohl zwei Messfahrten mit unterschiedlichen Sensorkonstellationen durchgeführt wurden, fokussieren die folgenden Untersuchungen auf Stereobilddaten des Hauptsystems nach vorne, das aus zwei 11 MP Kameras und einer kalibrierten Stereobasis von 905 mm besteht. Diese Stereokameras haben je eine Auflösung von 4008 x 2672 Pixel bei einer Pixelgrösse von 9 μm , eine Kamerakonstante von 21 mm und resultierende Sichtfelder von 81° in horizontaler und 60° in vertikaler Richtung. HD Kameras mit einer Auflösung von 1920 x 1080 Pixel, einer Pixelgrösse von 7.4 μm , einer Kamerakonstante von 8 mm und einem Sichtfeld von 83° x 53° komplettieren die Multi-Kamera Konfiguration. Bei der Messfahrt im Juli 2014 wurden zwei zusätzliche Stereobasen mit solchen Kameras mitgeführt, bei der Messfahrt im August 2015 lediglich eine dritte HD Kamera in der Mitte der 11 MP Kameras (siehe Abb. 1). Für die direkte Georeferenzierung wurde ein Inertial-Navigation-System vom Typ NovAtel SPAN verwendet. Bei guter GNSS Verfügbarkeit können mit Nachprozessieren horizontale Genauigkeiten von 10 mm und vertikale von 15 mm erreicht werden (NOVATEL 2016). Genauigkeiten der Orientierungswinkel Roll und Nick sind mit 0.005° spezifiziert, Gier mit 0.008°. Ein GNSS Ausfall von 60 Sekunden vermindert die horizontale Genauigkeit auf 110 mm und die vertikale auf 30 mm.



Abb. 1: Mobile Mapping System des IVGI (links), Sensor-Konfiguration für die Messfahrt Juli 2014 (rechts oben) und für die Messfahrt August 2015 (rechts unten)

2.2 Untersuchungsgebiet und Messfahrten

Das Untersuchungsgebiet in Abb. 2 umfasst eine verkehrsreiche Kreuzung mit fünf Strassen im Stadtzentrum von Basel. Grosse Geschäftsgebäude erschweren die GNSS Verfügbarkeit in den Strassenschluchten. Drei Strassenabschnitte dieses Untersuchungsgebiets wurden dreimal kartiert, einmal im Juli 2014 und zweimal am gleichen Tag im August 2015 (siehe Tab. 1). In allen neun Fällen erfolgte die Datenerfassung kurz vor Mittag und bei guten Wetterbedingungen. Für die Untersuchungen wurden 85 bis 191 Stereobildpaare vom vorwärts gerichteten Stereokamerasystem bei einer Sequenzlänge zwischen 108 m und 217 m verwendet. Es wurde eine Distanz zwischen nachfolgenden Projektionszentren von 1 m beabsichtigt. Bei einer Fahrgeschwindigkeit von mehr als 18 km/h erhöhte sich aber dieser Abstand, da die maximale Bildrate pro Sekunde 5 betrug.

Während die Kampagne im Juli 2014 Teil einer vollständigen Befahrung des Kantons Basel-Stadt war, wurde die Messfahrt im August 2015 spezifisch für unser Untersuchungsgebiet angelegt (siehe Abb. 2). Um optimale Trajektorien zu erzielen, erfolgte die kinematische Datenerfassung nach einem vom Hersteller empfohlenen Verfahren. Somit wurde zuerst eine statische Initialisierung von ca. 3 Minuten in einem offenen Bereich durchgeführt. Es folgte ein Einfahren und die erste Kartierung des Untersuchungsgebiets. Nach einer zusätzlichen Schleife wurde das Untersuchungsgebiet zum zweiten Mal kartiert und nachdem der Startbereich erreicht wurde, wurden Bilddaten vom Kalibrierfeld für die Bestimmung von Hebelarm und Fehlansrichtung erfasst. Eine letzte Schleife diente fürs Ausfahren gefolgt von einer statischen Beobachtung von etwa 4 Minuten in der Nähe des FHNW Gebäudes. Die GNSS Antenne auf dem Dach dieses Gebäudes diente als Basisstation. Die komplette Messfahrt lieferte 12220 Stereobildpaare über eine Länge von 22.756 km und eine Erfassungszeit von etwas mehr als einer Stunde.

Für die Untersuchungen wurden 51 Punkte vorwiegend an Ecken von Strassenmarkierungen mit einer 3D Genauigkeit von besser als 1 cm durch Tachymetrie erfasst. Diese 3D Punkte dienten entweder als Passpunkte (PP) oder als Kontrollpunkte (KP) (siehe Abb. 2 und Abb. 6).

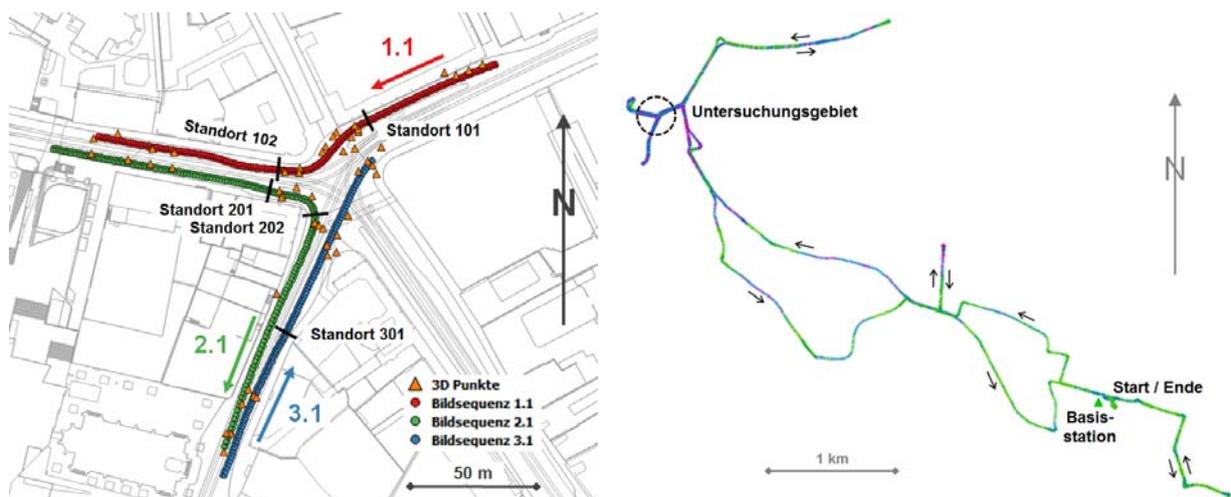


Abb. 2: Grundlagekarte des Untersuchungsgebiets mit überlagerten Projektionszentren ausgewählter Stereobildsequenzen, 3D Referenzpunkte und Standorte der Trajektorienunstetigkeiten (links, Quelle: Geodaten Kanton Basel-Stadt); Trajektorie der Messfahrt August 2015 (rechts)

Tab. 1: Kenndaten der neun gewählten Stereobildsequenzen x.y (x entspricht den Strassenabschnitten 1 bis 3 und y entspricht der Kampagne, 0 = 24./27.7.14, 1 = 20.8.15 10:30-10:37, 2 = 20.8.15 10:47-10:53)

Sequenz	Datum und Uhrzeit	Anzahl Bilder	Länge [m]	Distanz in Fahrtrichtung [m]	
				Mittelwert	Maximum
1.0	24.7.2014 10:20	246	164	1.34	1.97
1.1	20.8.2015 10:30	322	173	1.08	1.25
1.2	20.8.2015 10:47	312	175	1.13	1.48
2.0	27.7.2014 11:53	314	173	1.11	1.60
2.1	20.8.2015 10:34	342	212	1.25	2.06
2.2	20.8.2015 10:50	382	217	1.14	1.93
3.0	27.7.2014 11:57	170	108	1.29	1.49
3.1	20.8.2015 10:37	232	141	1.23	1.73
3.2	20.8.2015 10:53	192	146	1.54	2.37

3 Kalibrierung und Trajektorienprozessierung

Alle Bildsensoren des Mobile Mapping Systems wurden in einem umfangreichen Prozess kalibriert. Dieser beinhaltete zuerst die Bestimmung der inneren wie auch der relativen Orientierungsparameter zwischen allen Kameras. Hierfür erfolgte jeweils eine Bündelausgleichung aufgrund von Bilddaten eines Innenraum-Kalibrierfeldes, das viele codierte Zielmarken aufweist. In einem zweiten Schritt folgte die Bestimmung von Hebelarm und Fehlansrichtung zur linken Kamera des vorwärts gerichteten Stereosystems durch Verwendung von Bildern, die vom Aussenraum-Kalibrierfeld aufgenommen wurden (BURKHARD et al. 2012). Für die Bestimmung der direkten Georeferenzierung wurden die erfassten Navigationsdaten mit der Software Inertial Explorer von NovAtel prozessiert. Dies erfolgte im Tightly-Coupled Modus in mehreren Richtungen gefolgt von einer Glättung der Trajektorien. Abb. 2 kann entnommen werden, dass die Trajektorie im Untersuchungsgebiet nur eine mittlere bis niedrige Qualität aufweist. Unter Berücksichtigung der vorgängig berechneten Fehlansrichtung und Hebelarm wie auch der relativen Orientierungsparameter konnte die direkte Sensororientierung für alle Bilder bestimmt werden.

Die bildbasierte Georeferenzierung durch Bündelausgleichung aller neun Stereobildsequenzen erfolgte mit Agisoft PhotoScan. Dabei wurden die äusseren Orientierungsparameter von der direkten Georeferenzierung wie auch etwa 20 Passpunkte pro Sequenz miteinbezogen. Es resultierten mittlere Gesamt-Rückprojektionsfehler in der Grössenordnung von 0.42-0.89 Pixel. Während Rückprojektionsfehler für Verknüpfungspunkte im Bereich von 0.15-0.21 Pixel liegen, betragen diese 0.81-1.08 Pixel bei den Passpunkten. Werte von etwa 1 Pixel sind aber plausibel, da die Identifikation und die manuelle Messung der Passpunkte nicht trivial sind (siehe Abb. 3). Eine detaillierte Zusammenstellung aller Werte ist in CAVEGN et al. (2016) ersichtlich.



Abb. 3: Ein Beispielbild für alle 3D Punkte der Stereobildsequenz 2.1

4 Genauigkeit von Trajektorien und Kontrollpunkten

Im Folgenden sind die Ergebnisse der systematischen Untersuchungen dokumentiert, die es zum Ziel hatten, die Qualität der direkt bestimmten Sensororientierungen und dessen Verbesserung durch bildbasierte Georeferenzierung in einem anspruchsvollen Stadtgebiet mit regelmässigen GNSS Ausfällen zu beurteilen. Um die Aussagekraft zu erhöhen, wurden zudem Kontrollpunktuntersuchungen für die direkte wie auch für die bildbasierte Georeferenzierung vorgenommen.

4.1 Trajektorien- und Orientierungsdifferenzen zwischen direkter und bildbasierter Georeferenzierung

Differenzen der Projektionszentren und Orientierungswinkel zwischen direkter und bildbasierter Georeferenzierung wurden für alle neun Sequenzen berechnet (siehe Tab. 2). Die 3D Differenzen liegen im Bereich von 46 bis 803 mm und die Höhe ist die Komponente mit den grössten Residuen, ausser bei Sequenz 3.2. Eher kleine Differenzen konnten für Strassenabschnitt 3 ermittelt werden. Sequenz 1.1 weist mit 0.5° die grössten Orientierungsdifferenzen für alle Orientierungswinkel auf. Es wurden aber mittlere Omega-Phi-Kappa Differenzen von 0.3° erreicht.

Tab. 2: Differenzen der Orientierungswinkel und Projektionszentren zwischen direkter und bildbasierter Georeferenzierung

Sequenz	$\Delta\omega$ RMSE [°]	$\Delta\phi$ RMSE [°]	$\Delta\kappa$ RMSE [°]	$\Delta\omega\phi\kappa$ RMSE [°]	ΔO RMSE [mm]	ΔN RMSE [mm]	ΔH RMSE [mm]	$\Delta 3D$ RMSE [mm]
1.0	0.246	0.095	0.246	0.361	298	36	423	519
1.1	0.366	0.072	0.354	0.515	41	86	126	158
1.2	0.311	0.067	0.295	0.434	437	26	572	721
2.0	0.143	0.056	0.160	0.222	52	37	75	98
2.1	0.319	0.081	0.298	0.445	264	83	498	569
2.2	0.310	0.064	0.283	0.425	172	470	628	803
3.0	0.011	0.055	0.008	0.057	23	30	77	86
3.1	0.088	0.079	0.047	0.127	15	25	36	46
3.2	0.084	0.057	0.048	0.112	172	500	146	548
Mittel	0.209	0.070	0.193	0.300	164	144	287	394

Mögliche Verbesserungen der Projektionszentren-Differenzen und somit der Trajektoriengenauigkeit aus integrierter und bildbasierter Georeferenzierung im Vergleich zur direkten Georeferenzierung sind in Abb. 4 detailliert dargestellt. Trajektorien von Stereobildsequenzen der gleichen Strassenabschnitte weisen Differenzen von bis zu mehreren Dezimetern auf. Für Sequenzen 1.1 und 2.0 resultierten kleine Abweichungen. Diese sind aber signifikant grösser für die anderen Sequenzen dieser beiden Strassenabschnitte. Alle Differenzen von Strassenabschnitt 3 sind kleiner als 10 cm, mit Ausnahme der Nordkomponente von Sequenz 3.2 mit einem Wert von ca. 50 cm.

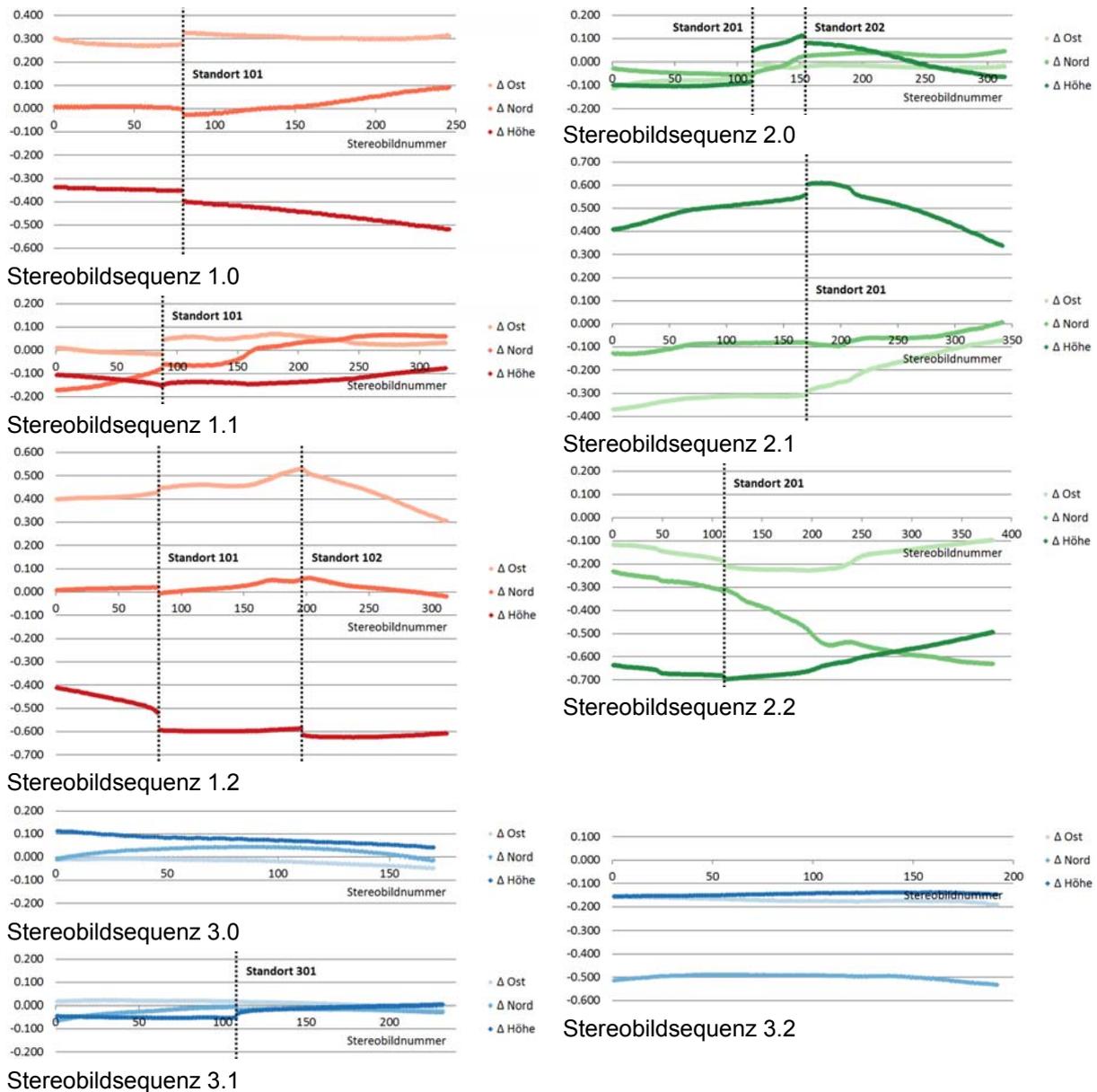


Abb. 4: Differenzen der Projektionszentren in m zwischen direkter und bildbasierter Georeferenzierung

4.2 Trajektorienunstetigkeiten aus direkter Georeferenzierung

Die Diagramme mit den Trajektorienunterschieden zwischen direkter und bildbasierter Georeferenzierung zeigen neun Trajektorienunstetigkeiten. Diese sind jeweils mit einer gepunkteten vertikalen Linie gekennzeichnet (siehe Abb. 4). Gemäss Abb. 2 und Abb. 5 wurden alle ausser einer (Standort 202) durch einen Fahrzeug-Stopp von mehreren Sekunden verursacht, meistens vor Fussgängerstreifen. Es konnte aber keine Korrelation zwischen Stopp-Dauer und 3D Wert der Unstetigkeiten festgestellt werden (siehe Tab. 3). Die 3D Unstetigkeiten betragen grösstenteils einige Zentimeter, reichen aber bis zu ca. 15 cm bei Sequenz 2.0 und Standort 201.

Sequenzen 3.0 und 3.2 weisen keine Unstetigkeiten auf, wobei das Fahrzeug bei diesen auch nicht anhalten musste.

Tab. 3: Ausmasse der Trajektorienunstetigkeiten

Standort	Sequenz	Stopp [s]	ΔO [mm]	ΔN [mm]	ΔH [mm]	$\Delta 3D$ [mm]
101	1.0	17	48	-21	-43	68
101	1.1	13	62	19	4	65
101	1.2	31	17	-27	-76	82
102	1.2	68	-3	2	-29	29
201	2.0	19	57	4	132	144
201	2.1	56	17	-3	46	49
201	2.2	41	-16	16	-12	26
202	2.0	0	-5	0	-30	30
301	3.1	16	-3	-12	18	22



Standort 101, Sequenz 1.0



Standort 101, Sequenz 1.1



Standort 101, Sequenz 1.2



Standort 102, Sequenz 1.2



Standort 201, Sequenz 2.0



Standort 201, Sequenz 2.1



Standort 201, Sequenz 2.2



Standort 202, Sequenz 2.0



Standort 301, Sequenz 3.1

Abb. 5: Mobile Mapping Bilder die an Standorte der Trajektorienunstetigkeiten erfasst wurden

4.3 Untersuchungen mit Kontrollpunkten für die direkte und bildbasierte Georeferenzierung

In Kapitel 4.1 wurden die 3D Koordinaten der Kameratrajektorien aus direkter Georeferenzierung und Bündelausgleichung verglichen. Obwohl dies einige Hinweise auf die entsprechende Qualität gibt, sind Genauigkeitsuntersuchungen der 3D Koordinaten von gemessenen Bildpunkten aussagekräftiger. Für die Berechnung dieser 3D Koordinaten durch

Vorwärtseinschnitt können Orientierungsparameter aus direkter Georeferenzierung und Bündelausgleichung verwendet werden. Somit wurden Gruppen bestehend aus zwei, drei oder vier Passpunkten (PP) angelegt und etwa die Hälfte der vorgängig verwendeten Passpunkte als Kontrollpunkte (KP) definiert (siehe Abb. 6). Danach wurden Kontrollpunkt-Residuen für zwei Szenarien berechnet. Zuerst für jeweils eine Passpunktgruppe an den Enden der Segmente und danach mit zwei zusätzlichen Passpunktgruppen nahe der entsprechenden scharfen Kurve. Für Szenario 1 mit zwei Passpunktgruppen wurden mittlere Kontrollpunkt-Residuen von ca. 15 cm ermittelt. Dies ist etwa dreimal besser als ein Wert von ca. 40 cm für die direkte Georeferenzierung (siehe Tab. 4). Szenario 2 mit vier Passpunktgruppen lieferte Kontrollpunkt-Residuen im Bereich von 21-73 mm, was einer Verbesserung um eine Zehnerpotenz im Vergleich zur direkten Georeferenzierung entspricht.

Tab. 4: Kontrollpunkt-Residuen (RMSE) der direkten und bildbasierten Georeferenzierung

Sequenz	Anzahl KP	Direkt	Bildbasiert 2 PP-Gruppen		Bildbasiert 4 PP-Gruppen	
		$\Delta 3D$ [mm]	$\Delta 3D$ [mm]	Verb.-Faktor	$\Delta 3D$ [mm]	Verb.-Faktor
1.0	15	555	137	4.1	27	20.4
1.1	11	168	42	4.0	21	8.0
1.2	11	774	121	6.4	26	29.4
2.0	11	131	76	1.7	48	2.7
2.1	12	593	432	1.4	73	8.1
2.2	11	813	425	1.9	36	22.5
3.0	8	174	42	4.2		
3.1	10	64	30	2.1		
3.2	10	568	53	10.8		
Mittel		427	151	2.8	39	11.1

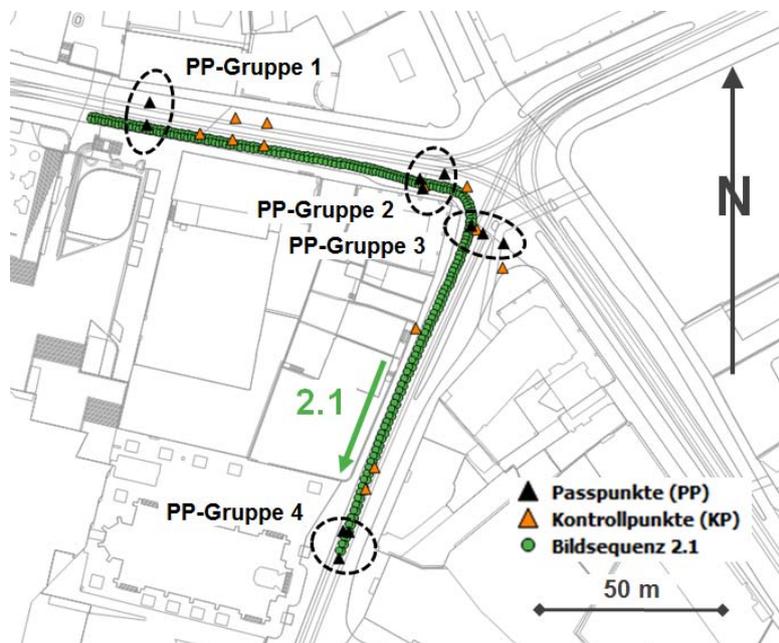


Abb. 6: Lage der Passpunktgruppen und Kontrollpunkte bei Stereobildsequenz 2.1 (Quelle: Geodaten Kanton Basel-Stadt)

5 Fazit und Ausblick

Umfassende Untersuchungen anhand von qualitativ hochstehender Referenz- und Bilddaten führten bei der direkten Georeferenzierung wiederholt zu 3D Trajektoriengenauigkeiten im Bereich von einem bis mehreren Dezimetern. Durch bildbasierte Georeferenzierung mit Bündelausgleichung und einigen Passpunkten konnten Objektpunktgenauigkeiten von ca. 4 cm konsequent erreicht werden, was einer Verbesserung um eine Zehnerpotenz entspricht. Da Bildbeobachtungen über die ganzen Sequenzen berücksichtigt wurden, konnten nicht nur grosse Offset- und Drift-Fehler der direkten Georeferenzierung eliminiert, sondern auch Trajektorienunstetigkeiten detektiert und kompensiert werden.

Die erreichten absoluten 3D Punktgenauigkeiten im Bereich von einigen Zentimetern erfüllen die Anforderungen typischer Anwendungen der Infrastrukturverwaltung. Die 3D Messungen basierten aber nur auf Bilddaten des vorwärts gerichteten Stereosystems und eine beträchtliche Anzahl Passpunkte ist noch immer notwendig. Effizienz und Genauigkeit werden künftig erhöht, indem Multi-View Stereobildsequenzen in der Bündelausgleichung berücksichtigt und zusätzlich Bedingungen für die kalibrierten Offsets und Rotationen zwischen den entsprechenden Kameras eingeführt werden. Die erreichten Sub-Pixel genauen relativen Orientierungen sind ausreichend, um eine dichte Bildkorrelation durchführen zu können. Die völlig unterschiedlichen Blickrichtungen erschweren jedoch die automatische Verknüpfungspunkt-Detektion und Zuordnung in Mobile Mapping Sequenzen. Dies wie auch die Entwicklung einer integrierten und bildbasierten Georeferenzierungslösung, um Bildsequenzen ganzer Städte zu prozessieren, werden bei künftigen Forschungsaktivitäten angegangen.

Anerkennung: Wir möchten uns bei der Firma iNovitas AG (Baden-Dättwil, Schweiz) und dem Kanton Basel-Stadt (Bau- und Verkehrsdepartement) für die Bereitstellung der Mobile Mapping-Daten von der Messfahrt Juli 2014 bedanken. Diese Arbeit wurde von der Kommission für Technologie und Innovation KTI im Rahmen des infraVIS-Projekts mitfinanziert.

6 Literaturverzeichnis

- BURKHARD, J., CAVEGN, S., BARMETTLER, A. & NEBIKER, S., 2012: Stereovision Mobile Mapping: System Design and Performance Evaluation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39**, Part B5, 453-458.
- CAVEGN, S., HAALA, N., NEBIKER, S., ROTHERMEL, M. & ZWÖLFER, T., 2015: Evaluation of Matching Strategies for Image-Based Mobile Mapping. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2** (3/W5), 361-368.
- CAVEGN, S., NEBIKER, S. & HAALA, N., 2016: A Systematic Comparison of Direct and Image-Based Georeferencing in Challenging Urban Areas. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41**, ICWG I/Va.
- ELUM, C. M. & EL-SHEIMY, N., 2006: New Strategies for Integrating Photogrammetric and GNSS Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**, Part 5.

- EUGSTER, H., HUBER, F., NEBIKER, S. & GISI, A., 2012: Integrated Georeferencing of Stereo Image Sequences Captured with a Stereovision Mobile Mapping System – Approaches and Practical Results. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39**, Part B1, 309-314.
- FORLANI, G., RONCELLA, R. & REMONDINO, F., 2005: Structure and Motion Reconstruction of Short Mobile Mapping Image Sequences. VII Conference on Optical 3D Measurement Techniques, Vienna, Austria, 265-274.
- HAALA, N., PETER, M., KREMER, J. & HUNTER, G., 2008: Mobile LiDAR Mapping for 3D Point Cloud Collection in Urban Areas - A Performance Test. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **37**, Part B5, 1119-1124.
- HASSAN, T., ELLUM, C. & EL-SHEIMY, N., 2006: Bridging Land-Based Mobile Mapping Using Photogrammetric Adjustments. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**, Part 1.
- NEBIKER, S., CAVEGN, S., EUGSTER, H., LAEMMER, K., MARKRAM, J. & WAGNER, R., 2012: Fusion of Airborne and Terrestrial Image-Based 3D Modelling for Road Infrastructure Management – Vision and First Experiments. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39**, Part B4, 79-84.
- NEBIKER, S., CAVEGN, S. & LOESCH, B., 2015: Cloud-Based Geospatial 3D Image Spaces – A Powerful Urban Model for the Smart City. *ISPRS International Journal of Geo-Information* **4** (4), 2267-2291.
- NOVATEL, 2016: SPAN UIMU-LCI. <http://www.novatel.com/assets/Documents/Papers/IMU-LCI.pdf> (28.4.2016).
- PUENTE, I., GONZÁLEZ-JORGE, H., MARTÍNEZ-SÁNCHEZ, J. & ARIAS, P., 2013: Review of Mobile Mapping and Surveying Technologies. *Measurement* **46** (7), 2127-2145.
- SILVA, J.F.C., LEMES NETO, M.C. & BLASECHI, V., 2014: Automating the Photogrammetric Bridging Based on MMS Image Sequence Processing. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (1), 389-396.

Automated Point Cloud Analysis for 3D Topographic Change Detection

THOMAS BAHR¹, NICOLAI HOLZER¹ & REBECCA LASICA²

Zusammenfassung: Eine Hangrutschung am 30. Juli 2014 führte zu großflächigen Massenbewegungen, welche Teile des Dorfes Malin in Maharashtra in Indien begruben. In dieser vorläufigen Studie wurde durch Differenzbildung zweier multi-temporalen Geländemodelle Oberflächenänderungen von bis zu 10 m als Folge der Hangrutschung bestimmt. Die ursprüngliche Geländesituation repräsentiert das Airbus DS WorldDEM™ Produkt. Ein aus Pléiades Stereo-Satellitendaten abgeleitetes Geländemodell beschreibt die Situation nach dem Ereignis. Die operationelle Umsetzung der Analytik basiert auf ENVI 5.3 und dem in IDL entwickelten, objektorientierten ENVITask-System. Diese modulare und erweiterungsfähige Implementierung zeigt die Potentiale einer vollautomatisierten Integration, z. B. in webgestützte, servicebasierte Umgebungen wie ArcGIS Server®.

1 Introduction

The fusion of stereo photogrammetric point clouds with LiDAR data or terrain information derived from SAR interferometry has significant potentials for 3D topographic change detection. In the present, preliminary case study latest point cloud generation and analysis capabilities are used to examine a landslide that occurred in the village of Malin in Maharashtra, India, on 30 July 2014, and affected an area of approximately 44.000 m². The landslide was triggered by heavy rainfall in the area (ERING et al. 2015; RAMASAMY et al. 2015).

The point cloud analysis focuses on Pléiades high resolution satellite imagery and the Airbus DS WorldDEM™ as a product of the TanDEM-X mission. The pre-event topography is represented by the WorldDEM™ product (called pre-DSM, Digital Surface Model), delivered with a raster of 12 m x 12 m and based on the EGM2008 geoid. For the post-event situation a Pléiades 1A/1B stereo image pair was obtained (Figure 1) with a stereo aperture of 19.7°, containing rational polynomial coefficients (RPCs). The panchromatic bands have a ground sample distance (GSD) of 0.5 m and serve for the generation of a photogrammetric point cloud, from which a so-called post-DSM is derived.

For the operation of the various processing steps in service-based production environments an automated workflow was designed. It can be embedded in virtually any existing geospatial workflow for operational applications.

This case study was performed using the Commercial of the Shelf (COTS) software package ENVI 5.3. Integration of custom processes and automation is supported by IDL (Interactive Data Language). Thus, ENVI analytics is running via the object-oriented and IDL-based ENVITask API.

¹ Harris Corporation, Talhofstraße 32a, D-82205 Gilching,
E-Mail: [Thomas.Bahr, Nicolai.Holzer]@harris.com

² Harris Corporation, 4990 Pearl East Circle, Boulder, Colorado 80301, U.S.A.,
E-Mail: rebecca.lasica@harris.com



Fig. 1: Pléiades panchromatic stereo pair of the Malin landslide area, used to run the photogrammetric generation of a point cloud from optical imagery (subset, GSD 0.5 m).
 Left: Pléiades 1A, acquired 8 Feb 2015, incidence angle 21.6°.
 Right: Pléiades 1B, acquired 9 Feb 2015, incidence angle 1.9°.

2 3D Topographic Change Detection

2.1 Key Components

For the application of the 3D topographic change detection we compiled the following processing steps:

- Photogrammetric extraction of a passive point cloud from the panchromatic Pléiades stereo dataset (Figure 2). For details on the algorithm see section 3.2. The point cloud is generated with one 3D point per pixel, resulting in a point density of 3-4 points / m², according to the 0.5 m GSD of the panchromatic bands.
- Derivation of a post-event digital surface model from the photogrammetric point clouds (called post-DSM) with a raster of 1 m x 1 m, referred to the ellipsoid. For details on the algorithm see section 3.2.
- Adding the geoid component (EGM 2008) to the previously generated post-DSM to get the real geoidal height, as the Airbus DS WorldDEM™ (pre-DSM) is referred to this geoid (RIEGLER et al. 2015).
- Pre-DSM reprojection to a common coordinate system (UTM Zone 43N, WGS-84) and resizing to the 1 m x 1 m grid size of the post-DSM.
- Subtraction of the pre-DSM from the post-DSM.
- Filtering of the DSM difference with a convolution, as the resolution difference between the two elevation datasets has caused some anomalies in the resulting raster grid.
- Creation of a classification raster by thresholding on selected data ranges of the DSM difference (Figures 3, 4).

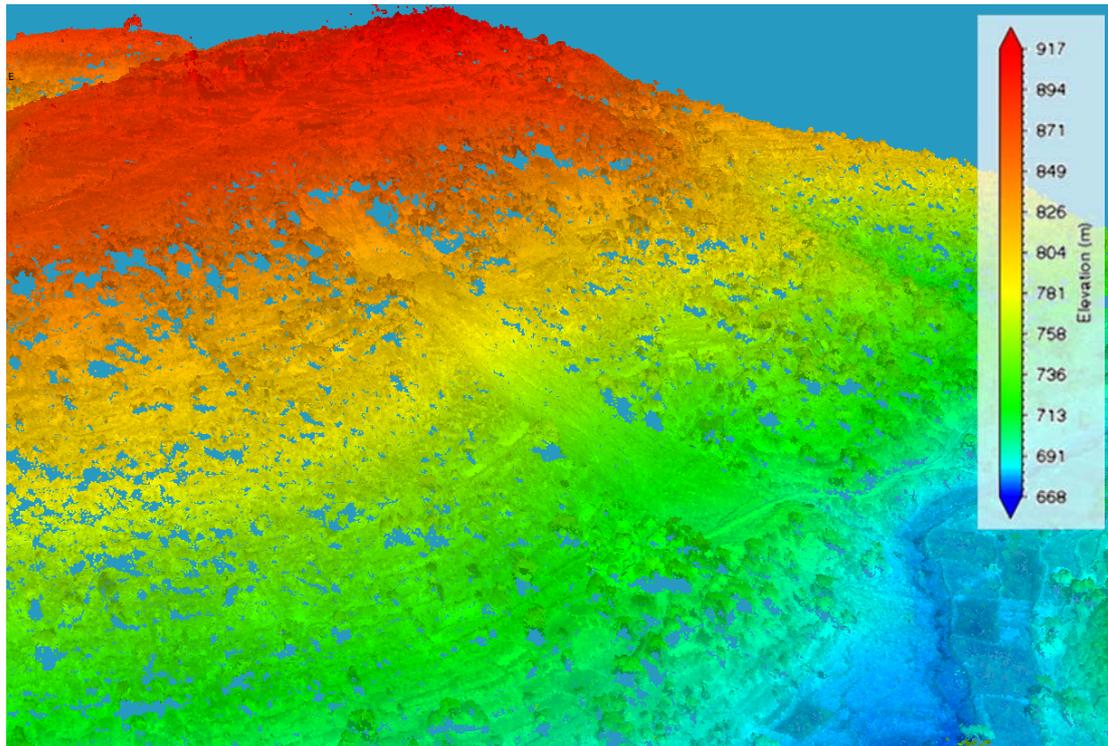


Fig. 2: Point cloud, composed of 158 *.las files, extracted from the panchromatic stereo datasets (subset). The Malin landslide with a length of 514 m is located in the center.

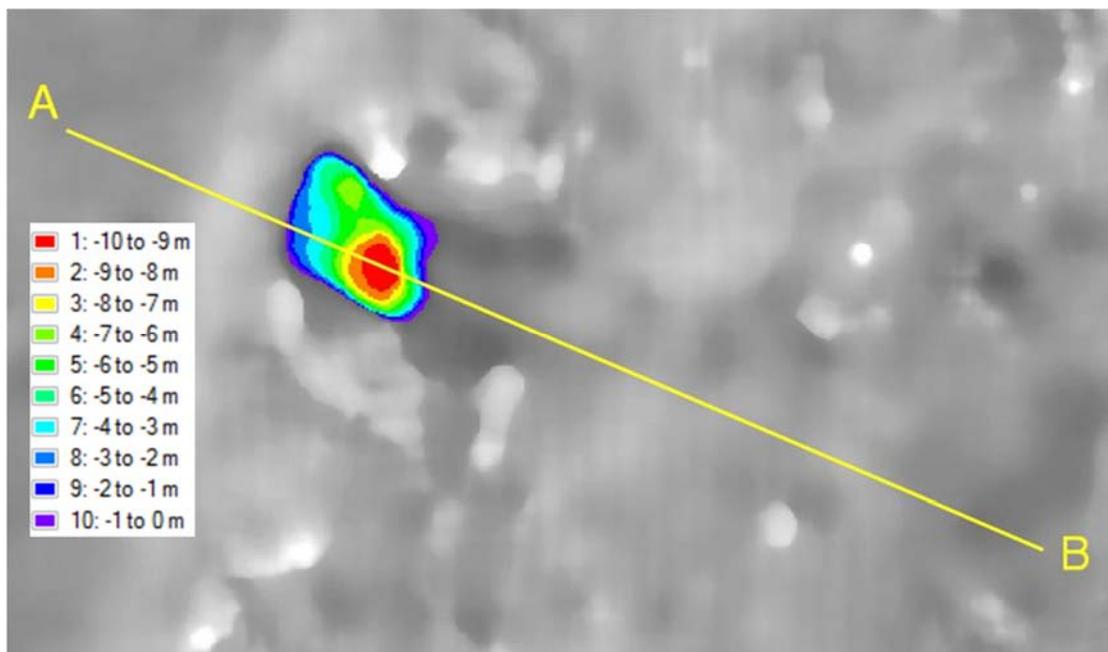


Fig. 3: DSM difference (post-DSM minus pre-DSM) with threshold based classification of the DSM difference (subset), scaled to a range from 0 to -10 m.

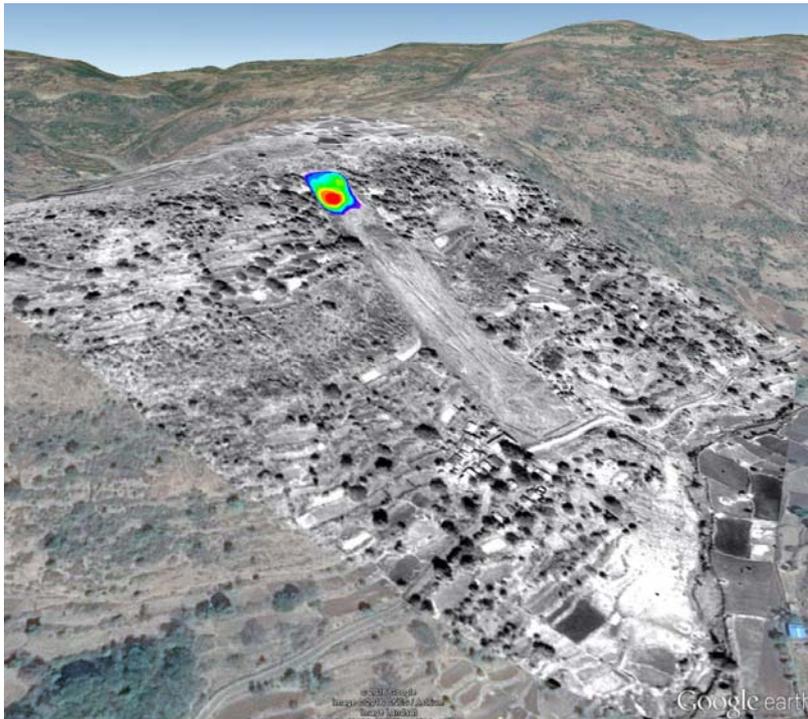


Fig. 4: 3D view of the Malin landslide area with point cloud density image and threshold based classification of the DSM difference overlaid.

2.2 Result Evaluation

ERING et al. 2015 examined the landslide using the finite difference program FLAC 2D (Fast Lagrangian analysis of continua). As a result, Figure 5 (right) shows the displacement contours [dm] of the slope. Accordingly, ERING et al. 2015 calculated a maximum slip surface thickness of 10 m and were able to confirm that by ground truth. This magnitude corresponds to our threshold based classification of the DSM difference, visualized in Figures 3 and 4 with a range from 0 m to -10 m. The pre-DSM and post-DSM transects in Figure 5 (left) clearly reveal the erosion within the slip zone. The analysis of ERING et al. 2015 demonstrates that the slope failure occurred due to the loss of suction strength at the interface between rock and local soil.

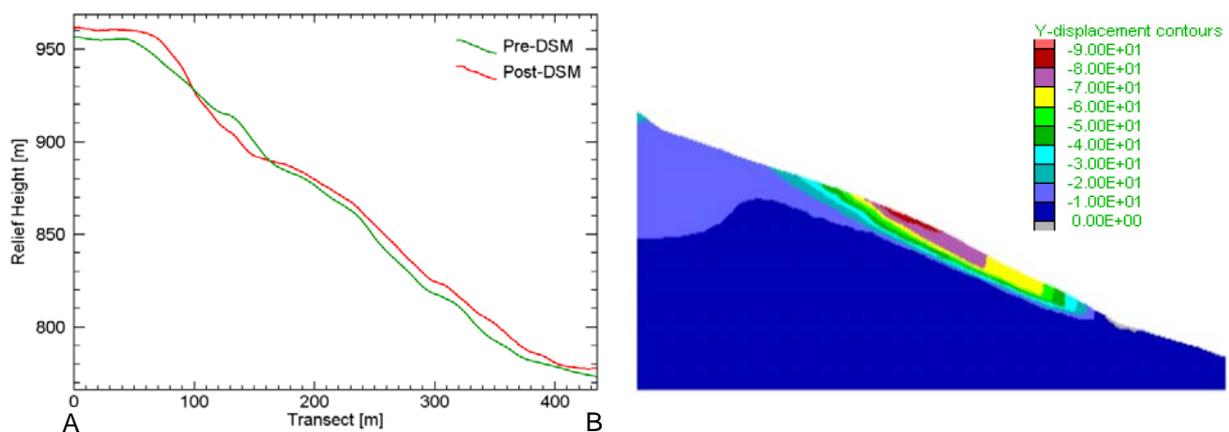


Fig. 5: Left: Terrain profiles along the transect A-B (see Figure 3) reveal the slip surface of the Malin landslide. Right: Modeled displacement contours of the slope at the time of failure after ERING et al. 2015. Maximum thickness of the slip surface is approximately 10 m.

3 Automated Change Detection Workflow

3.1 IDL-Based Task System

For the operation of the processing steps (as described in section 2) in service-based production environments we embedded them in an IDL-based Task System. IDL (Interactive Data Language) as the development language of ENVI allows to extend or customize ENVI features and functionality, and to integrate them in existing geospatial workflows.

Based on IDL, ENVI introduced a so-called ENVITask API for processing data. ENVITasks are an object-oriented approach to run ENVI analytics programmatically. ENVITasks are objects with properties and methods. After instantiation, their properties can be manipulated before executing them. An example for predefined ENVITasks is given in Figure 6.

```

PRO generatePointClouds, inStereo1, inStereo2, inDEM, outDIR
  COMPILE_OPT IDL2
  ; Start the application
  e = ENVI(/HEADLESS)
  ; Open input rasters
  rasterStereo1 = e.OpenRaster(inStereo1)
  rasterStereo2 = e.OpenRaster(inStereo2)
  rasterDEM     = e.OpenRaster(inDEM)
  ; Get the point cloud generation task from the catalog of ENVITasks
  Task = ENVITask('GeneratePointCloudsByDenseImageMatching')
  Task.INPUT_RASTERS      = [rasterStereo1, rasterStereo2]
  Task.INPUT_DEM_RASTER  = rasterDEM
  Task.DO_BLOCK_ADJUSTMENT = 55.00 ; default
  Task.MINIMUM_OVERLAP   = 55     ; default
  Task.MATCHING_THRESHOLD = 15     ; default
  Task.EDGE_THRESHOLD    = 5      ; default
  Task.QUALITY_THRESHOLD  = 60     ; default
  Task.OUTPUT_DIRECTORY  = outDIR
  Task.Execute
END

```

Fig. 6: Usage of the corresponding ENVITask to automate the photogrammetric generation of a point cloud from optical stereo imagery.

3.2 Change Detection Workflow in the Task System

The workflow introduced in section 2 contains the following predefined ENVITasks:

The “GeneratePointCloudsByDenseImageMatching” task was implemented to extract passive point clouds in LAS file format from the panchromatic stereo datasets:

- A dense image-matching algorithm is used to identify corresponding points in the stereo pair. For a given point in one image, it searches a two-dimensional grid of points in the second image. By having orientation data, the search is reduced to one dimension: along an epipolar line in the second image.
- Additionally, the WorldDEM™ was input to constrain the range of heights in the matching area, and subsequently the length of the epipolar line. This reduces both the risk of false matches and the time in matching by reducing the search space.

- A bundle adjustment is applied to refine the 3D coordinates that describe the scene geometry. It finds a set of parameters that most accurately predict the locations of the observed points in a set of images, thus minimizing the reprojection error between image locations of observed and predicted image points.

Figure 6 presents a complete application example of the corresponding ENVITask scripted in an IDL procedure. Not shown are optional steps such as error handling, etc., for sake of clarity.

The “PointCloudFeatureExtraction” task was executed to generate the post-DSM from the photogrammetric point clouds. The following parameters were set for the DEM extraction:

- DEM_GRID_RESOLUTION: 1 m (project density 3-4 points / m²).
- DEM_NEAR_TERRAIN_CLASSIFICATION: 0.5 m. Classifying points as near-terrain is useful to catch low objects that are not part of the terrain and are also not buildings, trees or power lines.
- DEM_FILTER_LOWER_POINTS: Rural Area Filtering. To filter, the area is divided to the grid resolution of 1 m. For each cell in the grid, the points are sorted by height and put into a histogram. Points that are below a multiplier of the standard deviation from the median are filtered out. Rural areas have less false points that need to be filtered; therefore the filtering is applied in a less rigorous way.

4 Implementation in Operational Environments

The presented automated point cloud generation and analysis approach can be embedded in virtually any existing geospatial workflow for operational applications. Three integration options were implemented as part of this case study:

- Integration within any ArcGIS® environment whether deployed on the desktop, in the cloud, or online. Execution uses a customized ArcGIS® script tool.

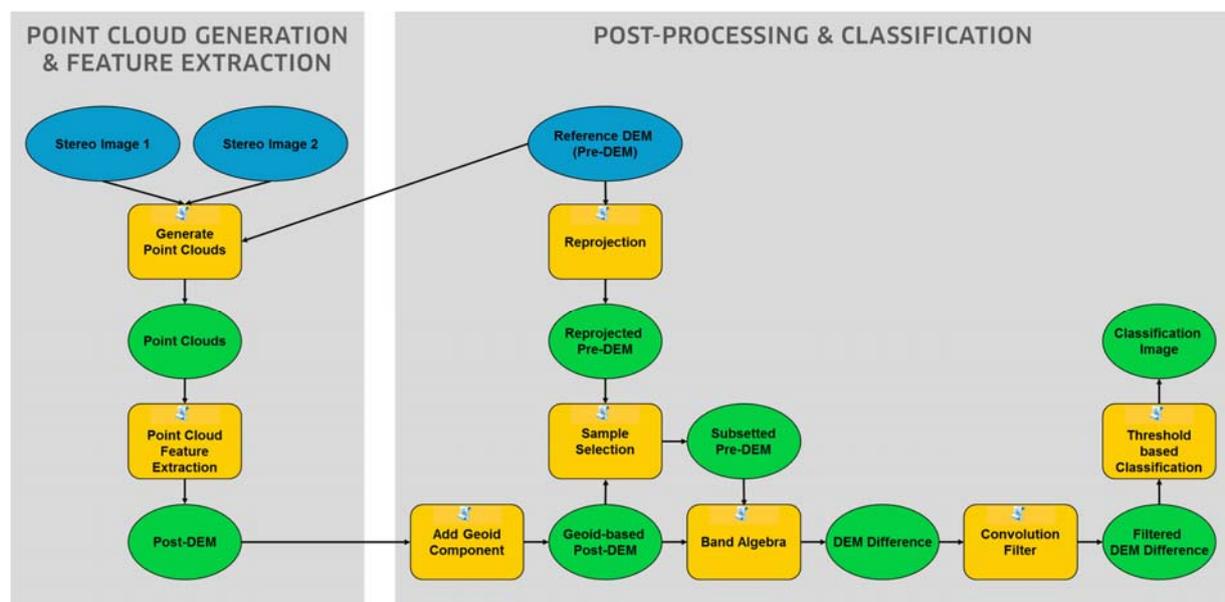


Fig. 7: Workflow for a point cloud based change detection implemented in the ArcGIS® ModelBuilder.

A Python script file retrieves the parameters from the user interface and runs the precompiled IDL code. That IDL code is used to interface between the Python script and the relevant ENVITasks. Figure 7 shows the workflow for a point cloud based change detection approach implemented in the ArcGIS® ModelBuilder.

- Publishing the point cloud processing tasks as services via the ENVI Services Engine (ESE). ESE is a cloud-based image analysis solution to publish and deploy advanced ENVI image and data analytics to existing enterprise infrastructures. For this purpose the entire IDL code can be capsuled into a single ENVITask.
- Integration in an existing geospatial workflow using the IDL bi-directional Python bridge. This mechanism allows calling IDL code within Python on a user-defined platform.

5 Conclusion

In the present, preliminary case study latest point cloud generation and analysis capabilities are used as part of an automated workflow to examine a landslide in India that affected an area of approx. 44.000 m². The topographic change detection is based on Pléiades high resolution satellite imagery and the Airbus DS WorldDEM™.

The results allow a 3D estimation of the topographic changes within the tectonically active and anthropogenically invaded area after the landslide event. Accordingly, the point cloud analysis was correlated successfully with modelled displacement contours of the slope. Based on optical satellite imagery, such point clouds of high precision and density distribution can be obtained in a few minutes to support the operational monitoring of landslide processes.

The future evolution of the automated workflow introduced here will focus 1) on the validation of the DSM products derived from photogrammetric point clouds by comparing them against existing reference products (e. g. DSMs, GPS points, etc.) and 2) on further geomorphologic analysis such as volume calculation of landslides.

6 References

- ERING, P., KULKARNI, R., KOLEKAR, Y., DASAKA, S.M. & BABU, G.L.S., 2015: Forensic analysis of Malin landslide in India. International Symposium on Geohazards and Geomechanics (ISGG) 2015, IOP Conf. Series: Earth and Environmental Science **26**, 012040, 1-12.
- RAMASAMY, S.M., MUTHUKUMAR, M. & SUBAGUNASEKAR, M., 2015: Malin–Maharashtra landslides: a disaster triggered by tectonics and anthropogenic phenomenon. Current Science **108** (8), 1428-1430.
- RIEGLER, G., HENNIG, S. D. & WEBER, M., 2015: WorldDEM – a novel global foundation layer. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-3/W2, 183-187.

Automatische Ableitung von Strukturlinien aus 3D-Punktwolken

**GOTTFRIED MANDLBURGER¹, JOHANNES OTEPKA¹, CHRISTIAN BRIESE², WERNER MÜCKE²,
GERHARD SUMMER¹, NORBERT PFEIFER¹, SVEN BALTRUSCH³, CARSTEN DORN⁴ &
HERBERT BROCKMANN⁵**

Zusammenfassung: Eine Methode zur automatischen Ableitung von 3D-Strukturlinien auf Basis von 3D-Punktwolken wird vorgestellt. Aufbauend auf einem Digitalen Geländemodell wird der 2D-Verlauf der Strukturlinien mittels des Canny Algorithmus näherungsweise ermittelt. Die detaillierte 3D-Modellierung erfolgt anschließend durch den Schnitt von elementaren Teilflächen (Ebene, Kegel, Polynomialer Zylinder), die aus den originalen 3D-Punkten abgeleitet werden. Das Verfahren wurde auf sechs Testgebiete mit Daten unterschiedlicher Messmethode (Airborne Laserscanning, Bildmatching) und Punktdichte (1-16 Punkte pro m²) angewendet. Bereiche mit ausgeprägtem Steigungswechsel (>8°) werden zuverlässig rekonstruiert. Manuelle Nachbearbeitung ist in Gebieten mit geringer Ausgeprägtheit und Datendichte erforderlich.

1 Einleitung

Ein modernes Digitales Geländemodell (DGM) repräsentiert die reale Geländeform objektiv auf Grundlage von möglichst homogen verteilten Messwerten ohne Hervorhebung oder Generalisierung. Auf Grund des großen, universell nutzbaren Informationsgehalts und der resultierenden hohen Wirtschaftlichkeit, ist Airborne Laserscanning (ALS) (SHAN & TOTH 2009; VOSSELMAN & MAAS 2010) das Standardaufnahmeverfahren im Sinne der anerkannten Regeln der Technik. Daneben gewinnt in letzter Zeit auch die Geländeerfassung durch Bildkorrelation (Dense Image Matching, DIM) zunehmend an Bedeutung (HIRSCHMÜLLER 2013; ROTHERMEL et al. 2015). In beiden Fällen ist das Ergebnis eine dichte Menge an 3D-Datenpunkten (3D-Punktwolke). Linienhafte interaktive Aufnahmen (Höhenlinien und/oder 3D-Strukturlinien) finden ressourcenbedingt nur noch punktuell Anwendung und sind eine Ausnahme bei speziellen Anforderungen.

Allerdings werden für vielfältige Aufgaben 3D-Strukturlinien (3D-Bruch/-Geländekanten oder 3D-Formlinien) als ein Basisprodukt zur Bereitstellung nutzerorientierter geotopographischer Produkte gewünscht/benötigt. U.a. zur:

¹ Technische Universität Wien, Department für Geodäsie und Geoinformation, Gusshausstr. 27-29, A-1040 Wien, E-Mail: [gottfried.mandlburger, johannes.otepka, norbert.pfeifer]@geo.tuwien.ac.at

² EODC Earth Observation Data Centre for Water Resources Monitoring GmbH, A-1040 Wien, E-Mail: [christian.briese, werner.mücke]@eodc.eu

³ Amt für Geoinformation, Vermessungs- und Katasterwesen, Postfach 12 01 35, D-19059 Schwerin, E-Mail: sven.baltrusch@laiv-mv.de

⁴ Hessisches Landesamt für Bodenmanagement und Geoinformation, Schaperstraße 16, D-65195 Wiesebanden, E-Mail: carsten.dorn@hvbh.hessen.de

⁵ Bundesanstalt für Gewässerkunde, Am Mainzer Tor 1, D-56068 Koblenz, E-Mail: Brockmann@bafg.de

- Vervollständigung des Grunddatenbestandes im Sinne der GeoInfoDok (AdV 2016) unter Beachtung des AAA-Standards,
- topographischen Modellierung im ATKIS Basis-DLM,
- kartographischen Ausgestaltung der ATKIS DTK-Reihe,
- optimierten Ableitung von Digitalen Othophotos (DOP),
- Modellierung von Digitalen Geländemodellen des Wasserlaufs (DGM-W),
- hydrodynamisch-numerischen Modellierung.

In diesem Kontext ist nach wie vor keine automatisierte oder zumindest weitestgehend automatisierte praxisreife Lösung zur Ableitung von 3D-Strukturlinien aus 3D-Punktwolken, aus hoch aufgelösten ALS- oder Bildkorrelationsdaten verfügbar. Eine adäquate Lösung wird aber aus fachlichen und wirtschaftlichen Gründen benötigt, da ressourcen- und zeitbedingt rein interaktive Erfassungen von 3D-Strukturlinien nicht mehr realisierbar sind. Im Zuge dessen wurde die TU Wien mit der Lösung dieser Fragestellung in Form einer Fachstudie beauftragt, die gemeinsam mit dem Earth Observation Data Centre for Water Resources Monitoring (EODC) durchgeführt wurde. Im vorliegenden Artikel werden die Ergebnisse dieser Studie präsentiert.

Der weitere Artikel ist wie folgt gegliedert. Abschnitt 2 enthält einen Abriss bestehender Ansätze. In Abschnitt 3 werden die Datensätze vorgestellt, auf denen die in Abschnitt 4 erläuterte Verfahrenslösung exemplarisch angewandt wurde. Die Präsentation und Diskussion der Ergebnisse erfolgt Abschnitt 5 und im Abschnitt 6 werden die wesentlichen Erkenntnisse zusammengefasst.

2 State-of-the-art

Existierende Ansätze für die Modellierung von Strukturlinien arbeiten entweder rasterbasiert oder direkt auf der 3D-Punktwolke. Zu Ersteren zählen u.a. die Arbeiten von BRÜGELMANN (2000), GOMES-PEREIRA & WICHERSON (1999) und SUI (2002). Typischerweise werden dabei, ausgehend von einem vorab erstellten DGM, durch die Anwendung unterschiedlicher Kantendetektionsalgorithmen (Sobel-Operator, Canny, etc.) 2D Linienpixel ermittelt, diese anschließend vektorisiert und die Höhe aus dem DGM übernommen. In WEHR & DUZELOVIC (2009) ist ein Verfahren beschrieben, welches die Detektion von Liniensegmenten anhand eines gefilterten TIN (Triangulated Irregular Network) vornimmt. Auch in diesem Fall erfolgt die Modellierung auf DGM-Basis. Bei allen gängigen DGM-Filterverfahren (PFEIFER & MANDLBURGER 2008) kommt es allerdings im Bereich von konvexen Geländekanten zum Ausrunden der in der Punktwolke relativ scharf abgebildeten Geländeunstetigkeiten (SITHOLE & VOSSELMANN 2004).

Im Gegensatz zu den o.a. Verfahren erfolgt die 3D-Modellierung bei BRIESE (2004) und BRZANK et al. (2009) ausgehend vom näherungsweise bekannten 2D-Strukturlinienverlauf daher direkt auf der 3D-Punktwolke und basiert auf dem Konzept des Verschnittes von lokalen Ebenenpaaren. Die Bestimmung der Ebenen erfolgt dabei für jedes Teilstück (im Folgenden: Patch) unter der Berücksichtigung von individuellen Punktgewichten. Das a-priori Gewicht eines Punktes ergibt sich aus seiner Lage im Bezug zur Näherungskante, wobei das Gewicht quer zur Kantenrichtung mit zunehmender Entfernung von der Näherung und in Längsrichtung mit zunehmendem Abstand

von der Mitte des Patches abnimmt. Im Zuge einer Ausgleichung werden die beiden Ebenen links und rechts der Kante derart bestimmt, dass die Quadratsumme der vertikalen Abstände zwischen den Datenpunkten und den Ebenen minimal wird. Die automatische Elimination von Vegetationspunkten gelingt dabei durch robuste Schätzung (d.h. iteratives Anpassen der a-priori Punktgewichte in Abhängigkeit der Verbesserungen). In BRIESE et al. (2010) ist dieser Ansatz zu einer vollautomatischen Methode weiterentwickelt worden, wobei die Kantendetektion durch Analyse der Hauptkrümmungen in den Datenpunkten erfolgt. Die Anwendung dieses Konzepts im hydraulischen Kontext ist in BROCKMANN et al. (2010) beschrieben. Die im Abschnitt 4 beschriebene Verfahrenslösung baut betreffend der 3D-Modellierung auf BRIESE (2004) auf und verwendet zur Kantendetektion einen rasterbasierten Ansatz ausgehend vom DGM.

3 Datensätze

Zur Erarbeitung der Verfahrenslösung für eine weitestgehend automatische Ableitung von 3D-Strukturlinien standen im Rahmen der in Abschnitt 1 erwähnten Fachstudie 3D-Punktwolken von 6 Testgebieten mit einer Fläche von jeweils 4 km² zur Verfügung. Die Datenerfassung erfolgte dabei für zwei Gebiete rein Luftbild-basiert mittels DIM (Punktabstand: 20-30 cm) und für drei Gebiete mittels ALS (Punktdichte: 1.5-9 Punkte/m²).

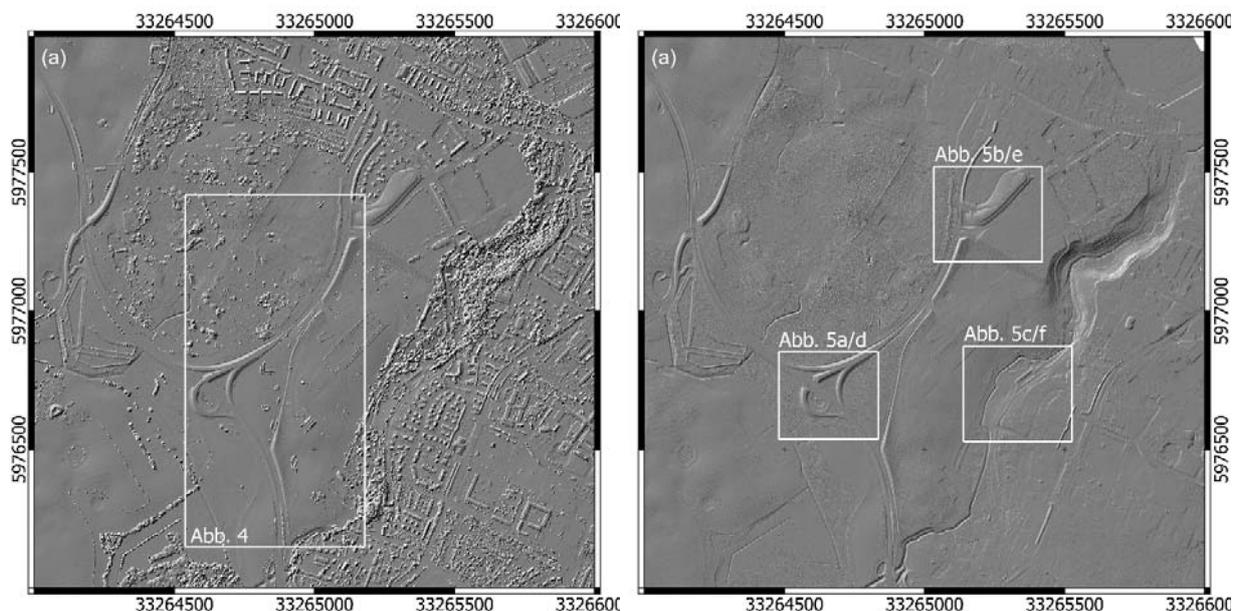


Abb. 1: Testgebiet Hansesstadt Wismar; (a) Schummerung DIM-Oberflächenmodell; (b) Schummerung ALS-Geländemodell; Bezugssystem: ETRS89/UTM, Daten: AfGVK M-V

Für den vorliegenden Artikel wurde ein Datensatz aus dem Bearbeitungsgebiet Hansesstadt Wismar (Abbildung 1) herangezogen, für den sowohl eine DIM- als auch eine ALS-Punktwolke vorlagen. Das dichte Matching wurde mit einem Bildmaterial aus dem Jahr 2012 (Überlappung: 80% längs, 60% quer; Ground Sampling Distance: 10 cm) durchgeführt, und die resultierende 3D-Punktwolke weist einen Punktabstand von 30 cm auf (Punktdichte: 11 Punkte/m²). Die ALS Daten stammen aus einer Befliegung aus 2014 mit einer mittleren Bodenpunktdichte von 16 Punkten/m².

4 Methode

Die Methode zur Ableitung von 3D-Strukturlinien ist zweistufig (Detektion/Modellierung) und ist in den folgenden Unterabschnitten beschrieben.

4.1 2D-Detektion von Strukturlinien

Aus der 3D-Punktwolke werden zunächst durch Anwendung geeigneter Filterverfahren (PFEIFER & MANDLBURGER 2008) die Bodenpunkte extrahiert und aus diesen mittels Flächeninterpolation (KRAUS 2000) ein DGM berechnet. Dieses DGM bildet die Basis für die Ableitung der 2D-Näherungen der Strukturlinien. Dazu wird zunächst ein Neigungsmodell in der Einheit [°] abgeleitet und darauf der aus der Bildverarbeitung bekannte Canny-Kantendetektions-Algorithmus (CANNY 1986) angewandt. Die resultierende binäre Rasterkarte wird vektorisiert und zusammenhängende Pixel zu Linienzügen zusammengefasst, die an Einmündungs- bzw. Kreuzungspunkten enden. In einem anschließenden Verarbeitungsschritt wird die Linientopologie automatisch überarbeitet mit dem Ziel, kleine Lücken im Linienverlauf zu überbrücken und ein Netzwerk möglichst langer Linienzüge herzustellen. Das Resultat ist ein topologisch korrektes Netz von 2D-Näherungslinien. Der gesamte Detektions-Workflow ist in Abbildung 2 dargestellt.

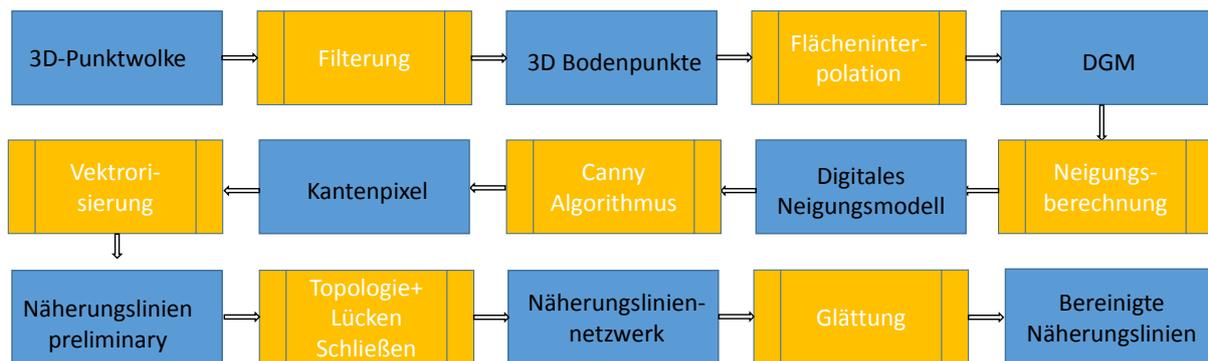


Abb. 2: Workflow zur Ableitung der 2D-Näherungslinien

Der Canny Algorithmus wird durch zwei Schranken gesteuert. Der obere Grenzwert gibt die Neigungsdifferenz an, ab der eine Kante detektiert wird und bestimmt die Anzahl der gefundenen Kanten. Der untere Schwellwert legt fest, bis zu welcher minimalen Neigungsdifferenz die Linie verfolgt wird und definiert daher die Länge der detektierten Kanten. Einerseits entstehen in Bereichen mit geringer Kantenausprägung Lücken im Liniennetz und andererseits kommt es bei komplexen Geländeformen zu Linienknoten, an welchen zusammenhängende Linien bei der Vektorisierung in Teilstücke aufgeteilt werden.

In einem Postprocessing-Schritt werden daher Linien innerhalb eines benutzerdefinierten Suchbereichs automatisch verbunden, sofern die Teillinien hinreichend gleich ausgerichtet sind. Befinden sich innerhalb des Suchradius mehrere Linienenden, dann wird die optimale Verbindung ausgewählt, wobei die Bewertung der Verbindungsgüte im Wesentlichen über den Abstand der Linienenden, sowie die Winkeldifferenz, die Gestrecktheit und die Länge der beteiligten Teillinien erfolgt. Um Unabhängigkeit von der Bearbeitungsreihenfolge zu erreichen wird die Suche der möglichen Verbindungen ausgehend von einem (primären) Linienende für alle gefunden

Linienenden in mehreren Generationen wiederholt. Das Verfahren liefert daher lokal optimale Linienverbindungen. Um den Berechnungsaufwand in Grenzen zu halten, kann die maximale Generationenanzahl vom Benutzer festgelegt werden. Darüber hinaus werden Teillinien, die in einem Linienknoten enden, durch die Anwendung des Kürzesten-Pfad-Algorithmus (Dijkstra 1959) zu möglichst langen Gesamtlinien zusammengesetzt.

4.2 3D-Modellierung von Strukturlinien

Die eigentliche 3D-Modellierung der Strukturlinien erfolgt auf Basis der 3D-Punktwolke und den 2D-Näherungslinien. Der Algorithmus startet mit dem Einlesen und dem Aufbau eines TIN aller Näherungslinien, wobei die Linien als Zwangskanten eingehen (Constrained Delaunay Triangulierung). Das TIN dient im Zuge der Modellierung als Begrenzung der Punktselektion auf maximal den Bereich bis zur benachbarten Linie. Außerdem werden aus dem TIN die Kreuzungspunkte extrahiert und für die Topologiebereinigung nach der Modellierung gespeichert. Die Modellierung erfolgt dann linienweise. Eine gesamte Linie wird in lokale, überlappende Patches dynamischer Länge und Breite aufgeteilt. Die Länge der Patches ergibt sich aus der Krümmung der Näherungslinie und der benutzerdefinierten minimalen/maximalen Patchlänge. Je geradliniger der Linienverlauf desto länger kann der Patch gewählt werden, ohne den Linienverlauf dabei zu generalisieren. Die Patchbreite wird generell vom Benutzer vorgegeben und zusätzlich über das TIN der Näherungslinien auf den Bereich bis maximal zur Nachbarlinie begrenzt. Für den derart bestimmten Patchbereich werden alle Punkte aus der 3D-Punktwolke selektiert. Eine weitere dynamische Patchlängenbeschränkung erfolgt noch auf Basis einer benutzerdefinierten maximalen Punktzahl.

Für jeden Patch liegen dann die 3D-Punkte aufgeteilt in zwei getrennte Mengen links und rechts der Näherungslinie vor. Für die nachfolgende Modellierung wird für jeden Punkt auf Basis seiner Lage innerhalb des Patches längs und quer zur Linienrichtung ein individuelles Gewicht berechnet (vgl. BRIESE 2004). Die Bestimmung der 3D-Lage und Orientierung des Patches erfolgt schließlich durch simultane Ausgleichung der folgenden Geometrie-Primitive: (i) Ebene, (ii) Kegel mit vertikaler Drehachse, (iii) polynomialer Zylinder, wobei jeweils die Quadratsumme der Vertikalabstände zu den Datenpunkten minimiert wird. Welche Flächenpaar-Kombination zum Einsatz kommt, hängt von der Krümmung der Näherungslinie und von der Böschungsform ab. Bei gestrecktem Verlauf kommt gemäß BRIESE (2004) das Ebenenpaar zum Einsatz. Weist die 2D-Näherungslinie gleichförmige Krümmung auf, bietet sich darüber hinaus das Ebene-Kegel Modell an, da dadurch Richtungsänderungen im Linienverlauf besser berücksichtigt werden können. Ist hingegen die Böschungsform durchhängend oder bauchig gewölbt, liefert ein polynomialer Zylinder bessere Ergebnisse, da die Krümmung der Böschung durch eine entsprechende Krümmung der Basispolynomkurve n -ten Grades erfasst werden kann. In der Praxis beschränkt man sich dabei auf Polynome des Grades $n=2$ (quadratische Parabel, kein Krümmungswechsel) und $n=3$ (kubische Parabel, z.B. S-Bogen). Falls der Schnitt über die o.a. Geometrieprimitive nicht erfolgreich durchgeführt werden kann (zu geringe Genauigkeit, zu geringer Schnittwinkel, zu wenige Punkte...), bleibt die Lage unverändert und die Höhe wird aus unabhängigen Ebenen links bzw. rechts der Näherung bestimmt. Ist auch das nicht möglich, erfolgt die Höhenschätzung aus nur einer (einseitigen) Ebene. Wenn alle genannten Ansätze fehlschlagen, wird der Patch als ungültig markiert. Für jeden gültigen Patch werden schließlich ein repräsentativer 3D-Punkt und dessen 3D-Richtung sowie weitere Metadaten (Standardabweichung der Modelleinpassung, Schnittwinkel, Berechnungsmethode, Krümmung, Punktzahl, etc.) gespeichert.

Nachdem alle Patches auf die oben beschriebene Art prozessiert wurden, erfolgt das Zusammensetzen der Patches zu einer Gesamtlinie. In diesem Zug erfolgt auch eine Prüfung aufeinander folgender Patches hinsichtlich Konsistenz der Fortschreitungsrichtung, etc., und eine Homogenisierung des Punktabstandes durch Einrechnen von möglichst gleichabständigen Zwischenpunkten entlang von stückweisen Bezierkurven dritten Grades. Ungültige Patches führen zu einem Aufspalten der modellierten 3D-Linie in mehrere Teile. Die Qualitätsbeurteilung der Gesamtlinie erfolgt durch statistische Analyse der Patch-Metadaten in Form eines Schulnotensystems (sehr gut bis nicht genügend). Vor dem abschließenden Export wird das Netz der unabhängig modellierten Kanten einer globalen Analyse und Korrektur der Linientopologie unterzogen. Dabei werden vor allem die aus dem TIN der Näherungskanten ermittelten Bereiche mit V-, T- und X-förmigen Netzknoten auf etwaige Klaffungen an Knoten bzw. auf windschiefe Linienkreuzungen hin untersucht und (i.d.R. durch Mittelbildung) automatisch korrigiert. Der gesamte Ablauf der 3D-Linienmodellierung ist schematisch in Abbildung 3 dargestellt.

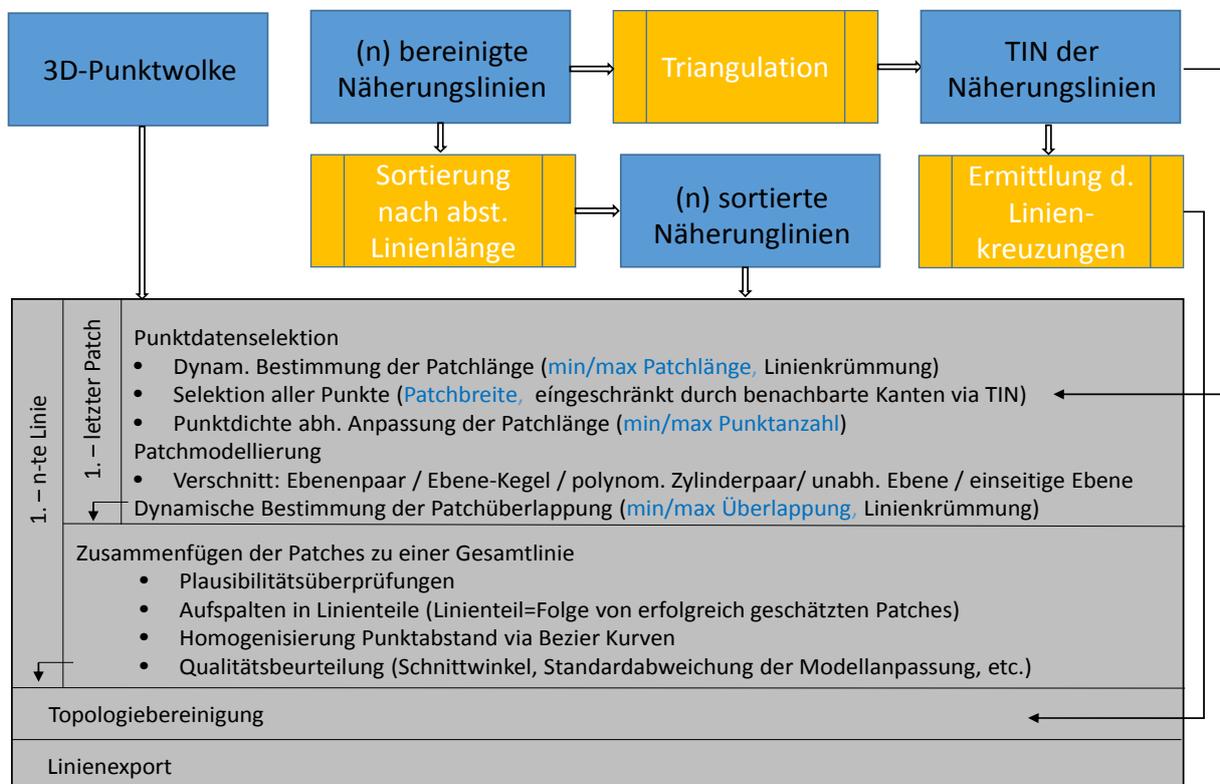


Abb. 3: Workflow zur Modellierung von 3D-Strukturlinien

5 Ergebnisse und Diskussion

In diesem Abschnitt werden die Ergebnisse der automatischen Detektion und Modellierung von 3D-Strukturlinien präsentiert und einer kritischen Diskussion unterzogen. Dazu wurde an der TU Wien das im Abschnitt 4 beschriebene Verfahren in Form einer Prototypsoftware in der Programmiersprache C++ implementiert (Vektorisierung des Linienrasters, Nachbearbeitung des 2D-Liniennetzes, 3D-Modellierung). Für den Canny Algorithmus wurde auf das open-source

Programmsystem Fiji/ImageJ (SCHINDELIN et al. 2012; SCHNEIDER et al. 2012) zurückgegriffen. Eine Implementierung des gesamten Workflows in der wissenschaftlichen Laserscanning Software OPALS (PFEIFER et al. 2014) ist in Arbeit.

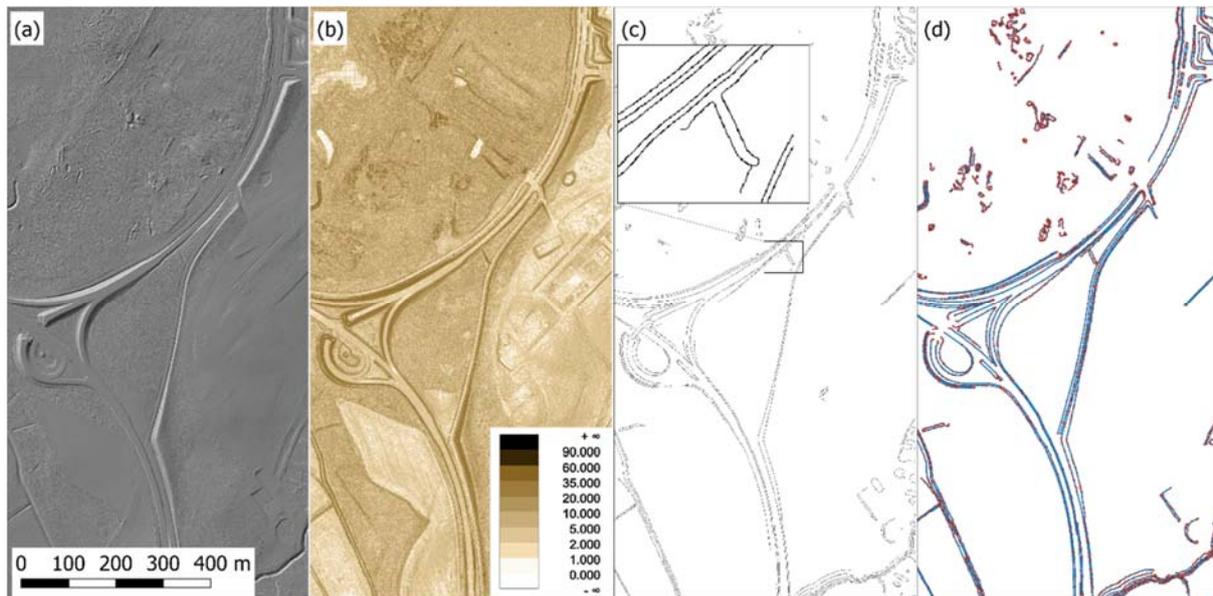


Abb. 4: Strukturliniendetektion; (a) DGM Schummerung; (b) Höhenkodierung des Neigungsmodells [°]; (c) binäre Rasterlinien-Karte (Canny Algorithmus); (d) Vektorisiertes, automatisch und teilweise manuell nachbearbeitetes 2D-Liniennetzwerk

Abbildung 4 zeigt die detektierten 2D-Näherungslinien für den ALS-Datensatz, wobei für den Canny Algorithmus als obere und untere Schranke Neigungsdifferenzen von 4° bzw. 1° verwendet wurden. Die resultierende binäre Rasterkarte ist in Abbildung 4c dargestellt. Die Vorgabe für die Ausweisung einer Strukturlinie liegt bei einem Steigungswechsel von mindestens 10° . Die Wahl des deutlich niedrigeren Wertes für die obere Schranke ist darin begründet, dass bei einem rein aus Massenpunkten berechneten DGM der Steigungswechsel nicht abrupt ausgeprägt ist. Beim ALS-DGM kommt darüber hinaus noch der Umstand zum Tragen, dass praktisch alle Filteralgorithmen zum Ausrunden vor allem der konvex gekrümmten Kantenbereiche neigen (SITHOLE & VOSSelman 2004). Die Ableitung von DGMen aus DIM-Punktwolken hingegen ist in Vegetationsbereichen wegen der mangelnden Durchdringung generell problematisch (RESSL et al. 2016). Es kann festgehalten werden, dass eine automatische Ableitung von 3D-Strukturlinien auf Basis von DIM-Punktwolken derzeit nur in offenen Bereichen zuverlässig gelingt, was die Möglichkeiten einer vollautomatischen Ableitung einschränkt. Für die weiter unten präsentierten DIM-basierten 3D-Strukturlinien Ergebnisse diene jedenfalls ein ALS-DGM für die Detektion der 2D-Näherungen.

Für die Anwendung im Kontext der hydrodynamisch-numerischen Modellierung sowie für das ATKIS-DGM1 sollten Strukturlinien neben dem Mindeststeigungswechsel von 10° ab einem seitlichen Abstand von ≥ 0.5 m und einem Höhenunterschied von ≥ 0.25 m detektiert werden. Während der Steigungswechsel, wie oben erwähnt, über die Schranken bei der Anwendung des

Canny Algorithmus berücksichtigt wird, kann der geforderte Linienabstand nur durch eine entsprechende DGM-Auflösung erreicht werden. Der Canny Algorithmus liefert eine binäre Pixelkarte, wobei die Linien auf eine Breite von einem Pixel reduziert sind. Der Mindestabstand zweier benachbarter Linien beträgt daher zwei Pixel, und folglich ist für die o.a. formulierten Anforderungen eine DGM-Auflösung von 25 cm erforderlich. Der vorliegende ALS Datensatz erfüllt mit einer mittleren Bodenpunktdichte von 16 Punkten/m² diese hohen Anforderungen. Für die derzeit im Kontext der Landes- und der hydrographischen Vermessung übliche Punktdichte von 2-10 Punkten/m² ist dies aber nicht gegeben. Hier ist realistischer Weise eher von einer Trennbarkeit von individuellen Linien ab einem Abstand von 1 m auszugehen. Abbildung 4d zeigt das 2D-Liniennetzwerk nach der Vektorisierung, der automatischen und manuellen Nachbearbeitung. Gegenüber dem automatischen Ergebnis (Abbildung 4c) sind vor allem im Bereich der Straßenkreuzung manuell zusätzlich Linien ergänzt, die zwar in der Schummerung und im Neigungsmodell (Abbildung 4a und 4b) erkennbar sind, aber deren Ausprägung nicht den oben genannten Kriterien entsprechen. Aus Sicht der Geländemodellierung fehlen sie daher zurecht, aus Gründen der kartographischen Vollständigkeit wurden sie aber ergänzt.

Die Ergebnisse der 3D-Modellierung sind in Abbildung 5 für beide Erfassungsarten (DIM/ALS) und für drei Teilbereiche (Straßenkreuzung mit Unterführung, offenes Gelände, Gerinne mit Vegetation und Siedlungsbereich) dargestellt. Für beide Datensätze wurde dabei die maximale Patchbreite mit 2.5 m und die (variable) Überlappung der einzelnen Patches mit 15-75% festgelegt. Eine größere Überlappung ist in Bereichen mit geringer Datendichte (z.B. Vegetation) und stark gekrümmten Bereichen notwendig, um die erforderliche Mindestpunktzahl (10 Punkte) nicht zu unterschreiten. Auf Basis der Linienkrümmung wird die Patchlänge vom Algorithmus automatisch zwischen vorgegebenen Schranken (hier: 3-15 m) bestimmt. Der höheren ALS Punktdichte wurde durch eine höhere Obergrenze betreffend die maximale Punktzahl pro Patch Rechnung getragen (DIM: 180 Punkte; ALS: 250 Punkte). Des Weiteren wurde die Punktanzahl entlang der modellierten 3D-Strukturlinien durch Einrechnen von Zwischenpunkten im 1 m-Abstand homogenisiert und alle Linien mit einer Länge kleiner als 10 m verworfen.

Die Linienfarben in Abbildung 5 spiegeln die ermittelte Qualitätsstufe wider (grün=sehr gut bis rot=nicht genügend). Auffällig ist dabei das deutlich höhere Qualitätsniveau der ALS-Ergebnisse (Abbildung 5d-f) gegenüber dem Resultat auf Basis von DIM (Abbildung 5a-c). Der errechnete Durchschnitt liegt für den DIM Datensatz bei 3.0 (1=sehr gut, 2=gut, usw.) und für ALS bei 1.4. Dies ist einerseits auf die deutlich bessere Präzision der ALS-Höhen und andererseits auf die stärkeren Abrundungseffekte beim DIM in Folge der Glattheitsbedingung (RESSL et al. 2016) zurückzuführen. Die ALS-basierten Linien verlaufen auch ruhiger und sind besser imstande, komplexe Strukturen mit engem Kantenabstand widerzugeben (vgl. Abbildung 5a/d bzw. 5b/e). Die Abbildung 5c/f verdeutlicht ferner, dass die Modellierung von Strukturlinien im Vegetationsbereich wegen der Durchdringungsfähigkeit nur für ALS möglich ist, wobei auch hier Laubfreiheit gegeben sein muss. Aus Abbildung 5f ist zu erkennen, dass auch der ALS-basierten Strukturliniendetektion und -modellierung durch die Vegetationsabschattung Grenzen gesetzt sind, da das Gerinne stellenweise in der vollen Trapezform mit jeweils zwei Ober- und Unterkanten, oftmals aber nur durch die beiden Uferböschungsoberkanten repräsentiert ist.

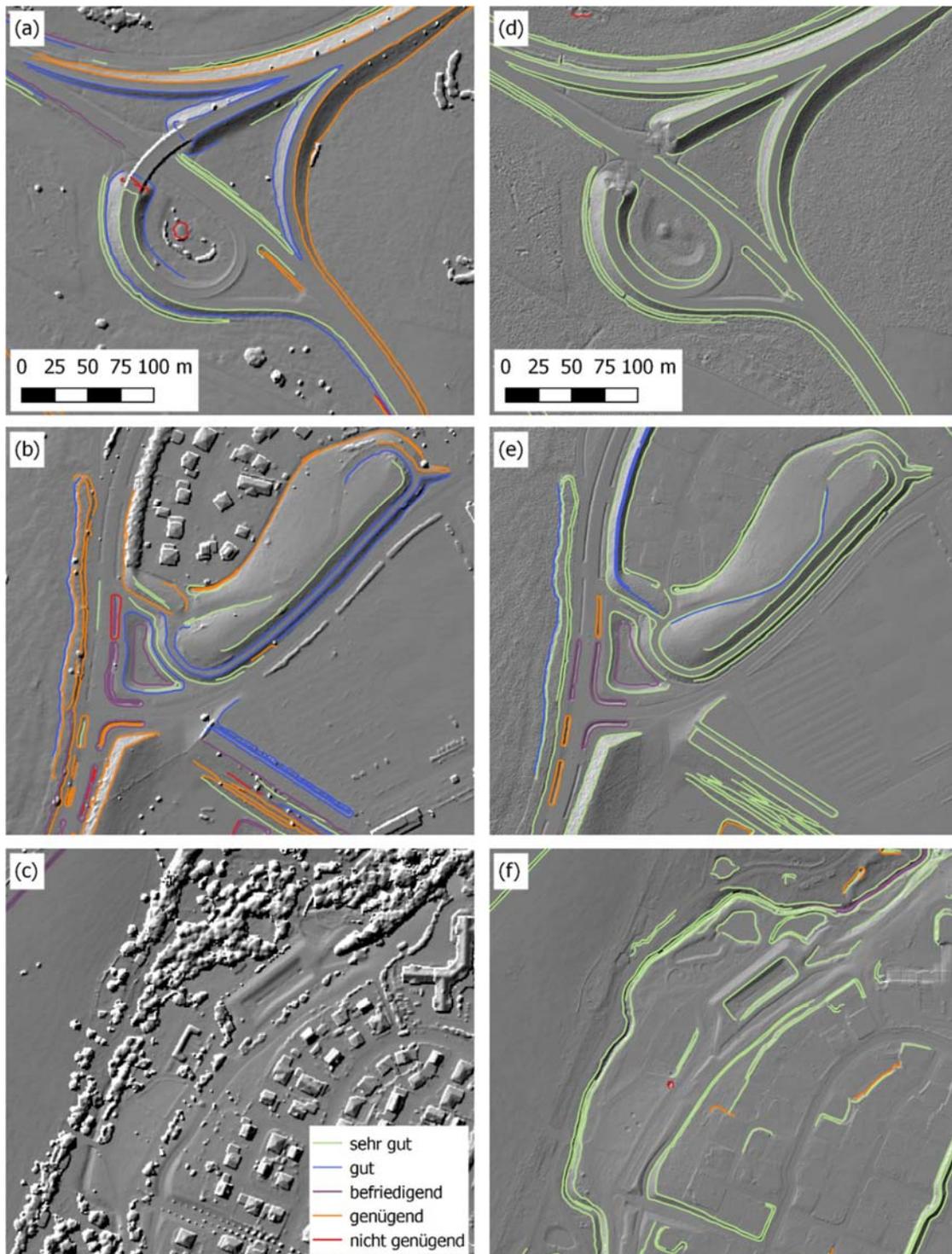


Abb. 5: Automatisch modellierte 3D-Strukturlinien; (a-c) Datenquelle: DIM, Bildhintergrund: DOM-Schummerung; (d-f) Datenquelle: ALS, Bildhintergrund: DGM Schummerung; Linienfarben entsprechen der Qualitätsstufe gemäß Legende in (c); (a+d) Straßenkreuzung mit Unterführung; (b+e) gemischtes, offenes Gelände; (c+f) Gerinne mit Vegetation und Siedlungsbereich

Unabhängig von der Datenerfassungsart werden nicht nur gestreckte Linienzüge sondern auch Stellen mit spitz zusammenlaufenden Linien (z.B. Straßenkreuzung in Abbildung 5a/d) oder engen Radien (z.B. Grabenstrukturen in Abbildung 5b/e) korrekt modelliert. Ermöglicht wird dies einerseits durch die konsequente Berücksichtigung von benachbarten Strukturlinien im Zuge der Datenselektion und andererseits durch die Erweiterung des Modellkatalogs (Kegel, polynomialer Zylinder). Obwohl die meisten Patches mit dem klassischen Ebenenpaar-Modell rekonstruiert wurden (DIM: 65%, ALS: 79%), liegt die Qualitätssteigerung bei der vorgestellten Verfahrenslösung vor allem in der Möglichkeit, komplexer geformte Geländebereiche besser rekonstruieren zu können. Eine manuelle Kontrolle und Überarbeitung bleibt zwar nach wie vor erforderlich, der Aufwand dafür konnte aber verringert werden. Bedarf für einen interaktiven Eingriff besteht vor allem bei der Bereinigung des 2D-Liniennetzwerkes, während sich die 3D-Modellierung auf Basis eines sauberen Gerüsts an Näherungen als robust und zuverlässig erwiesen hat.

6 Zusammenfassung und Ausblick

Es wurde ein weitestgehend automatischer, zweistufiger Ansatz zur Ableitung von 3D-Strukturlinien auf Basis von 3D-Punktwolken vorgestellt. In einem ersten Detektionsschritt wird der genäherte 2D-Linienverlauf durch Anwendung des aus der Bildverarbeitung bekannten Canny Algorithmus ermittelt. Die Basis dafür stellt das aus dem DGM abgeleitete Neigungsmodell dar. Das resultierende 2D-Liniennetz wird zum Überbrücken kleinerer Lücken und zur Ableitung möglichst langer Linienzüge automatisch nachbearbeitet und im Bedarfsfall manuell editiert. Gemeinsam mit der 3D-Punktwolke stellen die 2D-Näherungen die Eingangsdaten für die 3D-Modellierung mittels Verschnitt von lokalen Flächenpaaren dar. Aufbauend auf der von BRIESE (2004) eingeführten Methode, die auf dem Verschnitt von Ebenenpaaren basiert, erfolgte eine wesentliche Weiterentwicklung durch die Einführung weiterer Geometrieprimitive (Kegel, polynomialer Zylinder), welche eine bessere Berücksichtigung von Krümmungen sowohl längs als auch quer zur Linie erlauben. Das Kegelmodell ermöglicht die Modellierung von abrupten lagemäßigen Richtungsänderungen, und der polynomiale Zylinder erlaubt die Rekonstruktion von durchhängend oder bauchig gewölbten Böschungen. Weitere Verbesserungen des originalen Ansatzes konnten durch die dynamische Festlegung der einzelnen Linienteilstücke (Patches) mittels Analyse der Flächenkrümmung und der Berücksichtigung von angrenzenden Strukturlinien erreicht werden. Dadurch wurde eine Optimierung der Punktauswahl erreicht, ohne das Ergebnis zu generalisieren. Durch die konsequente Erfassung und Speicherung von Metadaten wie Schnittwinkel, Genauigkeit der Modellanpassung, etc. konnte auch eine Qualitätsbeurteilung der einzelnen modellierten Linien im Sinne eines Schulnotensystems durchgeführt werden, was vor allem bei der Ergebnisbeurteilung und manuellen Kontrolle durch den Operateur hilfreich ist. Das Verfahren wurde von der TU Wien in einer Prototypsoftware in C++ implementiert und auf Testdatensätze verschiedener Datenerfassungsart (DIM/ALS) sowie Punktdichte (1-16 Punkte/m²) angewendet. Die Ergebnisse für den Datensatz „Hansestadt Wismar“ wurden präsentiert und bestätigen das Potenzial des vorgestellten Verfahrens insbesondere hinsichtlich der möglichen Erfassung von abrupten Richtungsänderungen (Grabenenden, spitz zu laufende Linien) und der verbesserten Rekonstruktion von parabolisch gekrümmten Böschungen. Aufgrund der besseren

Höhengenauigkeit und Vegetationsdurchdringung konnten für den ALS Datensatz quantitativ und qualitativ bessere Ergebnisse erzielt werden. Für das 4 km² Testgebiet wurden auf Basis der ALS Punktwolke 1000 Strukturlinien mit einer Gesamtlinielänge von 90 km und einer durchwegs als „sehr gut“ eingestuften Qualität automatisch erfasst. Die DIM Punktwolke lieferte für dieselbe Fläche 500 im Mittel als „befriedigend“ eingestufte Linien mit einer Länge von 55 km. Zurzeit läuft die anwendungsbezogene Evaluierung durch den Auftraggeber der Fachstudie mit dem Ziel der Überführung der Ansätze und der Software in den Wirkbetrieb.

7 Literaturverzeichnis

- ADV 2016: Webseite der Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland. <http://www.adv-online.de/AAA-Modell/Dokumente-der-GeoInfoDok/>, Letzter Zugriff: 29.04.2016
- BRIESE, C., 2004: Three-dimensional modelling of breaklines from airborne laser scanner data. *International Archives of Photogrammetry and Remote Sensing* **35**, B3, Istanbul, Turkey.
- BRIESE, C., MANDLBURGER, G., MÜCKE, W., 2010: Maßstabsabhängige Modellierung von Strukturlinien aus Airborne Laser Scanning Daten. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, Band **19**, Wien, Österreich, 561-569.
- BROCKMANN, H., HAASE, A., MANDLBURGER, G., SCHRÖDER, M., 2010: LIDAR-Datenpotenziale für hydrodynamisch-numerische Modelluntersuchungen. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, Band **19**, Wien, Österreich, 542-551.
- BRÜGELMANN, R., 2000: Automatic breakline detection from airborne laser range data. *International Archives of Photogrammetry and Remote Sensing* **33**, B3, Amsterdam, Netherlands, 109-115.
- BRZANK, A., HEIPKE, C., GÖPFERT, J. & SOERGEL U., 2009: Ableitung Digitaler Geländemodelle im Wattenmeer aus luftgestützten Laserscannerdaten. *Die Küste* **76**, 91-121.
- CANNY, J., 1986: A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8** (6), doi: 10.1109/TPAMI.1986.4767851, 679-698.
- DIJKSTRA, E.W., 1959: A note on two problems in connexion with graphs. *Numerische Mathematik* **1**, 269-271.
- GOMES-PEREIRA L. & WICHERSON, R., 1999: Suitability of laser data for deriving geographical information – a case study in the context of management of fluvial zones. *ISPRS Journal of Photogrammetry and Remote Sensing* **54**, 105-114.
- HIRSCHMÜLLER, H., 2008: Stereo processing by semi-global matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30** (2), 328-341.
- KRAUS, K., 2000. *Photogrammetrie, Band 3, Topographische Informationssysteme*. Dümmler.
- PFEIFER, N. & MANDLBURGER, G., 2008: Filtering and DTM Generation. *Topographic Laser Ranging and Scanning: Principles and Processing*, J. Shan, C. Toth (ed.); CRC Press, ISBN: 978-1-4200-5142-1, 307-333.

- 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.
- RESSL, C., BROCKMANN, H., MANDLBURGER, G. & PFEIFER, N., 2016: Dense Image Matching vs. Airborne Laser Scanning – Comparison of two methods for deriving terrain models. *Photogrammetrie Fernerkundung Geoinformation*, Heft 2, 57-73.
- ROTHERMEL, M., WENZEL, K., FRITSCH, D. & HAALA, N., 2012: Sure: Photogrammetric surface reconstruction from imagery. *Proceedings of the LC3D Workshop*, Berlin.
- SHAN, J. & TOTH C., 2009: *Topographic Laser Ranging and Scanning: Principles and Processing*, CRC Press, ISBN: 978-1-4200-5142-1.
- SCHINDELIN, J., ARGANDA-CARRERAS, I., FRISE, E., KAYNIG, V., LONGAIR, M., PIETZSCH, T., PREIBISCH, S., RUEDEN, C., SAALFELD, S., SCHMID, B., TINEVEZ, J., WHITE, D.J., HARTENSTEIN, V., ELICEIRI, K., TOMANCAK, P. & CARDONA, A., 2012: Fiji: an open-source platform for biological-image analysis. *Nature methods* **9** (7), 676-682.
- SCHNEIDER, C. A., RASBAND, W. S. & ELICEIRI, K. W., 2012: NIH Image to ImageJ: 25 years of image analysis. *Nature methods* **9** (7), 671-675.
- SITHOLE, G. & VOSSelman, G., 2004: Experimental comparison of filtering algorithms for bare-earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* **59**, 85-01.
- SUI, L., 2002: Processing of laser scanner data and automatic extraction of structure lines. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **34** (Part 2), 429-435.
- WEHR, A. & DUZELOVIC, H., 2009: Bruchkantendetektion mit hochaufgelösten Laserscannerdaten, *Vermessung & Geoinformation* **4**, 408-415.
- VOSSelman, G. & MAAS, H.-G., 2010: *Airborne and Terrestrial Laser Scanning*. Whittles Publishing, ISBN 978-1904445-87-6.

Von der Punktwolke zum Terrainmodell

ROLF BART¹

Zusammenfassung: Es wird eine Methodik vorgestellt, aus 3D-Punktwolken automatisiert ein digitales Terrainmodell abzuleiten. Dazu werden die Daten in eine Datenbank übertragen und von Berechnungstools zur Weiterverarbeitung eingelesen. Die Daten werden in Pseudozellen, kleinräumigen Aggregationen von Punktmessungen, mit Clusteranalysen auf je einen besten Repräsentanten der untersten und obersten Ebene reduziert. Die Baum- und Strauchvegetation wird erkannt und als solche klassiert. Alle Objekte, die sich durch scharfe Umbiegungen der Oberfläche erkennen lassen, werden getrennt klassiert und einzeln als Punktansammlungen verwaltet. Die Umrandungen der Objekte sind Bruchkanten. Zu den Objekten und weiteren Geländebruchkanten erlauben Metadaten Zugriff auf Objekte und Bruchkanten. Die Daten werden nach Qualitätskriterien bewertet und können für die DTM-Bildung gezielt selektiert und in verschiedenen Formaten ausgegeben werden.

1 Einleitung

Die nachfolgend vorgestellte Methode dient der Nachbearbeitung von Punktwolken, die aus Drohnenaufnahmen erstellt werden. Verarbeitet werden 3D-Punktwolken, die möglichst keine Vorbereinigungen aufweisen. Für die Erarbeitung der Methode und die bisherigen Tests und praktischen Anwendungen wurden Punktwolken verarbeitet, die mit Pix4D erstellt wurden. Die Erarbeitung der Methode basiert auf Farbbildern mit einer Pixelgrösse zwischen ca. einem und fünf Zentimetern.

Ebenfalls von Bedeutung ist die Messung der Punktwolke mit automatisierten Verfahren über die Autokorrelation. Wie sich das Verfahren im gegenwärtigen Stand etwa bei Rastermessungen bewährt, ist nicht abgeklärt.

Beschrieben wird die Datenbereinigung im Hinblick auf die Erstellung eines digitalen Terrainmodelles (DTM) also der Abbildung der Erdoberfläche. Es ist klar, dass andere Zielsetzungen der Datenauswertung genauso ihre Berechtigung haben, hier jedoch nicht das Thema sind.

Die Beschreibung der Methode ist als Ergänzung zum Referat zu verstehen, welches die Methode mit Fallbeispielen illustriert.

2 Problemstellung

2.1 Grundlegendes

Wir verwenden digitale Terrainmodelle (DTM) seit den 1990er Jahren namentlich für zweidimensionale Überflutungsmodellierungen. Dies verlangt gute Daten im Gerinneschlauch und in den potentiellen Überflutungsgebieten. Die Abklärungen werden besonders in besiedelten Gebieten durchgeführt, wo die Gefährdung durch Hochwasser grosse Relevanz aufweist und eine

¹ Ingenieure Bart AG, Waisenhausstrasse 15, CH-9000 St. Gallen, Schweiz, E-Mail: rolf@Bart.ch

hohe Abklärungstiefe rechtfertigt. Nur die Kombination von hydraulisch guten Modellen zusammen mit einem guten DTM ergibt insgesamt befriedigende und in sich konsistente Ergebnisse.

Die heute verfügbaren technischen Mittel erlauben kostengünstige Bildflüge und effiziente Aerotriangulationen, inkl. automatisierter Punktmessungen in hoher Dichte und guter Qualität. Die unstrukturierten Punktwolken lassen sich mit den bisher verfügbaren Werkzeugen nur mit grossem Aufwand inhaltlich auf jene Daten reduzieren, welche ein DTM ausmachen. Nachstehende Punkte erweisen sich als wichtigste Probleme:

- Bruchkanten können nicht automatisiert erstellt werden
- Die Vegetation wird nicht zuverlässig genug entfernt
- Objekte aller Art (Gebäude, Fahrzeuge, Zäune, Kandelaber, Leitungen etc.) bleiben erhalten oder können nur unzulänglich entfernt werden
- Zur Qualität der einzelnen Punkte stehen keine Informationen zur Verfügung
- Den Punktdaten fehlen zusätzliche Attribute, um auf bestimmte Gruppen gezielt zuzugreifen
- Aufgrund fehlender Attribute kann eine Bereinigung der Daten zu einem DTM nicht differenziert erfolgen
- Es fehlen Grundlagen, die Daten nach Qualitätskriterien zu analysieren und eine übersichtliche Bewertung zu erstellen
- Die verbleibende Datenmenge ist generell zu gross
- Der Bearbeitungsaufwand, um von einer Punktwolke zu einem DTM zu gelangen ist zu hoch
- Eine verlässliche Beurteilung der resultierenden DTM Qualität ist ohne systematische, praktisch flächendeckende Kontrolle nicht möglich

Die aufgezählten Probleme erschweren es, dem nicht spezialisierten Praktiker Daten abzugeben, welche einfach und mit den gängigen Werkzeugen zu handhaben sind. Diese erschwert es, viele der möglichen Datennutzer als Kunden zu gewinnen. Der nötige händische Nachbearbeitungsaufwand erhöht die Kosten deutlich.

2.2 Beispiele von Datenproblemen und Lösungsansätzen

Nachfolgend wird anhand einiger typischer Datenausschnitte einer Punktwolke erläutert, welche Probleme sich ergeben und welche Lösungsansätze bestehen.

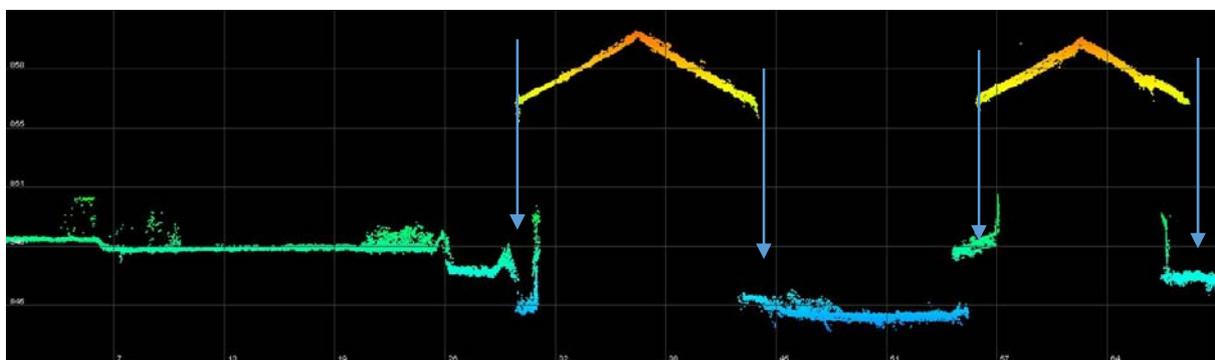


Abb. 1: Schnitt durch einen Parkplatz, zwei Wege und dann zwei Gebäude (von links nach rechts).

Die vier Pfeile an den Dachrändern zeigen auf, welche der gebäudenahen Bodeninformation bei einer Draufsicht verloren gehen. Im vorgestellten Verfahren werden die Bodenpunkte und Dachpunkte ohne Einschränkung erhalten, indem erstere in der Ebene 1 und letztere in Ebene 2 gespeichert werden.

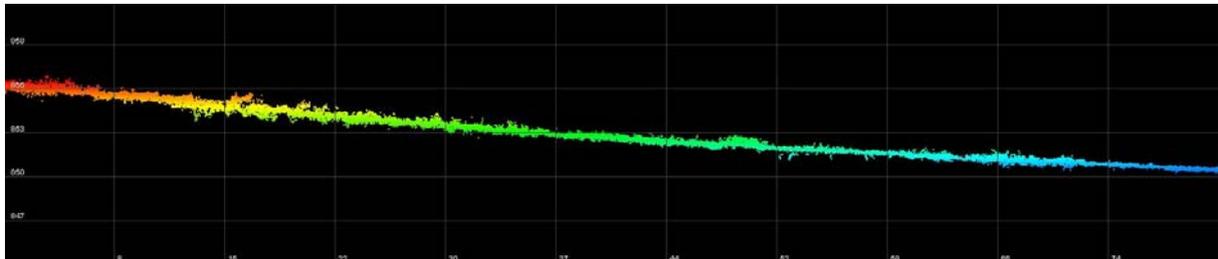


Abb. 2: Schnitt durch eine Strasse, die in Natura keine Unterschiede in der gleichmässig glatten, ebenen Oberfläche aufweist.

3 Lösungskonzept

Kontrollmessungen haben ergeben, dass der höher gelegene Abschnitt einen deutlich grösseren Fehler aufweist als der Teil rechts im Bild. Mit den vorgeschlagenen Qualifikationen der Messungen können Bereiche mit mutmasslich problematischen Messungen leicht erkannt und flächenhaft zugewiesen und dargestellt werden.

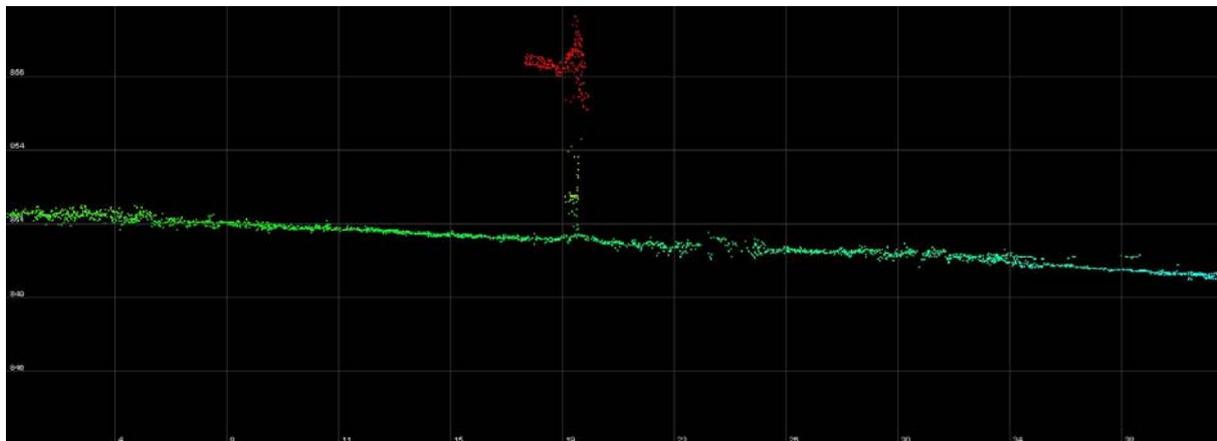


Abb. 3: Wegweiser in einer Strassenkreuzung.

Der Pfosten des Wegweisers verschwindet nach der Datenanalyse als Folge der Clusteranalyse. Die Punkte an den Schildern (rot) fallen in die Ebene 2 oder bilden ein Objekt, welches ausgeblendet oder dargestellt werden kann.

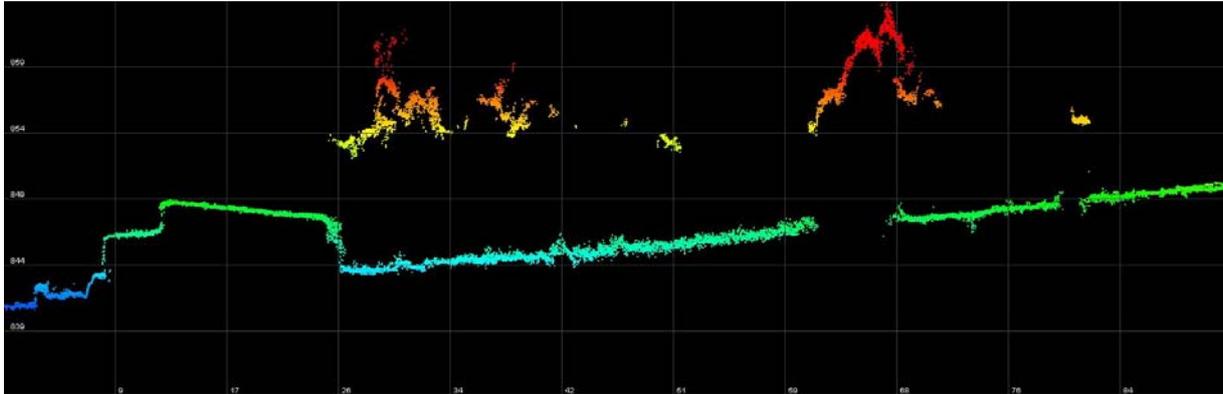


Abb. 4: Schnitt durch ein Gebäude und anschliessend eine Strasse mit Bäumen am Rand (von links nach rechts).

Die Messungen an den Bäumen (rot und gelb) können auf drei Arten von den noch vorhandenen Bodenpunkten separiert werden. Sie werden teils als Vegetation erkannt, teils fallen die Punkte in die Ebene 2 und an Orten, wo die Bodenpunkte fehlen, kann allenfalls ein Objekt entstehen, das wiederum verwendet oder verworfen werden kann.

3.1 Erste Auswahl der Datenpunkte

3.1.1 Art der Oberfläche und Datenstruktur

Als Information zu den Daten stehen in der Punktwolke nur die Koordinaten x und y sowie die Höhe zur Verfügung. Bei glatten, ebenen Flächen können die streuenden Höhen als Messfehler verstanden werden. Ist die Oberfläche uneben, wie beispielsweise bei einer Kiesbank eines Flusses, entspricht die sichtbare Streuung nicht alleine Messfehlern an Höhen. Die Messungenauigkeiten betreffen die Lage und die Höhe. Liegen zwei Punkte dicht beieinander, so kann sich der eine auf die Spitze eines Steines beziehen und der andere auf einen tiefer liegenden Punkt unmittelbar daneben. Umgekehrt können zwei Punkte in wenigen Zentimetern Abstand die Messung desselben Punktes darstellen. Die Unsicherheit nimmt bei Messungen in einer Wiese weiter zu. Auf engstem Raum finden sich potentielle Messpunkte verteilt auf mehreren Dezimetern Höhenunterschied. Es ist unmöglich, sicher zu unterscheiden, welche Abweichungen Messfehler zeigen und welche die Messung unterschiedlicher Objekte.

3.1.2 Grundgerüst der verwendeten Messungen

Die Daten werden zuerst in eine Mischform aus Vektor- und Rasterdaten überführt. Dazu wird jedem Punkt eine Rasteradresse zugewiesen. Man kann von einer hybriden Datenstruktur sprechen, die Pseudozellen bildet. Dadurch bleibt ein Datenpunkt im Vektorformat erhalten, gleichzeitig gehört er aber auch einer räumlich klar definierten Pseudozelle an. Die Pseudozelle bildet die Grundeinheit, die darin enthaltenen Messungen zu analysieren, zu klassieren und statistisch auszuwerten.

Die Wahl der Grösse der quadratischen Pseudozellen richtet sich nach den Ausgangsdaten (Punktwolke) und den Zielsetzungen des Anwenders. Für das Gros der interessierenden Flächen sollten in einer Pseudozelle drei bis sieben Punkte liegen. Die bisherigen Erfahrungen mit den eigenen Datenauswertungen zeigen, dass die günstige Kantenlänge einer Pseudozelle zwei bis drei

Pixel beträgt. Es ist kein Problem, wenn eine definierte Pseudozelle keinen Punkt enthält. In speziellen Fällen können es ohne weiteres dutzende von Punkten oder gar hunderte bis tausende sein.

3.1.3 Clusteranalyse

Wie bereits ausgeführt, darf nicht erwartet werden, dass innerhalb einer Pseudozelle die Streuung der Messungen als Fehler an einem konkreten Geländepunkt verstanden werden darf. Damit fallen übliche Verfahren, Messungen auszugleichen, weg. Als Verfahren wird eine Clusteranalyse pro Pseudozelle durchgeführt. Dabei werden die einzelnen Messungen so aufgefasst, dass sie ein Objekt in der a priori zu erwartenden Genauigkeit räumlich bestimmen. Weiter wird die Pseudozelle so definiert, dass sie lediglich eine Höhe pro Datenebene enthalten kann. Jeder Punkt erhält auf seiner exakten Höhe den Clusterwert 1. Dieser Wert nimmt gegen oben und unten bis zur vom Benutzer geschätzten Höhenmessgenauigkeit linear bis auf 0 ab. Werden die Messungen innerhalb einer Pseudozelle einander in der vertikalen Verteilung überlagert, so überschneiden sich die Messgenauigkeitsbereiche teilweise und werden addiert. Der Punkt mit dem grössten Clusterwert wird als Repräsentant dieser Ebene für die Pseudozelle bestimmt.

In mit Vegetation bedeckten Flächen zeigen die Datenanalysen, dass die oben liegenden Objekte oder Objektteile mit der Messung durch Autokorrelation begünstigt werden. Das heisst, dass in einer Pseudozelle die hoch an der Vegetation erfassten Punkte übervertreten sind. Um der gesuchten Erdoberfläche möglichst nahe zu kommen, werden die tief liegenden Punkte leicht begünstigt. Weisen innerhalb eines Clusters zwei Punkte denselben Wert auf, so wird immer der tiefer liegende gewählt. Ausreisser gegen unten werden unter bestimmten Kriterien einem höher liegenden Cluster vorgezogen. Ob es sich tatsächlich um einen Ausreisser handelt, wird in einer der nachgeschalteten Nachbarschaftsanalysen ermittelt.

3.1.4 Datenebenen

Beispielsweise in Randbereichen von Bäumen oder Gebäuden treten Überdeckungen auf, welche in den Daten erfasst sind. In derselben Pseudozelle können Punkte am Boden, von darüber hängenden Ästen und einem vorspringenden Dach überlagert werden. Die Clusteranalyse ist so umgesetzt, dass sie die Datenebenen nach Kriterien des Anwenders unterscheiden kann. Die Arbeiten an der Methode zeigten, dass der Umgang mit einer beliebigen Anzahl von Höhenebenen zu erheblichen Schwierigkeiten führt. Angesichts des nicht augenfälligen Nutzens von Zwischenebenen sind die Datenebenen zurzeit auf zwei begrenzt. Die zweite Ebene ist die oberste aller vorhandenen. Es werden maximal zwei Datenebenen verwaltet und in den Analysen vollumfänglich, wenn auch unterschiedlich, genutzt. Damit eröffnet sich eine erweiterte Palette an möglichen Endprodukten.

3.1.5 Auswirkungen von Datenebenen, Pseudozellen und deren Grösse

Die zentralperspektivische Datenerfassung erfasst an jenen Orten, an welchen die Überdeckungen einen gewissen Abstand zum Boden aufweisen, ebenfalls Bodenpunkte. Bei einer Datenanalyse aus der Draufsicht, gehen diese Daten verloren. Mit der Möglichkeit, zwei Datenebenen zu erfassen, lässt sich die Messung der Erdoberfläche in oft relevanter Weise vergrössern. Die

Kontrolle über die Daten wird ausgeweitet, da der Anwender Zugriff auf diese Information zu den Daten hat.

Werden die Pseudozellen deutlich grösser gewählt als die mittlere Pixelgröße misst, kann die Lage eines Punktes in der Pseudozelle tatsächlich genauer festgelegt sein als durch die Annäherung mit einem Zellmittelpunkt einer üblichen Rasterzelle. In der Clusteranalyse werden die Koordinaten des gewählten Punktes in seiner Vektorform übernommen. Die Verwendung von Pseudozellen in einer gewissen Grösse lassen Messungen “verschwinden”, welche in der Regel unerwünscht sind. Ein Lattenzaun oder ein Drahtgeflechtzaun kann eine grosse Zahl an Messungen in der Vertikalen liefern. Mit der Clusteranalyse ist es nahezu sicher, dass ein Bodenpunkt innerhalb der Pseudozelle als Repräsentant der Ebene 1 gewählt wird. Ist der Zaun hoch genug, so wird dessen obere Begrenzung in der Ebene 2 erhalten. Für die Abbildung der Ebene 1, der Erdoberfläche sind die störenden Messungen ohne weiteres Zutun weg klassiert.

3.2 Umgang mit Vegetation

Die spezielle Erkennung und Bezeichnung der Vegetation in den Daten bezieht sich auf hohe Vegetation. In der Regel handelt es sich um Bäume und Sträucher, einzeln stehend, in Gruppen oder Beständen. Derselben Kategorie gehören hohe krautige Vegetationsformen wie Mais, Schilf und ähnliches an. Diese hohen Vegetationsformen werden aus Merkmalen in den Pseudozellen und grossräumigen Nachbarschaftsanalysen erkannt und attribuiert.

Zwischen der Vegetation sichtbare Bodenoberfläche muss eine bestimmte räumliche Ausdehnung erreichen, damit sie ausgeschieden wird. Grundsätzlich wird Vegetation eher grosszügig als solche bezeichnet. Die digitale Photogrammetrie liefert zwischen hoher Vegetation kaum brauchbare Messungen. Daher werden mit diesem Verfahren zur Datenbereinigung an Stellen mit hoher Vegetation keine Messungen mehr verbleiben und dort andere Daten verwendet werden müssen. In der Vegetationsanalyse werden zwei Arten von Bruchkanten resp. Umrandungen gebildet. Die Bruchkante im eigentlichen Sinne ist die Verbindung aller Punkte, welche dem Trauf der Vegetation folgt. Die zweite Linie ist die Verbindung der letzten Bodenpunkte (von Vegetation bereits überschirmt), die photogrammetrisch noch erfasst werden konnten.

Sehr dichte, kompakte Vegetationsoberfläche kann nicht als solche erkannt werden. Typische Beispiele sind Einzäunungen mit stark geschnittenen Pflanzen wie Thuja, Buchs, Buchen etc. Diese Vegetation wird meist als Objekt erkannt und lässt sich so entfernen.

3.3 Objekte

Objekte überragen die Umgebung mit einer scharfen Umbiegung der Oberfläche. (Vorderhand führen nur Umbiegungen nach oben zu einem Objekt. Ein Schwimmbecken ist daher (noch) kein Objekt.) Die Schärfe der Umbiegung wird auf einer maximal möglichen Distanz untersucht. Dadurch ist indirekt auch die minimale Höhendifferenz im Bereich der Umbiegung definiert, welche zur Abgrenzung eines Objektes führt.

Die gewählte Definition des Objektes hat zur Folge, dass ein Kamin auf einem Dach ein Objekt ist und das Haus mit dem Kamin ist ein Objekt auf der Erdoberfläche. Jedes Objekt erhält eine im Projekt eindeutige Nummer, welche, wie die Daten selbst, in einer Datenbank verwaltet werden. Die Objektnummern werden als Identifikatoren jedem Punkt des Objektes zugewiesen. Zu den

Objekten werden Attribute mitgeführt, welche die Selektion von Objektgruppen erlauben. Die wesentlichen davon sind der Umfang, die Grundfläche und die maximale Meereshöhe.

Ein Subobjekt (Kamin) liegt auf einem Objekt (Haus) und ist somit in der Projektion vollständig von diesem umschlossen. Ein Subobjekt (Kamin) kann ebenfalls zu einem Objekt werden, wenn darauf wiederum ein Subobjekt (z.B. Kaminaufsatz) erkannt wird. Objekte und Subobjekte werden logisch verknüpft.

Geometrisch wird das Objekt von einer Fusslinie (Bruchkante) umschlossen. Die Stützpunkte der Fusslinie sind jene Punkte, an welchen die massgebende Bedingung des Objektrandes erkannt wurde (minimale Umbiegung). Auf dem Objekt können weitere Bruchkanten vorhanden sein, welche alle die Objektnummer des jeweiligen Objektes tragen

Ein Objekt hat den Zustand gewählt oder nicht gewählt. Ist ein Objekt nicht gewählt, so bleibt die Fusslinie des logisch höchsten (topographisch tiefsten) Objektes als Teil der Erdoberfläche erhalten und umschliesst den Ort des nicht ausgewählten Objektes als inneren DTM-Rand.

3.4 Keine Daten vorhanden

Versteht man die Daten als reine Vektordaten ist die Bezeichnung „NoData“ zumindest unpräzise. Obwohl in den gespeicherten Daten keinerlei Information zu Bereichen ohne Daten erfasst sind, ergeben sich diese indirekt durch das Fehlen von Rasteradressen mit Messungen. Damit können Bereiche ohne Daten (aus Sicht der Rasterstruktur) mit einer Begrenzungslinie umschlossen werden. Die Stützpunkte dieser Linie sind die Punkte der Ebene 1 unmittelbar vor einer Datenlücke. Diese muss das definierte Mass der Relevanz aufweisen.

Die Information „NoData“ kann für grössere, mit Linien umgrenzte Flächen dazu verwendet werden, Ersatzdaten gezielt einzulesen. Je nach Anwendung kann es genügen solche Daten beispielsweise mit DATM-AV Daten minimal zu bestücken.

3.5 Bruchkanten

Die Bruchkanten, welche im Zuge der Objekterkennung und der Vegetationserkennung ermittelt werden, sind in den entsprechenden Kapiteln bereits behandelt. Daneben werden Geländebruchkanten bestimmt. Als Bedingung müssen sie eine minimal vorgegebene Geländeumbiegung wiedergeben. Die Schärfe der Umbiegung muss nicht derart ausgeprägt sein wie bei der Objekterkennung. Der dazu nötige Raum wird in Abhängigkeit der gewünschten DTM-Auflösung bestimmt. Die Bruchkanten werden nach der topografischen Lage als konvex oder konkav unterschieden. Bruchkanten mit wechselnder topographischer Lage dürfen sich in kürzeren Abständen folgen als solche mit gleicher topografischer Lage.

Zu den Geländebruchkanten werden ähnlich den Objekten Metadaten gespeichert. Dazu gehören eine im Projekt eindeutige Nummer, die topographische Lage, die Länge und Informationen zur Geländeumbiegung an der Bruchkante.

3.6 Zusatzinformationen zu den Punkten

3.6.1 Qualitätsmerkmale

Mit Zusatzinformationen zu den Punkten sollen verlässliche Aussagen zur Messqualität eines Punktes abgeleitet werden. Die Qualifikationen liefern Zahlenwerte, dennoch kann lediglich von

einer qualitativen Beurteilung der Messungen gesprochen werden. Die quantitative Beurteilung setzt Vergleiche mit Kontrollmessungen voraus. Diese können zwar jederzeit erfolgen, werden in der vorgestellten Methode verfahrenstechnisch jedoch nicht vorausgesetzt oder direkt verwendet. Aufgrund der vorgenommenen Analysen bieten sich insbesondere die folgenden Merkmale an, die Messqualität zu beurteilen:

Der **absolute Clusterwert** gibt das Gewicht eines Punktes an, der als Repräsentant der Pseudozelle ausgewählt wurde. Der Wert 1 ist das Minimum. Liegen beispielsweise drei Punkte annähernd auf derselben Höhe geht der Wert gegen 3. Es darf vermutet werden, dass der Punkt mit guter Wahrscheinlichkeit repräsentativ ist.

Der **relative Clusterwert** setzt den absoluten mit der Anzahl der Punkte in der Pseudozelle ins Verhältnis. Ergibt sich ein relativer Clusterwert von knapp 3 aus 5 Messungen, so zeugt dies von einer besseren Messung als wenn dazu 23 Messungen vorliegen. Dass die grössere Anzahl an Messungen auf die Grösse der Pseudozellen zurückzuführen sein könnte, spielt für die Wertung keine Rolle.

Die **Streuung** aller Messungen in einer Pseudozelle ist ein geeignetes Mass, um die Sicherheit der Messung einzuschätzen. Die Streuung wird nicht mit der erwarteten Messgenauigkeit verglichen, da innerhalb einer Pseudozelle Messungen erfolgen, die sich in der Realität auf unterschiedliche Punkte beziehen können.

Die **Plausibilität** wird anhand der Abweichung von Nachbarpunkten beurteilt. Dabei wird eine geringe Abweichung zu Nachbarn als günstig und eine grosse als ungünstig beurteilt. In der Beurteilung der Abweichungen wird die Streuung unter den Vergleichspunkten mit berücksichtigt. Die **Anzahl** der Messungen erlaubt eine Einschätzung, ob an der betrachteten Stelle Messprobleme auftauchen. Über die Art der Probleme kann aus den Daten keine vernünftige Vermutung angestellt werden. Eine für das Operat atypisch geringe Zahl von Messungen weist auf Schwierigkeiten bei der Autokorrelation.

3.6.2 Topografische Lage

Die topographische Lage der Punkte ist generell in konkav und konvex unterteilt. Der Wechsel der topographischen Lage der Punkte kann in (deutlich) engeren Abständen erfolgen als die gewählte DTM-Auflösung. Bei der späteren Komposition der DTM-Daten ist zu beachten, dass in sehr rauem Gelände (Blockfelder, gepflügter Acker und ähnliches) die mittlere Höhenlage und die übernommene Rauigkeit ins DTM stark davon abhängt, ob konkave oder konvexe Punkte übervertreten sind.

3.7 Vorauswahl der Massenpunkte

Am Ende der Datenanalyse und -klassierung werden die Punkte, welche zu keiner besonderen Klasse von Umrandungen oder Bruchkanten gehören, ein letztes Mal bewertet. Die zu bewertenden Punkte werden so gewählt, dass der gewählte Bereich nirgends Bruchkantenstützpunkte überschreitet. Bei der Bewertung der Plausibilität (Abweichung zu Nachbarn) hätte der Einbezug von Punkten, die jenseits einer Bruchkante liegen, eine Verfälschung zur Folge. Die Unterbrechung der Datenanalyse an den Bruchkanten trägt zur Erhaltung der Geländekanten bei und verhindert die Bildung von Deformationen in unmittelbarer Nähe von Bruchkanten. Die Qualitätsmerkmale werden ein letztes Mal nachgeführt.

4 Konzept der Datennutzung

Die Verarbeitung ist für Punktwolken konzipiert, welche den vollen 3D-Datengehalt noch aufweisen. Die vorgestellte Methode nutzt einzig die aus der Verteilung der Punkte ableitbaren Informationen. Die Daten werden in eine Datenbank eingespielt und in eine hybride Datenstruktur überführt, die gleichermassen Eigenschaften von Vektordaten und Rasterdaten aufweist. Der Anwender lenkt die Datenanalyse mit seiner Parametersetzung hinsichtlich der gewünschten Auflösung und Datenqualität. Die resultierende Datensammlung ist geeignet, unterschiedliche Endprodukte zu erzeugen.

Die Anwendung ist sowohl für die Annahme von Daten wie auch den Export zu verschiedenen Formaten offen.

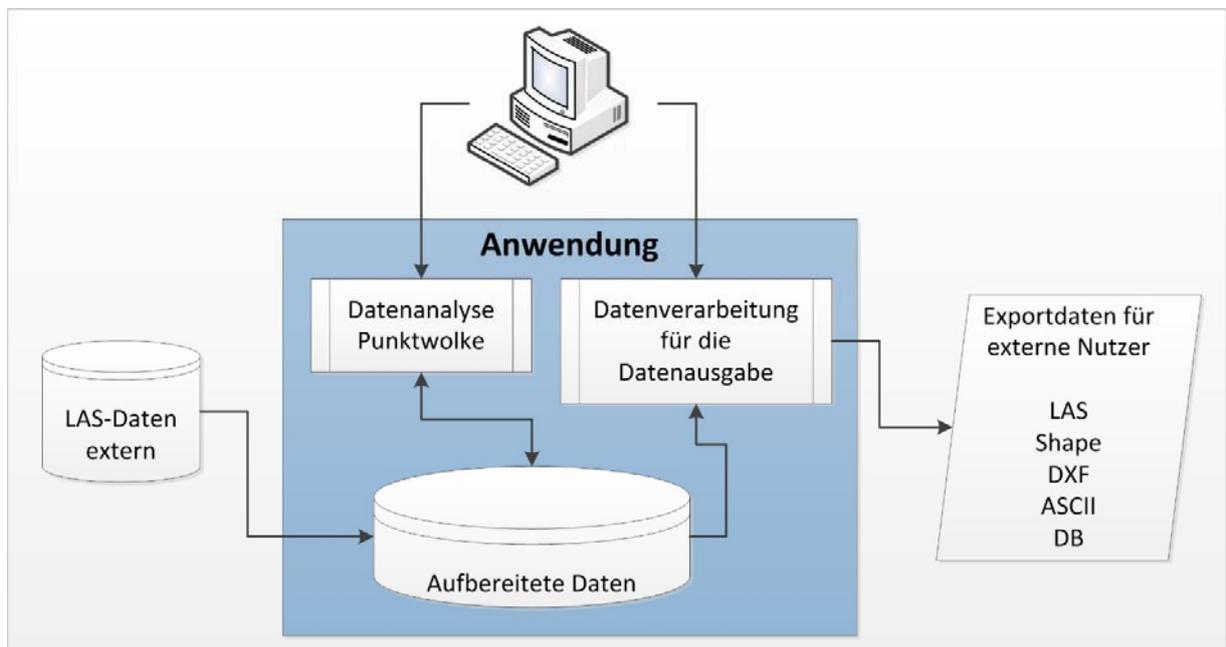


Abb. 5: Schema der konzeptionellen Datennutzung

Die differenzierte Datenauswertung und -ablage erlauben es, die Daten nach unterschiedlichen Zielsetzung zu Endprodukten zu komponieren. Die folgenden, schematischen Darstellungen zeigen das Wesentliche auf.

Mit der vorgestellten Methodik welche Objekte erkennt und ausblenden kann, Vegetation erfassen oder einblenden kann und Überdeckungen in zwei Ebenen aufteilt, können unterschiedliche DTM resp. DOM (digitale Oberflächenmodelle) erstellt werden.

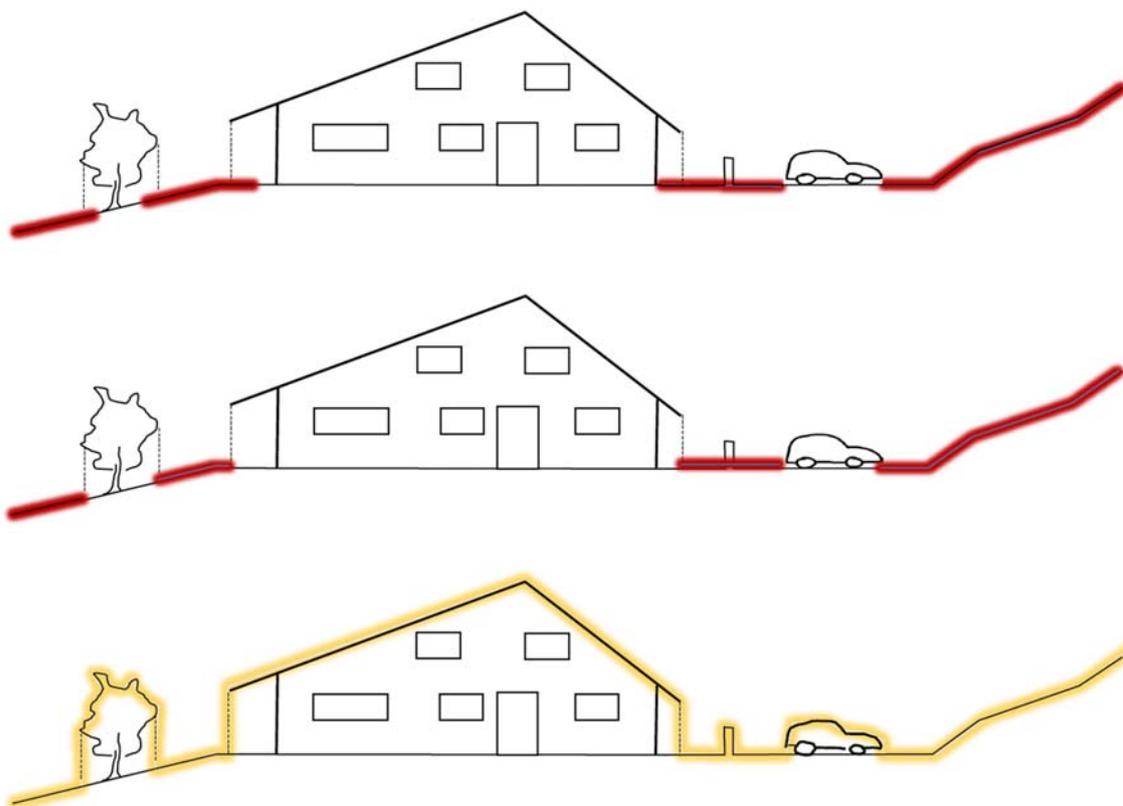


Abb. 6: Schemata von DTM und DOM

Die oberste Version entspricht dem üblichen DTM. Die Daten werden soweit sie vorhanden sind, an ein Gebäude heran als Erdoberfläche verwendet. Objekte wie Bäume, Mauern oder Fahrzeuge gehören nicht zum DTM. Die zweite Version verwendet die Daten nur bis an die Überdeckung, jedoch nicht unter diese. Auch bei Bäumen hört die Erdoberfläche am Trauf auf. Die letzte Version ist ein Dom, welches zwischen Objekten und Vegetation die bereinigte Erdoberfläche des DTM verwendet und ergänzend die Objekteinformation und die Baumvegetation enthält. Ein DTM kann allerdings auch gezielt mit einzelnen Objekten, beispielsweise Mauern ergänzt werden. Letztere sind bei 2D Überflutungsmodellierungen wichtig.

5 Fazit und Ausblick

Die gewählte Methode erzeugt aus Punktwolken automatisch Daten als Grundlage für DTM. Die aufbereiteten und klassierten Daten werden in einer Datenbank gehalten. Dadurch ist ein gezielter Zugriff auf die Daten möglich, um die gewünschten Produkte zu komponieren.

Es dürfte sich lohnen, die Automatisierung noch weiter zu verbessern. Bei der Aufbereitung der DTM-Daten ist vorstellbar, das gewünschte Endprodukt zu spezifizieren und das Programm die geeignete Parameterwahl anhand von gewählten Referenzflächen bestimmen zu lassen.

Die Möglichkeit, Bruchkanten automatisch zu erfassen, könnte künftig auch auf senkrecht stehende Flächen angewandt werden. Mit der Erweiterung um transformierte Koordinatensysteme, könnten Felswände oder Fassaden detailliert analysiert werden.

Recognizing Separate Structural Elements of Churches Using Neural Networks

MAKSIM BRODOVSKII¹, DMITRII KOROVIN¹, MARIA CHIZHOVA²,
ANSGAR BRUNN² & UWE STILLA³

Abstract: In this paper we develop a new approach to recognize structural elements of orthodox churches. We will work with 3D point clouds, received as a result of 3D point cloud acquisitions of churches, e.g. from laser scanning. Because of the large amount of points in such clouds, we have to use a projection (elevation) to decrease the calculation effort. To get meaningful images from the projection of the point cloud we do some prior segmentation of the 3D cloud. Images binary, with a predefined resolution that depends on the resolution of the 3D point cloud. To recognize elements we decide to use neural networks (Perceptron and Counter propagation neural networks) as they allow the automation of the process and have a broad range of methods to recognize images. For the subsequent 3D modeling we use analytic expressions, that describe each of the structural church elements. A further step is the deduction of those expressions that describe each of the recognized sectional views.

1 Introduction

1.1. Motivation

In a previous article (KOROVIN et al. 2016) a new approach for the reconstruction of Russian Orthodox churches have been developed on the basis of a Bayesian network and cellular automaton. The aim of that work was to present an algorithm, which develops the optimal process sequence of the automatic search, detection of the set of geometric objects and reconstruction with high probability of buildings and building components from a point cloud independent of its destruction. The first step of it was to recognize the single components of the church. In our work we would like to present the method, with which the geometry of such components could be recognized.

1.2. Previous Work

The recognition of object components with the aid of point clouds is one of the biggest tasks of reverse engineering used in many branches like industry, architecture, 3D modeling and robotics. The extraction of geometric information from the point cloud and its interpretation play a key role in this process. There are some common algorithms and mathematical techniques using 3D data e.g. from laserscanning as well as 2D data: RANSAC, Hough Transform, Least square fitting etc. The RANSAC-algorithm finds a mathematically described model from a set of data e.g.

¹ Ivanovo State Power Engineering University (Russia)

² Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Labor für Photogrammetrie & Fernerkundung, Röntgenring 8, D-97070 Würzburg, E-Mail: [maria.chizhova, ansgar.brunn]@fhws.de

³ Technische Universität München, Photogrammetrie und Fernerkundung, Arcisstraße 21, 80333 München, E-Mail: stilla@tum.de

laser scanning point data and estimates the model parameters by the interpretation of data as inliers (belong to model) and outliers. The algorithm has its advantages and disadvantages.

Advantages are:

- robust model estimation
- feasibility to find a geometric object in a noisy point cloud

Disadvantages are:

- the algorithm can estimate only one model and is not suitable for complex objects,
- the algorithm correctness depends from the number of iterations and mistake threshold,
- time involved.

SCHNABEL et al. (2007) uses an efficient RANSAC for the fitting of geometrical forms. The algorithm finds a best-fit geometrical form from candidates for the surface.

Hough Transform is a common method for detecting simple shapes like straight lines and circles in digital images, a 3D-shape evaluation (e.g. Kernel-based Hough Transform) can be used for simple geometric shapes like plane, cylinder and sphere. Geometries detection from point cloud with Hough Transform is presented in VOSSELMAN et al. (2004) and RABBANI et al. (2005). It is possible to find curves, too. However, the algorithm is generally suitable for simple forms and has some limitations like:

- the object size must be relative big, because the number of votes can fall to neighboring objects;
- the large number of model parameters can lead to mistakes;
- the quality of data plays an important role for algorithm efficiency (data denoising is preferred)

The method of least square fitting have been considered in AHN (2004). MARSHALL et al. (2001) presents an algorithm for the least squares fitting of spheres, cylinders, cones, and tori to point data. Although the least square method became a standard method for model estimation, it is not robust and can lead to the finding of incorrect model by noisy dataset. Huber (1964) developed in his work the method of robust parameter estimation. The assumption of this method is an explicit parameter modeling, which we do not need using neural networks.

The fitting of architectural models is considered in CANCIANI (2013). The method bases on the modelling paths definition, i.e. the profile of single component was extracted by the point cloud, analytical described. Then a building component is generated from knowledge-based model of profile line using different paths like circular, square, octagon etc. In ALBY & GRUSSENMEYER (2012) the architectural elements were modeled using the comparison of a knowledge-based model with the point cloud: a geometric primitive was closely adjusted to the point cloud and the surface distance for more as 90% of, by geometric primitives simplified, model within 5 cm of the point cloud becomes a resulting model.

An interesting algorithm of a surface fitting from unorganized, sparse, noisy point cloud using neural networks based on regression algorithm is presented in YUMER & KARA (2012). A free form surface was generated from a 3D point sets through a parametric embedding and tessellation in 2D-coordinates space. Then a neural network will be trained to learn a mapping by corresponding 3D to 2D coordinates, resulting in the synthesis of 2D surface in the model space. The activation function effects at the same time on the surface on the resulting surface.

ALEXANDRE (2014) presents different ways to train convolutional neural networks (CNN) with RGB-D data (color and depth). The approaches of independently for each channel trained CNN and by transfer learning trained CNN have been compared and analyzed. A wide investigation of neural networks is considered in CICHY et al (2016). The human brain capacity to recognize visual object has been presented through Deep Neural Networks based on anatomical analyses. The aim of our work is to recognize complex structural components, which cannot solve as geometric primitives.

2 Preprocessing and data preparation

2.1 Prepared data

The patterns for the training of neural network are symmetrical black-white 2D images of church structure elements that illustrate vertical and horizontal projections of its elements.

The input data are point clouds of church individual structural elements with the correct orientation, i.e. they are straight with no tilting. In our case, these elements are the domes, roofs and crosses. For pattern recognition, we will use 2D images, because it drastically reduces the computation due to a significant reduction in the number of points considered. These images will be black and white. Black pixels will correspond to the points of the point cloud. All other pixels are white.

In our work we have simulated in raster graphics editor Adobe Photoshop the point cloud cuts of structural elements (Fig. 1d) with taking in account the average resolution (with distance between closest points), which is usually suitable by laser scanning of such elements, and possible noise. Moreover, we try our algorithm with real laser scanning data namely point cloud cuts from Russian Orthodox Church in Wiesbaden (Germany) using 3D scanning software Trimble Realworks.

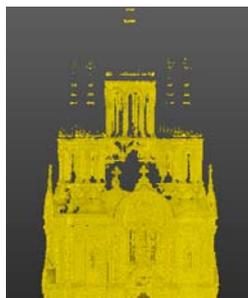


Fig.1a: Point cloud of the Russian Church in Wiesbaden



Fig.1b: Vertical projection (cuts on different distances) of the whole Church

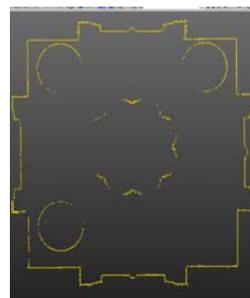


Fig.1c: Horizontal projection (cuts on different heights) of the whole Church



Fig.1d: Simulated vertical projection of the cupola

2.2 Processing of 3D point cloud (obtaining 2D black and white images)

Our first task is to get proper 2D image. To do this, we will use the vertical and horizontal projections. The horizontal projections will be secondary, as they will be used to obtain information

about analyzed object. In our case, there could be two types of the horizontal projections: a circle and a polygon. If the horizontal projection is a circle, we are dealing with body of rotation, and to obtain the desired vertical projection we should find a section, that would pass through the center of the circle. If the horizontal projection is a polygon, in general, the section for vertical projection would be parallel to one of the polygon sides. In polygon is a rectangle, the section for vertical projection would be parallel to the longest (or shortest) polygon side. If the polygon is not a rectangle, the projection would pass through the polygon center and intersection of two sides.

Thus, upon obtaining a horizontal projection, we should detect what is it, a circle or a polygon. One of the main differences between a polygon and a circle is presence of intervals that lie along lines. Thus, with the help of Hough transformation we could calculate the number of points that lie on such intervals. Using a predetermined threshold value for quantity of points on a segment (for example, 5 points) we can determine the number of such segments. Hough transformation also could help us to determine the analytical formulas for those lines. Using these formulas we could calculate the angles of the polygon and determine, is it a rectangle or not. The longest side of a polygon can be determined by the highest number of points that belong to the corresponding line that defines the segment. As to the center of a circle or polygon, it can be determined as mass center of black points on horizontal projection (1). Thus, using the horizontal projection, we can get the desired vertical projection, which we could use as the image for recognition. However, to increase the probability of correct recognition we could "improve the quality" of the resulting image, i.e. increase the density of black pixels. If horizontal projection is a circle, we could rotate the section plane λ (for vertical projection) that pass through the center of circle, with a certain step α (for example, 1 degree) and project this section on the plane δ so that the position of the line corresponding to the center of the circle remain constant (Fig. 2a). When the horizontal projection is a rectangle, we could move the section plane λ (for vertical projection) along the side, that is perpendicular to this section, with a certain step h (for example, 1 pixel). The vertical projection plane δ remain constant during moving of the section plane λ (Fig. 2b).

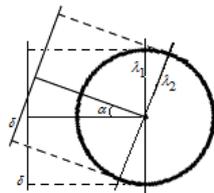


Fig. 2a: Increasing density for circle projection

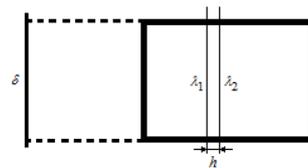


Fig. 2b: Increasing density for rectangle projection

3 Image recognition using counter propagation neural network

3.1 Outline of the method

Counter propagation neural network operates on the principle "winner takes all". It enables to apply clustering of input images and on the output it supplies the number of cluster for an input image. Counter propagation neural network consists of two layers: Kohonen layer and Grossberg layer (Fig. 3).

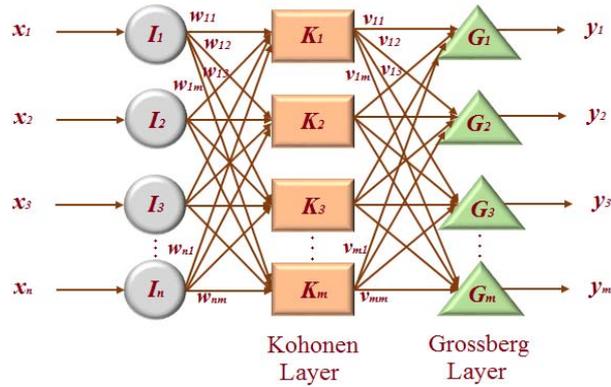


Fig. 3: Structure of Back propagation neural network

Values of Kohonen layer neurons are calculated as follows:

$$K_i = \sum_{j=1}^n x_j w_{ji}, \quad i = 1, 2, \dots, m. \quad (\text{Eq. 1})$$

A neuron with maximum value considered as "winner" neuron and on the output it supplies 1, other neurons supply 0 to the output. Let K_1, K_2, \dots, K_m be the output values Kohonen layer. Grossberg layer allows us to match the "winner" neuron with the recognized cluster of objects. Values of Grossberg layer neurons are calculated using following formula:

$$G_i = \sum_{j=1}^m k_j v_{ji}, \quad i = 1, 2, \dots, m. \quad (\text{Eq. 2})$$

In our application number of neurons in Kohonen and Grossberg layers are equal. First we train the Kohonen layer until each object cluster will activate only one unique neuron. Training of Kohonen layer is unsupervised training. For training we use formula:

$$w_{ij}^* = w_{ij} + \eta(x_i - w_{ij}), \quad i = 1, 2, \dots, n. \quad (\text{Eq. 3})$$

Where w_{ij}^* — new weight value, w_{ij} — current weight value, η — speed of training ($\eta < 1$). Then we train the Grossberg layer. We supply to the input of the network image from a specific cluster of objects, it activates corresponding neuron on Kohonen layer, then we set the weights for that neuron in the Grossberg layer so that on the output our network they activates neuron with the number of the corresponding cluster. In our case all weights for "winner" neuron will be equal to 0 except for the weight v_{ij} that is equal to 1, where i is the number of neuron winner, and j is the number of corresponding cluster.

To distinguish images supplied to the input of the counter propagation neural network, it is necessary to find characteristics that will allow differentiating one image from another. For this purpose we decide to use "characteristic" points, which we determine with the help of two following ways:

1. Calculation the mass center and searching for two of the most nearest and two of the most distant points from the mass center.
2. Searching for critical points, i.e. points that have slope of the curve close to 0 or 90 degrees.

Let us review each of the methods in detail.

3.2 Calculation the mass center and searching for two of the most nearest and two of the most distant points from the mass center

The choice of this method was determined by the fact that the distribution of points on a projection of the resulting three-dimensional point cloud and the original image must be approximately the same. In theory, the projection should be different from the original image only by density of points, but the pattern of point allocation should be the same.

For calculation of coordinates for mass center of points on the image we used the following formula:

$$x_c = \frac{\sum_{k=1}^n x_k}{n}, \quad y_c = \frac{\sum_{k=1}^n y_k}{n} \quad (\text{Eq. 4})$$

Where (x_k, y_k) is the coordinates of black pixels on the image, and n is the number of black pixels in the image. Here and after by coordinates x and y we mean number of column and number of row on the intersection of which the pixel of interested is located. Numbering of columns and rows begin in the left low corner of image. Pixel in left low corner has coordinates $(1;1)$.

Upon calculation of mass center (A, Fig. 4b) for black pixels on the image we search for 2 black pixels that are closest to the mass center (D and C, Fig. 4b) and 2 black pixels that are farthest from the mass center (D and E, Fig 4b), we also calculate the distances to those pixels from mass center. Moreover, when searching for the second closest or farthest pixels we do not take in account pixels from the circular neighborhood of the first corresponding point (Fig. 4b). The neighborhood radius is predefined (for example, 5% of the diagonal length of the image supplied to the input). For all received coordinates and distances we perform normalization, i.e. we divide each value by the length of the image diagonal. Obtained coordinates and lengths we supply to the input of the neural network in specific order.

The advantages of this method:

- It does not require the processing of original image except cropping on marginal black pixels;
- Density of points on the recognizable image may be less than the density of points on the reference image;
- Selected characteristic points allow differentiating a wide range of images.

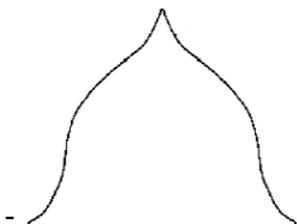


Fig. 4a: Example of recognized image

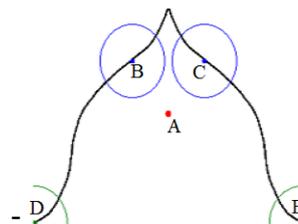


Fig. 4b: The image with mass center and related points

3.3 Searching for critical points

With “critical points” we understand points on the pattern and input data, which have slope of the curve close to 0 or 90 degrees. This method is also suitable to us, since regardless of the distribution density of pixels on the reference image and the projection, the location of the critical points and its number pro cut (object) should be approximately the same.

For each black pixel of the image we search for two closest black pixels and consider two pairs of pixels $B(x_0, y_0)$, $A(x_1, y_1)$ and $B(x_0, y_0)$, $C(x_2, y_2)$ where $B(x_0, y_0)$ —point of interest, $A(x_1, y_1)$ and $C(x_2, y_2)$ —two closest points (Fig. 4a). First, we calculate the slope coefficients of lines that pass through the selected pair of pixels:

$$k_1 = \frac{y_1 - y_0}{x_1 - x_0}, \quad k_2 = \frac{y_2 - y_0}{x_2 - x_0} \quad (\text{Eq. 5})$$

Then, with the help of the coefficients, we calculate the angle of slope (Fig. 5b and 5c):

$$\alpha_1 = \arctg(k_1), \quad \alpha_2 = \arctg(k_2) \quad (\text{Eq. 6})$$

Then calculate the angle for the point of interest as the average of those two angles of slope (Fig. 4d):

$$\alpha_0 = \frac{\alpha_1 + \alpha_2}{2} \quad (\text{Eq. 7})$$

For each image we search for pixels, slope angle values of which are close to 0 or 90 degrees. We consider those points as critical and use them as characteristic points to compare images. Coordinates of such points we supply to the input of neural network in specific order. The number of such points is predefined.

Advantages of this method:

- It does not require the processing of original image except cropping on marginal black pixels;
- Density of points on the recognizable image may be less than the density of points on the reference image;
- Selected characteristic points allow differentiating a wide range of images.

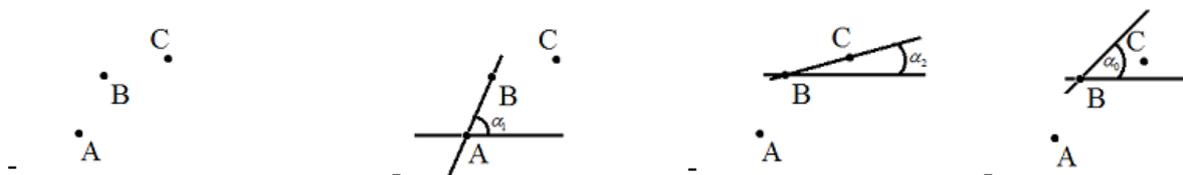


Fig. 5a: Three nearest points Fig. 5b: Obtaining α_1 Fig. 5c: Obtaining α_2 Fig. 5d: Obtaining α_0

4 Pattern recognition using perceptron

4.1 Outline of the method

Perceptron is one of the simplest neural networks, which allows variety of applications.

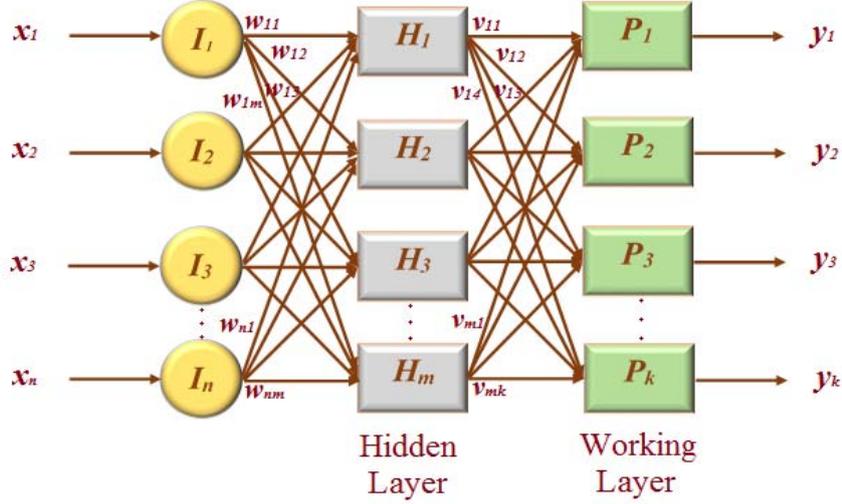


Fig. 6: Structure of perceptron with the hidden layer

Here we decide to use the hidden layer with the predefined number of neurons in it to increase an accuracy of image recognition. Values for neurons in the hidden layer are calculated as follows:

$$H_i = \sum_{j=1}^n x_j w_{ji}, \quad i = 1, 2, \dots, m. \quad (\text{Eq. 8})$$

Then, to determine output values of hidden layer we apply function f to each neuron:

$$h_i = f(H_i), \quad i = 1, 2, \dots, m, \quad (\text{Eq. 9})$$

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (\text{Eq. 10})$$

Calculating of output values for working layer neurons is similar:

$$P_i = \sum_{j=1}^m h_j v_{ji}, \quad i = 1, 2, \dots, k, \quad (\text{Eq. 11})$$

$$y_i = f(P_i), \quad i = 1, 2, \dots, k. \quad (\text{Eq. 12})$$

Perceptron training is supervised. We supply data from reference images to the input of perceptron and make an adjustments of weights based on deviation from intended output. Weights adjustment is made using back propagation training rule:

$$\delta_i^{(1)} = z_i - y_i, \quad i = 1, 2, \dots, k, \quad (\text{Eq. 13})$$

$$\delta_i^{(2)} = \sum_{j=1}^k v_{ij} \delta_j, \quad i = 1, 2, \dots, m, \quad (\text{Eq. 14})$$

$$w_{ij}^* = w_{ij} + \eta \delta_j^{(2)} (h_j (1 - h_j)) x_i, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m, \quad (\text{Eq. 15})$$

$$v_{ij}^* = v_{ij} + \eta \delta_j^{(1)} (y_j (1 - y_j)) h_i, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, k. \quad (\text{Eq. 16})$$

Where z_i — intended output, $\delta_i^{(1)}$ — errors of working layer, $\delta_i^{(2)}$ — errors of hidden layer, w_{ij}^* and v_{ij}^* — new weight values, η — speed of training ($\eta < 1$).

In this case to the input of perceptron we supply a matrix with elements that equal 0 or 1. We will use adjacency matrix as an input, obtaining of which will be described below. The number of neurons in our working layer of perceptron coincide with the number of reference images. At the output we get 0 values from all neurons except the one, number of which, corresponds to cluster of recognized image.

4.2. Dividing the image into rectangles, adjacency of which is determined based on connected pixels

The point of this method is to divide an image into predetermined number of rectangles (for example, 400 (20x20) rectangles). All rectangles have the same size. The width and height of the rectangles are calculated as follows:

$$w_s = \frac{w_i}{n}, \quad h_s = \frac{h_i}{n}. \quad (\text{Eq. 17})$$

Where w_i и h_i are width and height of the image, respectively, and $n \times n$ is predetermined number of rectangles. To get rectangles width and height as integers we add missing rows and columns of white pixels to the image edges.

To determine connection between pixels we start to move along the bottom border of the lower left corner of the image, we search for first pixel. As we find it we connect it to the nearest pixel. Next for this nearest pixel we also search for the nearest pixel, excluding from our search the first pixel, and so on. Thus, we consistently connect all black pixels on the image in one curve. Two neighboring rectangles (each rectangle have only 8 neighboring rectangles) are considered adjacent if there are two connected points, one of which belongs to the one rectangle, and the other — to the other rectangle. So we define the adjacency of neighboring rectangles (Fig. 7c). Then we create adjacency matrix of these rectangles as follows. We enumerate all rectangles from 1 to k where $k = n \times n$. Then we take a zero matrix with the size $k \times k$, and write in it 1 at the intersection of i -th row and the j -th column, if rectangles with numbers i and j are adjacent. Obviously, such a matrix would be symmetric. This matrix we supply to the input of the perceptron.

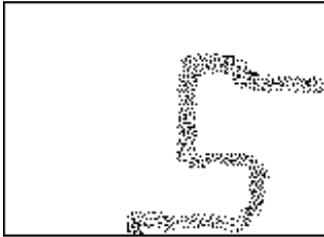


Fig. 7a: Example of image prior processing

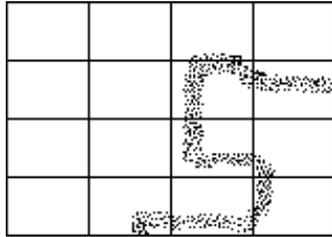


Fig. 7b: Image divided on equal rectangles

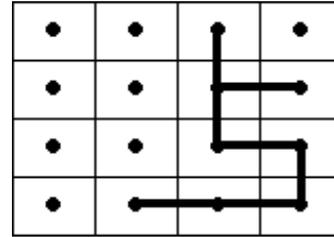


Fig. 7c: Interpretation of adjacency matrix after image processing

Advantages of this method:

- It does not require the processing of original image except cropping on marginal black pixels;
- Density of points on the recognizable image may be less than the density of points on the reference image;
- It is less sensitive to minor changes (within the same rectangle) and preserves the overall structure of the image;
- Selected characteristic points allow differentiating a wide range of images.

5 Results

To test our methods we require some training set. We simulated vertical projections of corresponding point clouds. The only way to do it is to change reference images somehow to make them similar to projections from 3D point cloud. We decide to take black pixels from reference images and change their location with the help of probability theory (Fig. 8a and 8b). We predefine to argument p and r . First argument is responsible for density of pixels, second argument is responsible for location change. For every black pixel on the image we generate random variable V that is evenly distributed on the interval $[0;1]$. If $V < p$ than we changed the location for the pixel of interest to $(x_0 + t, y_0 + t)$ where t is integer random variable that is evenly distributed on the interval $[-r; r]$. If $V \geq p$ then we delete the pixel of interest, i.e. replace it with white pixel. So if $V = 1$ we do not delete any black pixel, if $V = 0$ we delete all black pixels. For test and training sets we use for p integer values from 50 to 100, and for r we use integer values from 1 to 4.

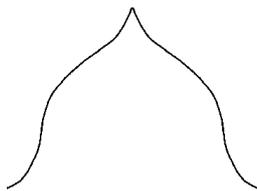


Fig. 8a: Example of reference image



Fig. 8b: Image after simulation vertical projection with $p=50, r=4$

The perceptron showed the best results. For domes we divided images on 20x20 rectangles and used perceptron with 28 neurons in hidden layer, it recognized correctly 13 of 14 images at average. For roofs we divided images on 20x20 rectangles and used perceptron with 15 neurons in

hidden layer, it sometimes made only 1 mistake for roofs generated with $r=4$, but mostly it recognized all roofs with no mistake. For crosses we divided images on 20x20 rectangles and used perceptron with 15 neurons in hidden layer, it recognized all crosses with no mistake.

Counter propagation neural network and mass center calculation method also showed good results. At average for domes 12 of 14, for crosses 6 of 7 and for roofs 5 of 6 images were recognized correctly. For the neighborhood radius we used the value equal to 10% of the diagonal length of the image.

Unfortunately, counter propagation neural network and critical points method showed poor results. For reference images it allows to correctly recognize all the images but for simulated images it mistakes for 3 and more images for every structural element. It could be explained by the fact that for proper angle of slope calculation curves on the image should have width of 1 or 2 pixels. Otherwise proposed method of slope angle calculation shows wrong results.

6 Conclusion

Obtained results tell us that not always we could determine structural element of the church correctly with 100% accuracy. But proposed methods showed their efficiency and more than that it is only the first steps and there is some room for further enhancements. For example, new algorithm for slope angle calculation in counter propagation neural network and critical points method could be developed. If we would be able to differentiate outer pixels out of others those method could be successful. As for counter propagation neural network and mass center calculation method there could be developed some algorithms that resolve issues with unsymmetrical results (for example, if point B would be in the uppermost black pixel, point C with equal possibility could be in the left side and on the right side of the image; Fig. 3b). And for perceptron there are plenty of opportunities. We could increase the number of hidden layers, we could change neuron activation functions and we could enhance method of rectangle adjacency determine.

7 Acknowledgements

We would like to thank the students of Ivanovo State Power Engineering University – Anatolii Bolshakov and Maksim Lobanov - for program realization of our algorithms.

8 References

- AHN, S., 2004: Least Squares Orthogonal Distance Fitting of Curves and Surfaces in Space. *Lecture Notes in Computer Science* **3151**, Springer Berlin Heidelberg, 125 pages.
- 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.
- ALEXANDRE, L., 2014: 3D Object Recognition using Convolutional Neural Networks with Transfer Learning between Input Channels. *Intelligent Autonomous Systems* **13**, Springer International Publishing, 889-898.
- CANCIANI, M., FALCOLINI, C., SACCONI, M., & SPADAFORA, G., 2013: From Point Clouds to Architectural Models: Algorithms for Shape Reconstruction. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (5/W1), 27-34.
- CICHY, R. M., KHOSLA, A., PANTAZIS, D., TORRALBA, A., & OLIVA, A., 2016: Deep Neural Networks Predict Hierarchical Spatio-temporal Cortical Dynamics of Human Visual Object Recognition. arXiv:1601.02970, <http://arxiv.org/pdf/1601.02970v1.pdf>.
- HUBER, P., 1964: Robust Estimation of a Location Parameter. *Annals of Mathematical Statistics* **35** (1), 73-101.
- KOROVIN, D., CHIZHOVA, M., BRUNN, A. & STILLA, U., 2016: Probabilistic Reconstruction of 3D Buildings using Bayesian Nets. Submitted manuscript for the journal *Photogrammetrie Fernerkundung Geoinformation (PFG)*.
- MARSHALL, D., LUKACS, G. & MARTIN, R., 2001: Robust Segmentation of Primitives from Range Data in the Presence of Geometric Degeneracy. *IEEE Transactions in pattern analysis and machine intelligence* **23** (3), 304-314.
- RABBANI, T. & VAN DEN HEUVEL, F., 2005: Efficient Hough Transform for Automatic Detection of Cylinders in Point Clouds. *Proceedings of the ISPRS Workshop Laser scanning 2005*, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences Volume* **36** (3/W19), 60-65
- ROSENBLATT, F., 1957: The Perceptron: a perceiving and recognizing automaton. Report 85-460-1, Cornell Aeronautical Laboratory.
- RUMELHART, D.E. & MCCLELLAND, J.L., 1986: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Volume 1 – Foundations, MIT Press Cambridge MA, USA.
- SCHNABEL, R., WAHL, R. & KLEIN, R., 2007: Efficient Ransac for Point-Cloud Shape Detection. *Computer Graphics Forum* **26** (2), 214-226.
- VOSSELMAN, G., GORTE, B., SITHOLE, G. & RABBANI, T., 2004: Recognizing Structure in Laser Scanner Point Clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **46**, 33-38.
- YUMER, M. & KARA, L., 2012: Surface creation on unstructured point sets using neural networks. *Computer-Aided Design* **44**, (7), July, 644-656.

Robuste Superpixel-basierte Segmentierung von spektralen Fernerkundungsaufnahmen zur Überwachung der Stadtentwicklung anhand von Straßennetzen

BENJAMIN SEPPKE¹, LEONIE DRESCHLER-FISCHER¹ & CHRISTIAN WILMS¹

Die Überwachung der Stadtentwicklung anhand von Fernerkundungsbildern, die im sichtbaren spektralen Bereich aufgenommen wurden, besitzt im Rahmen der Bildverarbeitung und Szenenanalyse bereits eine lange Tradition. Leider führen gut erforschte, generische Segmentierungsverfahren oft zu nicht-zufriedenstellenden Lösungen. Ein Weg, um dieses Problem zu umgehen, ist der Wechsel der Basisentitäten von Pixeln hin zu Superpixeln, die jeweils lokal kohärente Cluster zusammenhängender Pixel beschreiben. Obwohl diese Superpixel generell flexiblere Algorithmen ermöglichen, benötigen sie dennoch eine Repräsentation, die sowohl geometrische als auch topologische Merkmale beinhaltet, verfügbar macht, und gleichzeitig Operationen, wie z.B. das Verschmelzen von Superpixeln, unterstützt. In dieser Arbeit stellen wir ein derartiges, kombiniertes graphenbasiertes Rahmenwerk vor, den sog. RS-Graphen, und zeigen dessen Verwendung am Beispiel der halbautomatischen Straßenverfolgung. Als Bilddaten wurde für diese Studien auf öffentlich verfügbare Daten zurückgegriffen. Die Anwendung der Superpixel-Methode liefert, verglichen mit anderen Ansätzen, wie z.B. SHUKLA et al. 2002, hervorragende Ergebnisse. So ist das Verfahren beispielsweise robust gegenüber teilweisen Verdeckungen der aufgenommenen Straßen.

1 Einleitung

Die Segmentierung von Fernerkundungsbildern im sichtbaren spektralen Bereich in mehrere Zusammenhangskomponenten besitzt in der Photogrammetrischen Bildauswertung bereits eine lange Tradition. Zudem stellt sie häufig den Ausgangspunkt der Untersuchungen in Bereichen der allgemeinen Bildverarbeitung, der Bildanalyse und der Szeneninterpretation dar. Dennoch kann es je nach Anwendungsfall nur schwer möglich oder sogar unmöglich sein, die gewünschte Segmentierung der Bildsignaturen in einem Verfahrensschritt zu erhalten. Gründe hierfür sind z.B. wechselnde Farbwerte oder Verschattungen der abgebildeten Objekte. Zieht man als Beispiel die Abbildungen von Straßen oder Flüssen heran, so sind die spektralen Signaturen zwar lokal häufig recht ähnlich, können sich global jedoch deutlich voneinander unterscheiden.

Um diese Probleme zu umgehen, stellen wir in dieser Arbeit einen Superpixel-basierten Ansatz vor, der es erlaubt Straßennetze halbautomatisch aus Fernerkundungsbildern zu gewinnen. Der Ansatz beginnt zunächst mit einem übersegmentierten Bild, dessen einzelne Regionen wir als Superpixel benennen. Superpixel bezeichnen im allgemeinen lokal kohärente Cluster benachbarter Pixel, die die (meiste) notwendige Struktur der Segmentierung auf einem zu wählenden Maßstab bewahren (vgl. REN & MALIK 2003). Basierend auf einer neuartigen

¹ Universität Hamburg, MIN-Fakultät, Fachbereich Informatik, Arbeitsgruppe: Szenenanalyse und Visualisierung (SAV), Vogt-Kölln-Str. 30, D-22527 Hamburg, E-Mail: [seppke, dreschler, wilms]@informatik.uni-hamburg.de

geometrisch-topologischen Repräsentation dieser Superpixel, dem sog. RS-Graphen, stellen wir eine Methode zur halbautomatischen Verfolgung von Straßenzügen dar, die eine Verfolgung der Straßensignaturen von Superpixel zu Superpixel ermöglicht. Da Straßen sich hervorragend zur Klassifizierung und Überwachung von urbanen Gebieten eignen, erhalten wir somit einen wichtigen Indikator für die Entwicklung ebensolcher Gebiete. Zudem ist das Verfahren so angelegt, dass es robust gegenüber lokalen Verdeckungen der Straße ist, und eine parallele Verarbeitung erlaubt, um die Ausführung zu beschleunigen und moderne Hardware-Architekturen bestmöglich zu unterstützen.

2 Verwandte Arbeiten

Sowohl die Repräsentation von Bildsegmentierungen als auch die Extraktion und Verfolgung von Straßen in Fernerkundungsbildern werden schon seit geraumer Zeit erforscht. Daher erfolgt in diesem Abschnitt zunächst eine Einführung in die wichtigsten Aspekte beider Gebiete.

2.1 Geometrisch-topologische Repräsentation von Bildsegmentierungen

Nach der initialen Segmentierung eines Bildes in Superpixel ist die naive Repräsentation ein sogenanntes Label-Bild. Hierbei handelt es sich um ein einkanaliges Bild, wobei der Wert jedes Bildpunktes die ID des zugehörigen Superpixels widerspiegelt. Ein solches Bild wird üblicherweise direkt von den Segmentierungsverfahren erzeugt. Es kann allerdings auch nachträglich im Rahmen des Findens der Zusammenhangskomponenten generiert werden. Diese Art der Repräsentation ist allerdings wenig wertvoll, da die einzige direkt zugreifbare Information die Zugehörigkeit von Pixeln zu Superpixeln ist. Falls Fragestellungen, wie z.B. „Welche Superpixel sind direkte Nachbarn von Superpixel 124?“, beantwortet werden sollen, sind erweiterte geometrisch-topologische Repräsentationen nötig.

Eine der ersten derartigen Repräsentationen beschreibt der sog. Regionen-Nachbarschafts-Graph (engl. Region Adjacency Graph, kurz RAG). Dieser ist definiert als ein Graph, in dem die Knoten den Superpixeln entsprechen und die Kanten die Relation „benachbart“ abbilden (vgl. PAVLIDIS 1977). Aufgrund seiner Kompaktheit und einfachen Verständlichkeit findet der RAG auch aktuell noch Anwendung etwa bei Segmentierungsverfahren (vgl. FELZENSZWALB & HUTTENLOCHER 2004 und KROPATSCH 1995). In Bezug auf die möglichen Fragestellungen bleibt der RAG aufgrund seiner Definition allerdings auf Nachbarschaftsanfragen beschränkt.

Um die Beschränkungen des RAG zu überwinden, existieren generell zwei verschiedene Vorgehensweisen. So hat KROPATSCH (1995) eine Repräsentation vorgeschlagen, die Duale Bildgraphen zur Charakterisierung von Superpixeln verwendet. Diese kombinieren, korrigieren und erweitern den Ansatz des RAG mit einem hierzu dualen Graphen.

Einen gänzlich anderen Ansatz verfolgen hingegen die sogenannten Kombinatorischen Karten, die Permutationen verwenden, um die Segmentierung darzustellen. So beschreiben beispielsweise KÖTHE (2002) und MEINE (2009), welche Möglichkeiten und Einschränkungen bei dieser Art der Repräsentation vorliegen. Obwohl beide Arten der Darstellung, die Dualen Graphen und die Kombinatorischen Karten, in der Lage sind die geometrisch-topologischen Zusammenhänge zu beschreiben, sind sie dennoch limitiert. Teilweise stellen sie nicht alle Informationen explizit bereit, führen Hilfskonstrukte ein oder benötigen sehr komplexe

Modellierungen von Verfahren zur Manipulation von Superpixeln, wie z.B. die Verschmelzung zweier Superpixel.

2.2 Straßenverfolgung anhand von spektralen Signaturen

Pixel-basierte Verfahren zur automatischen und halbautomatischen Verfolgung und Extraktion von Straßen anhand spektraler Signaturen auf Fernerkundungsbildern werden bereits seit einiger Zeit erforscht. Einen umfassenden Überblick der verschiedenen Klassen von Verfahren sowie der Verfahren selbst beschreiben MENA (2003) und LI et al. (2009).

Als einen Stellvertreter der Klasse der modellbasierten Verfahren beschreiben GEMAN & JEDYNAK (1996) einen statistisch motivierten Ansatz, der in der Lage ist halbautomatisch Straßen in spektralen Bilddaten zu verfolgen. Hierbei muss zunächst durch den Benutzer ein Ausgangspunkt auf der Straße und eine zu verfolgende Richtung vorgegeben werden. Zudem müssen zwei Wahrscheinlichkeitsdichtefunktionen angegeben werden: eine für die spektrale Verteilung aller Intensitätswerte der Klasse „Straße“ und eine für den negativen Fall. Da diese Wahrscheinlichkeitsverteilungen nicht vorliegen, werden sie oft anhand von Trainingsdaten und überwachten Lernverfahren ermittelt (vgl. GEMAN & JEDYNAK 1996). Laut den Autoren führt diese Vorgehensweise zu vielversprechenden Resultaten. So waren sie unter anderem in der Lage hunderte Kilometer Autobahn zu verfolgen.

Einen anderen Ansatz, der speziell für hochaufgelöste Bilddaten entwickelt worden ist, beschreiben SHUKLA et al. (2002). Wie der bereits erwähnte Ansatz benötigt auch dieser halbautomatische Ansatz als Vorgabe einen Bildpunkt und eine Richtung, in der die Straße verfolgt werden soll. Allerdings benötigt dieses Verfahren kein a priori Wissen über die Wahrscheinlichkeiten der einzelnen spektralen Anteile. Stattdessen wird in diesem Verfahren eine Kostenfunktion definiert, die beschreibt, mit welchen Kosten ein Folgepixel ebenfalls zur Straße gehört. Die Kostenfunktion setzt sich hierbei aus einer Vielzahl von Unterfunktionen zusammen. Die Extraktion der Straße stellt sich anschließend als ein Minimierungsproblem dar: Die minimalen Kosten korrespondieren zu dem extrahierten Pfad, der im Idealfall der abgebildeten Straße entspricht. Ein Hauptkritikpunkt dieses Verfahrens ist die hohe Komplexität und die Verwendung einer Menge von anderen Verfahren und Heuristiken, deren Parameter erst ermittelt werden müssen. So werden beispielsweise mit dem Canny-Verfahren (siehe CANNY 86) Kanten erkannt und anhand eines Schwellenwerts bestimmt, wie breit die zu verfolgende Straße ist. Ähnliches gilt für die Verfolgung an Kreuzungen.

Weitere Ansätze zur Straßenverfolgung verwenden Vorlagen von Straßensignaturen oder Texturen derselben und versuchen diese im Bild zu erkennen (vgl. JIN et al. 2006; LIN et al. 2012). Eine weitere populäre Klasse von Verfahren beschreibt die Straßenverfolgung anhand einer initialen Position und einem Startwinkel sowie einem linearen Änderungsmodell von Bildpunkt zu Bildpunkt. Dieses Modell wird mit einem Kalman-Filter ausgewertet, was ebenfalls in einer Straßenverfolgung resultiert (vgl. ZHOU et al. 2006; MOVAGHATI et al. 2008). Darüber hinaus führen BAUMGARTNER et al. (2002) im Rahmen der Kalman-Filterung orthogonale Profile zur Straßenverfolgung ein.

Anstatt Straßen auf der Pixelebene zu verfolgen bzw. zu extrahieren, erscheint es sinnvoll, die Extraktion auf einer höheren Ebene zu vollziehen, z.B. auf der Ebene der Superpixel. Viele der möglicherweise benötigten statistischen Eigenschaften liegen hier bereits explizit vor. Allerdings

müsste hierfür zunächst eine geeignete Repräsentation zum interaktiven Umgang mit Superpixeln geschaffen werden.

3 Superpixel-Repräsentation mit dem RS-Graph

Um eine Menge von Superpixeln angemessen repräsentieren zu können, wurde in WILMS (2016) eine neuartige Repräsentationsform eingeführt: der sogenannte Regionen-Schalen-Graph (RS-Graph). Im Gegensatz zum RAG, der lediglich die Nachbarschaften der einzelnen Superpixel modelliert, oder anderen Verfahren, wie z.B. MEINE (2009), repräsentiert der RS-Graph nicht nur die Regionen, sondern auch die Schalen an den Kanten der jeweiligen Regionen explizit als Knoten im Graph. Hierbei werden die Bildpunkte an den Grenzen zwischen Superpixeln als Schalen der jeweiligen Regionen bezeichnet. Jeder Superpixel besitzt hierbei jeweils eine eigene Schale zu jedem benachbarten Superpixel, die innerhalb des Superpixels liegt. Eine beispielhafte grafische Darstellung eines RS-Graphen ist in Abbildung 1 gegeben.

Durch die explizite Repräsentation von Regionen und deren Begrenzungen steht mit dem RS-Graphen ein nützliches Werkzeug für eine Vielzahl von Anwendungen zur Verfügung. Zudem führt diese Art der Repräsentation noch zu einem weiteren Nebeneffekt: Durch die reichhaltige Beschreibung ist mit einem RS-Graphen nicht nur die Geometrie, sondern auch die Topologie eindeutig beschrieben. Wo andere Verfahren, wie z.B. KROPATSCH (1995) und KÖTHE (2002), künstliche Konstrukte benötigen, besitzt der RS-Graph eine einheitliche und geschlossene Beschreibung.

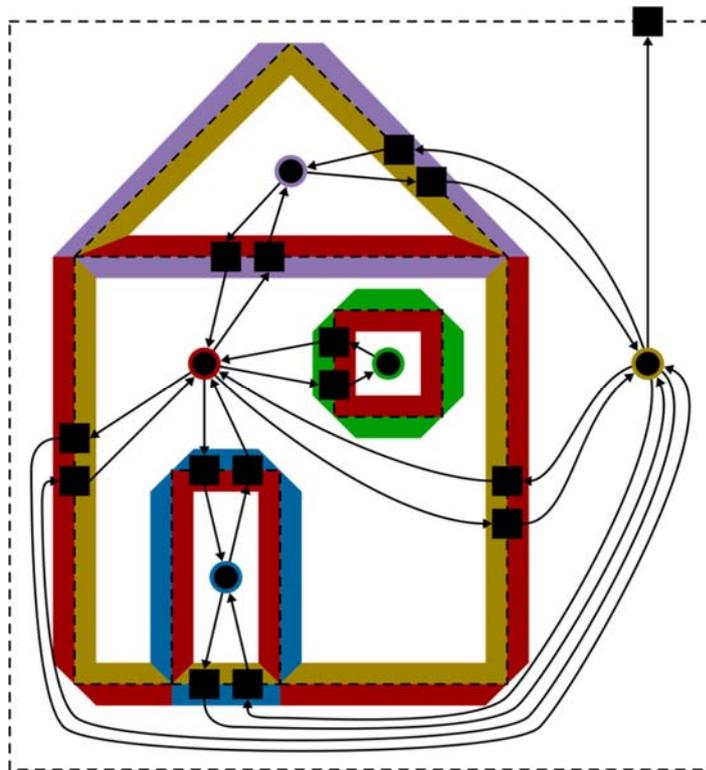


Abb. 1: Beispiel eines RS-Graphen zur Segmentierung eines Hauses anhand von fünf Superpixeln (Punkte). Die Schalen, welche die Superpixel voneinander trennen, sind als Quadrate im Graphen zu erkennen. Oben rechts: Äußere Schale zur Bildbegrenzung.

Weiterhin wurde in WILMS (2016) mithilfe von Verschmelzungsoperationen eine vollständig navigierbare irreguläre Pyramide von Superpixeln implementiert und somit auch verwendbar gemacht. Um die Erstellung der Datenstruktur des RS-Graph anhand eines Label-Bildes zu beschleunigen, wurden einige Schritte parallelisiert sowie auf die Grafikkarte ausgelagert. Somit wird das Potential aktueller Hardware bestmöglich ausgenutzt. Eine detaillierte Beschreibung der Erstellung des RS-Graphen findet sich in WILMS (2016).

4 Superpixel-basierte Straßenverfolgung mithilfe des RS-Graphen

Um die Vielseitigkeit des RS-Graphen zu zeigen, stellen wir im Folgenden eine Vorgehensweise zur Verfolgung von Straßen auf Basis dieser Repräsentation vor. Analog zu den halbautomatischen Verfahren, die in Abschnitt 2.2 beschrieben worden sind, gibt der Benutzer auch bei diesem vorgeschlagenen Verfahren einen Startpunkt und eine Richtung vor. Zudem sollte die Breite der Straße (in Pixel) angegeben werden. Im Anschluss daran wird zunächst das SLIC-Verfahren verwendet, um eine Segmentierung durchzuführen (siehe ACHANTA et al. 2004). Aufbauend auf den so entstandenen Superpixeln wird der RS-Graph erstellt.

Die spektralen Werte der Bildpunkte um den Startpunkt herum werden als initiale Farbwerte der Klasse „Straße“ interpretiert. Anschließend startet das Verfahren ausgehend von dem Superpixel, in dem der Startpunkt liegt. Die Suche nach den folgenden (besten) Verlängerungen der Straße erfolgt analog zu SHUKLA et al. (2002) anhand einer Kostenfunktion c , die sich aus drei Teilen zusammensetzt: Die Funktion c_{angle} bewertet den Winkel des Verlaufs von einem Superpixel zum nächsten, und bestraft zu enge Kurven. Die Funktion c_{border} verwendet die Schaleninformationen des RS-Graphen um festzustellen, ob die Straße überhaupt in die Richtungen der jeweils benachbarten Superpixel weiterführt. Schließlich vergleicht die Funktion c_{region} die spektrale Signatur der benachbarten Superpixel (jeweils die ähnlichsten 20%) mit der der Klasse „Straße“.

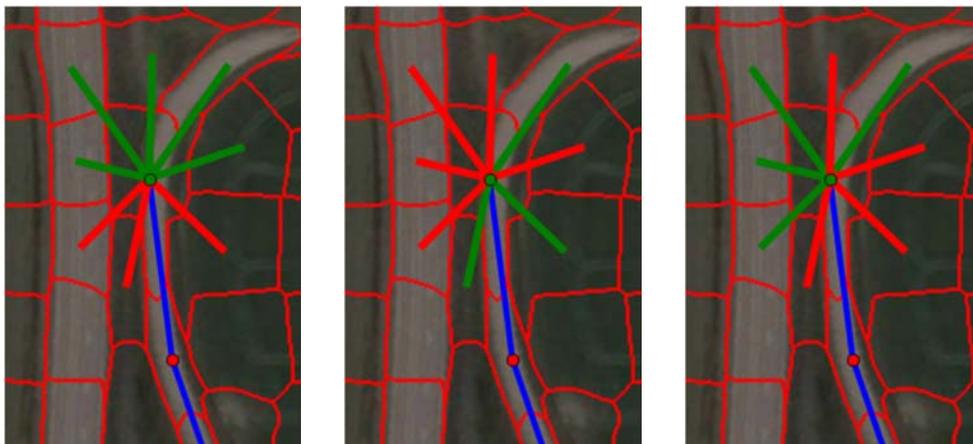


Abb. 2: Ergebnisse der Teilkostenfunktionen, die in der Summe die Kostenfunktion c bestimmen. Von links nach rechts: c_{angle} , c_{border} , c_{region} . Rot: Verlängerung der Straße nicht möglich. Grün: Verlängerung der Straße möglich.

Die einzelnen Kostenwerte werden summiert und führen somit zu einer Rangordnung der möglichen Verlängerungen. Anschließend wird jeweils der Superpixel mit den geringsten Kosten ausgewählt, sofern dieser existiert, andernfalls wird das Verfahren abgebrochen. Falls es mehr als einen Kandidaten zur Verlängerung gibt, werden die Alternativen gespeichert und zur späteren Abzweigungsverfolgung zurückgestellt.

Um die Robustheit des vorgeschlagenen Ansatzes zu verbessern, ist es zudem möglich auch indirekte Superpixel-Nachbarn mit in die Suche einzubeziehen, falls keine passende Verlängerung gefunden werden konnte. Dies ermöglicht auch lokal begrenzte Bereiche, in denen die Straße nicht sichtbar bzw. verschattet ist, zu überspringen. Um die Effizienz des Verfahrens zu steigern, wurde das Tracking zudem soweit möglich parallelisiert.

5 Ergebnisse

Um das entwickelte Verfahren zu testen, wird es auf öffentlich verfügbaren Bilddaten (Google Earth) angewendet und getestet. Die Anwendung auf den Bilddaten zeigt insbesondere die Robustheit des Verfahrens gegenüber Verdeckungen aber auch die Wichtigkeit der Verwendung von Schaleninformationen. In Abbildung 3 ist das Ergebnis einer Anwendung gezeigt, in der die Straße sowohl eine Kreuzung als auch zwei Verdeckungen aufweist.

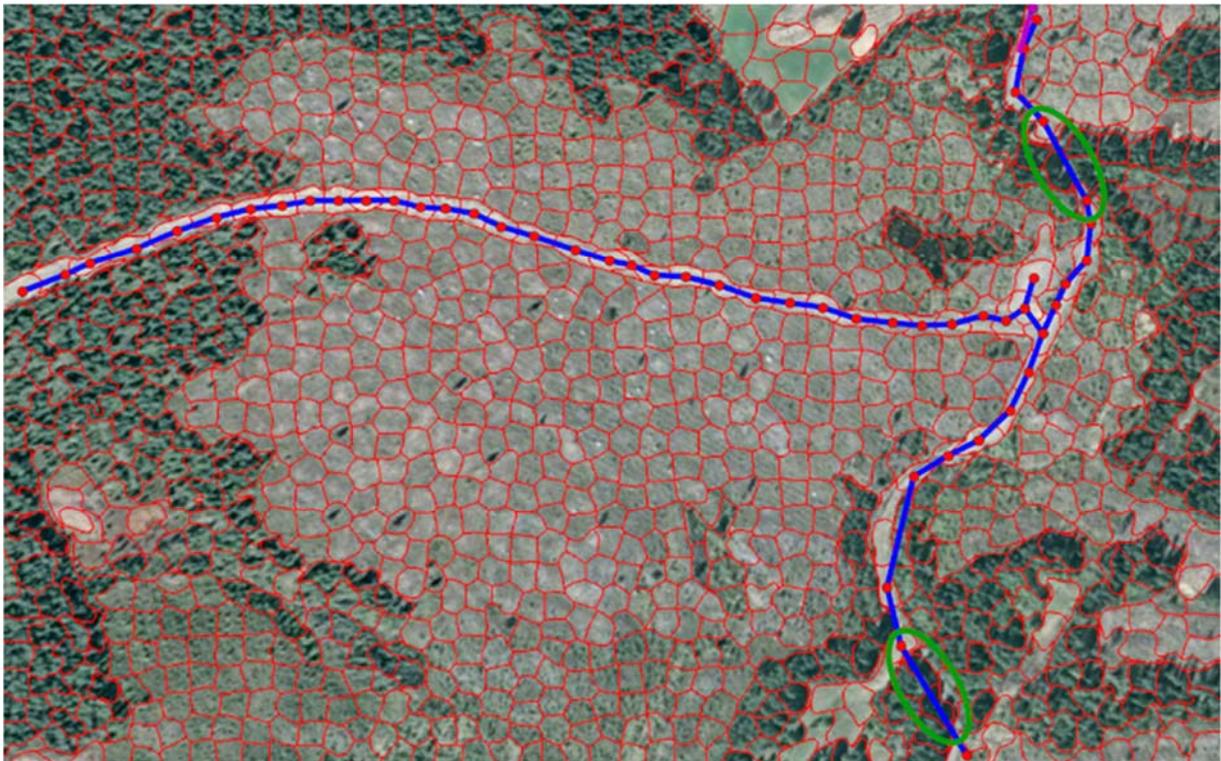


Abb. 3: Ergebnis des vorgeschlagenen Verfahrens zur Straßenverfolgung. Der Startpunkt und die initiale Richtung sind oben rechts (Magenta) markiert. Die Grenzen der SLIC-Superpixel sind in Rot hervorgeben. Die Straße konnte trotz einer Abzweigung und zweier Verdeckungen (Grün) zuverlässig erkannt werden. Bildquelle: Google Earth, Copyright Google Inc.

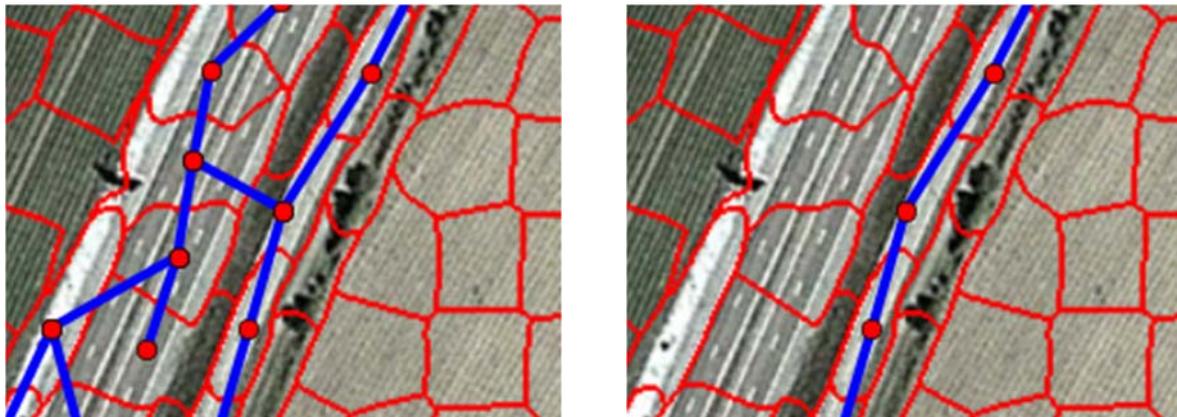


Abb. 4: Einfluss der Verwendung von Schaleninformation in Bezug auf die Kostenfunktion (C_{border}) auf das Resultat der Straßenerkennung. Links: ohne Verwendung von Schaleninformationen, rechts: Ergebnis bei Verwendung von Schaleninformation. Bildquelle: Google Earth, Copyright Google Inc.

In weiteren Untersuchungen wurde das Verfahren zudem in unterschiedlichen ländlichen und städtischen Szenen getestet. Hierbei konnten sowohl unterschiedliche Straßenbeläge als auch Straßen über sich verändernde Umgebungen, z.B. über Brücken hinweg, robust und erfolgreich verfolgt werden.

Um den Mehrwert des vorgeschlagenen Verfahrens in Bezug auf andere Repräsentationsformen wie z.B. den RAG zu erkennen, wird die Kostenfunktion, die auf den Schaleninformationen basiert, nun genauer untersucht. In Abbildung 4 ist ein Teil einer Szene zu erkennen, in der die Straßenverfolgung einmal mit und einmal ohne Zuhilfenahme der Kostenfunktion C_{border} durchgeführt worden ist. Hier wird sofort der große Einfluss der Schaleninformationen auf die Qualität der Straßenverfolgung deutlich. Ohne diese Information über die Trennung zwischen zwei Superpixeln wird die kleine Straße neben der Autobahn während der Verfolgung der Straße fälschlicherweise mit dieser verbunden, da es Superpixel gibt, die über den Grenzbereich hinweg benachbart sind. Nimmt man die Schaleninformationen hinzu, werden derartige Verbindungen unterbunden.

6 Evaluierung

Um die Ergebnisse des vorgeschlagenen Verfahrens quantitativ vergleichen zu können, wurde neben dem vorgestellten Verfahren auch jenes von SHUKLA et al. (2002) implementiert. Dieses ebenfalls kostenbasierte Verfahren stellt als pixelbasiertes Verfahren zur Straßenverfolgung einen guten Kontrast zum vorgeschlagenen Superpixel-basierten Ansatz dar. Für die Evaluierung werden beide Verfahren auf zwei Teilmengen der Testdaten zu je drei Aufnahmen angewendet. Die Bilder der ersten Teilmenge beinhalten allesamt Verdeckungen der Straßen, wohingegen die zweite Teilmenge nur Straßen beinhaltet, die ohne Verdeckungen aufgenommen worden sind. Die Ergebnisse der Evaluierung für diese beiden Teilmengen sind in Abbildung 5 gegeben. Im Gegensatz zu dem Verfahren von SHUKLA et al. 2002 sind hier die Vorteile des vorgeschlagenen Verfahrens sehr deutlich zu erkennen. Solange es keine Verdeckungen auf den Straßen gibt,

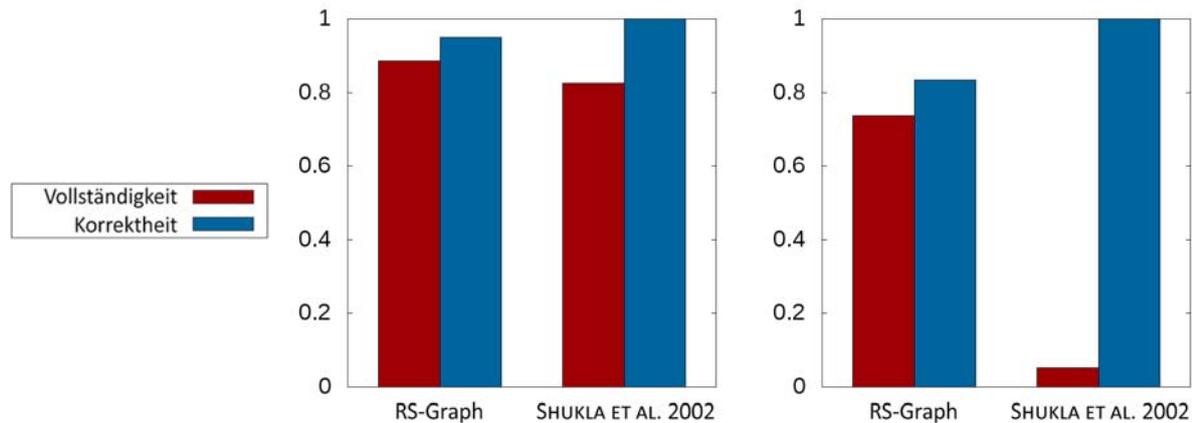


Abb. 5: Vergleich der vorgestellten Methode (RS-Graph) mit dem von SHUKLA et al. (2002) vorgeschlagenen Ansatz zur pixelbasierten Straßenverfolgung auf zwei Testmengen von jeweils drei Bildern. Links: Resultate bei Verwendung von Aufnahmen unverdeckter Straßen, rechts: Ergebnisse bei teilweise verdeckten Straßen.

liegen beide Ansätze etwa gleich auf, sowohl was die Vollständigkeit als auch die Korrektheit betrifft. Allerdings wird hier bereits sichtbar, dass das vorgeschlagene Verfahren stets komplexere Straßenzüge extrahiert – auch wenn diese nicht immer dem Goldstandard (Korrektheit) entsprechen.

Sobald Verdeckungen in den Bilddaten auftreten, wandelt sich dieses Bild noch stärker zu Gunsten des vorgeschlagenen Verfahrens. Zwar extrahiert das Verfahren von SHUKLA et al. (2002) noch einige wenige durchaus korrekte Teile der Straße. Andererseits ist diese Extraktion allerdings in höchstem Maße unbefriedigend, da nur ein verschwindend geringer Teil der Straße ($< 10\%$) extrahiert werden konnte. Hier erzielt das vorgeschlagene Verfahren durch die Möglichkeit des Überspringens von Superpixeln einen deutlich höheren Wert, der nahe an dem Goldstandard liegt.

7 Zusammenfassung und Fazit

In diesem Artikel wurde ein neuartiger, Superpixel-basierter Ansatz zur Verfolgung und Extraktion von Straßen aus spektralen Fernerkundungsbildern präsentiert. Da die Entwicklung von Straßennetzen ein wichtiger Indikator für die städtische Entwicklung ist, ist das Verfahren zudem in der Lage wichtige Daten zur Bestimmung weiterer Kennzahlen bereitzustellen.

Das vorgestellte Verfahren verwendet eine spezielle Art der Repräsentation der Superpixel, den sogenannten RS-Graphen (siehe WILMS 2016). Die Erfordernisse der Anwendung der Straßenverfolgung, die geometrisch-topologische Eigenschaften sowohl über die Superpixel als auch deren Grenzen benötigen, machten diese Art der Repräsentation nötig. Darüber hinaus eignet sich der RS-Graph auch für eine Vielzahl anderer Arten der Merkmalsextraktion (vgl. WILMS 2016).

Die Resultate, die durch das neue Verfahren auf frei verfügbaren Bilddaten erreicht werden konnten, sind allesamt vielversprechend. Im Vergleich mit anderen etablierten Verfahren fällt vor allem die größere Robustheit in Bezug auf die Vollständigkeit auf. Diese kommt

insbesondere dann zum Tragen, wenn die Abbildungen der Straße teilweise Verdeckungen beinhalten. Dieser Vorteil macht das vorgeschlagene Verfahren vor allem in ländlichen bzw. abgelegenen Gebieten besonders gut einsetzbar, da in diesen durch Vegetation natürlicherweise mit Verdeckungen oder Schattenbildung zu rechnen ist. Um die Qualität der Resultate noch weiter zu steigern, könnten zudem weitere Wissensquellen in das Verfahren integriert werden. Da dieses modular angelegt ist, könnte z.B. auch eine Kostenfunktion hinzugefügt werden, die aufgrund von digitalen Geländemodellen einen Ähnlichkeitswert ermittelt. Falls verfügbar, könnten auch bestehende Straßenkarten im Sinne weiterer Kostenfunktionen integriert werden.

8 Danksagung

Diese Forschungsarbeit hat in enger Zusammenarbeit mit dem Weizsäcker Zentrum für Friedensforschung der Universität Hamburg (ZNF) stattgefunden. Ziel der Zusammenarbeit war eine Förderung der Überwachung von Infrastrukturveränderungen in ländlichen Gebieten.

9 Literaturverzeichnis

- ACHANTA, R., SHAJI, A., SMITH, K., LUCCHI, A., FUA, P. & SUSSTRUNK, S., 2004: SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34** (11), 2274-2282.
- BAUMGARTNER, A., HINZ, S. & WIEDEMANN, C., 2002: Efficient methods and interfaces for road tracking. *International Archives of Photogrammetry and Remote Sensing* **34** (3/B), 28-31.
- CANNY, J., 1986: A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8** (6), 679-698.
- FELZENSZWALB, P.F. & HUTTENLOCHER, D.P., 2004: Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision* **59** (2), 167-181.
- GEMAN, D. & JEDYNAK, B., 1996: An active testing model for tracking roads in satellite images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18** (1), 1-14.
- KÖTHE, U., 2002: XPMAPS and Topological Segmentation - A Unified Approach to Finite Topologies in the Plane. *Discrete Geometry for Computer Imagery*, Springer, 22-33.
- KROPATSCH, W., 1995: Building Irregular Pyramids by Dual Graph Contraction. *IEE Proceedings - Vision, Image and Signal Processing* **142** (6), 366-374.
- JIN, X., PASWATERS, S. & DAVIS, C.H., 2006: Road tracking by maximization of mutual information. *Proceedings of the American Society of Photogrammetry and Remote Sensing Annual Conference*, 1-6.
- LI, Y., XU, L. & PIAO, H., 2009: Semi-automatic Road Extraction from High-Resolution Remote Sensing Image: Review and Prospects. *9th International Conference on Hybrid Intelligent Systems* **1**, 204-209.
- LIN, X., SHEN, J. & LIANG, Y., 2012: Semi-Automatic Road Tracking using Parallel Angular Texture Signature. *Intelligent Automation & Soft Computing* **18** (8), 1009-1021.
- MEINE, H., 2009: *The GeoMap Representation: On Topologically Correct Sub-pixel Image Analysis*. Universität Hamburg, Imprint Akademische Verlagsgesellschaft **323**, Dissertations in Artificial Intelligence, IOS Press.

- MENA, J. B., 2003: State of the Art on Automatic Road Extraction for GIS Update: A Novel Classification. *Pattern Recognition Letters* **24** (16), 3037-3058.
- MOVAGHATI, S., MOGHADDAMJOO, A. & TAVAKOLI, A., 2008: Using Unscented Kalman Filter for Road Tracing from Satellite Images. *2nd Asia International Conference on Modeling Simulation*, 379-384.
- PAVLIDIS, T., 1977: *Structural Pattern Recognition*. 1st edition, Springer.
- REN, X. & MALIK, J., 2003: Learning a Classification Model for Segmentation. *9th IEEE International Conference on Computer Vision*, IEEE Computer Society, 10-17.
- SHUKLA, V., CHANDRAKANTH, R. & RAMACHANDRAN, R., 2002: Semi- Automatic Road Extraction Algorithm for High Resolution Images Using Path following Approach. *Proceedings of the Third Indian Conference on Computer Vision, Graphics and Image Processing* **2**, 231-236.
- ZHOU, J., BISCHOF, W. F. & CAELLI, T., 2006: Road Tracking in Aerial Images Based on Human-Computer Interaction and Bayesian Filtering. *ISPRS Journal of Photogrammetry and Remote Sensing* **61** (2), 108-124.

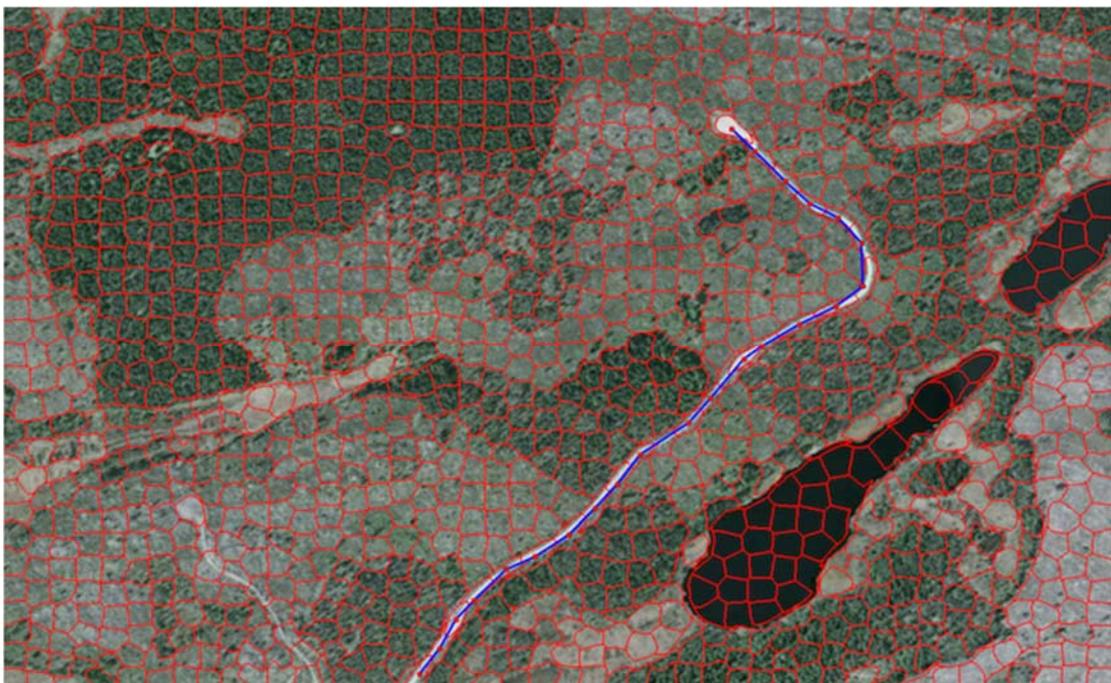
10 Anhang

Im Anhang dieser Arbeit befinden sich zwei weitere Fallbeispiele zur Straßenextraktion. Die Bilddaten stammen, wie die im Text beschriebenen Beispiele, ebenfalls von Google Earth (Copyright: Google Inc.). Neben dem Goldstandard sind zu jedem Fallbeispiel die Geokoordinaten angegeben, um eine Nachvollziehbarkeit der Bilddaten zu gewährleisten.

10.1 Fallbeispiel: Skålbygge (Schweden) – 61°24'54.51" N, 15°04'42.80" E



Goldstandard für Skålbygge mit manuell gesetztem Startpunkt und initialer Richtung (unten).



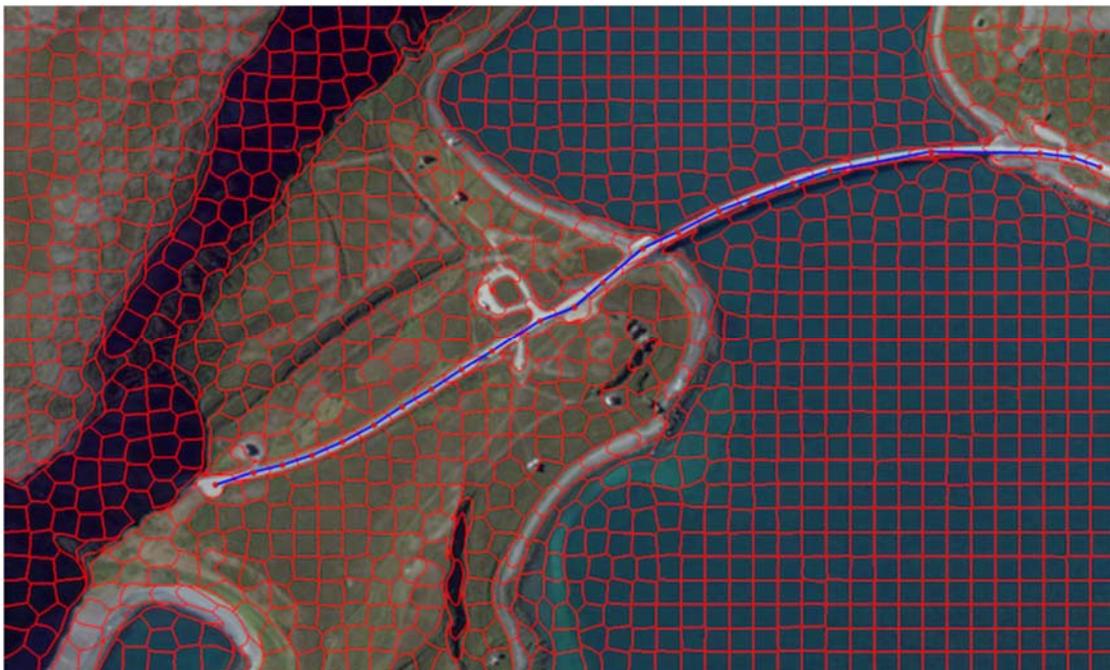
Ergebnis der Straßenverfolgung für Skålbygge. Der blau eingefärbte Kantenzug beschreibt die extrahiert Straße, die roten Punkte die Schwerpunkte der beteiligten Superpixel. Die einzelnen SLIC Superpixel sind durch ihre Begrenzungen rot eingefärbt.

In diesem Beispiel war das vorgeschlagene Verfahren in der Lage, die Straßenverfolgung oben korrekterweise abzubrechen, da dort keine gültige Vervollständigung der Straße möglich ist.

10.2 Nordkap (Norwegen) – 70°57'27.77" N, 25°41'56.54" E



Goldstandard für das Nordkap mit manuell gesetztem Startpunkt und initialer Richtung (rechts).



Ergebnis der Straßenverfolgung für das Nordkap. Der blau eingefärbte Kantenzug beschreibt die extrahiert Straße, die roten Punkte die Schwerpunkte der beteiligten Superpixel. Die einzelnen SLIC Superpixel sind durch ihre Begrenzungen rot eingefärbt.

Der vorgeschlagene Ansatz war in diesem Beispiel in der Lage, dem Straßenverlauf entlang verschiedener Umgebungstexturen auch über eine Brücke hinweg zu folgen.

Personendetektion in Grauwertbildern eines Stereo-Fahrzeugkamarasystems in Nachtsituationen mit variierenden Beleuchtungsverhältnissen

ALEXANDER HANEL¹, HORST KLÖDEN², LUDWIG HOEGNER¹ & UWE STILLA¹

Zusammenfassung: In Fahrzeugen der neuesten Generation werden Stereobilder vom Fahrweg des Fahrzeugs bei Tageslicht zur Detektion von Personen im Fahrweg genutzt. In diesem Beitrag wird untersucht, ob ein derartiger Detektionsalgorithmus für Bilder aus der Nacht geeignet ist. Dazu wird für verschiedene Beleuchtungssituationen je ein separates Detektormodell angelernt und ausgewertet. Weiter wird in diesem Beitrag eine Methode vorgestellt, mit der Falschdetektionen anhand der Disparitätskarte erkannt und verworfen werden können. Dazu wird die metrische Größe jeder Detektion bestimmt und die Plausibilität anhand realer Größenangaben von Personen geprüft. In einer Messfahrt wurden für verschiedene Beleuchtungssituationen am Tag und in der Nacht Stereobilder des Fahrwegs aufgezeichnet. Es hat sich gezeigt, dass der Detektor durch separate Modelle für Beleuchtungssituationen bei Nacht eine hohe Detektionsgüte mit einer mittleren Präzision über 65% erreichen kann. Dies ist wahrscheinlich auf die geringe Varianz des Erscheinungsbildes von Personen bei Nacht zurückzuführen. Die geringe Detektionsgüte bei Tag mit unter 30% lässt sich durch Plausibilitätsprüfung auf annähernd 80% steigern.

1 Objektdetektion im automobilen Umfeld

Heutige Fahrzeuge besitzen eine Vielzahl von Fahrerassistenzsystemen zur Verhinderung von Unfällen oder zur Verringerung ihrer Auswirkungen. Wesentliche Voraussetzung für solche Systeme ist es, zuverlässige Informationen über Objekte im Fahrweg eines Fahrzeugs zu besitzen. Insbesondere Fußgängern als leicht verletzliche Verkehrsteilnehmer kommt bei der Auslegung solcher Systeme eine hohe Bedeutung zu. Deswegen werden häufig Detektionsalgorithmen verwendet, die Personen in Bildern des Fahrwegs detektieren. Die neueste Generation von Fahrzeugen sind zur Erfassung des vorausliegenden Fahrzeugumfelds dazu oft mit einem Stereokamerasystem ausgestattet (Abb. 1).

Heute verwendete Detektionsalgorithmen sind in der Regel für das Erscheinungsbild von Personen bei Tageslicht ausgelegt. Bei Nacht stehen mit derartigen Algorithmen arbeitende Fahrerassistenzfunktionen dem Fahrer nicht zur Verfügung. Weiterhin können Bilder analysiert werden, um die Beleuchtungssituation der abgebildeten Straßenverkehrsszene zuverlässig zu bestimmen (HANEL et al. 2015). Dadurch wird es möglich, Detektionsalgorithmen durch verschiedene Parametersätze an die vorherrschende Beleuchtungssituation anzupassen und somit auch für Nachtfahrten auszuliegen.

¹ Technische Universität München, Photogrammetrie und Fernerkundung, Arcisstraße 21, D-80333 München, E-Mail: [alexander.hanel, ludwig.hoegner, stilla]@hcu-hamburg.de

² BMW Forschung und Technik GmbH, Hanauer Straße 46, D-80992 München, E-Mail: horst.kloeden@bmw.de



Abb. 1: Unfallvermeidung durch Detektion eines Fußgängers im Fahrweg des Autos mit Hilfe eines Stereokamerasystems am Innenspiegel [CONTINENTAL, 2011]

Zur Einschätzung des Kollisionsrisikos mit einem Fußgänger spielt dessen Abstand vom Auto eine wesentliche Rolle. Immer leistungsfähigere Prozessoren ermöglichen es heutzutage, statt Monobilddaten auch Stereobilddaten während der Fahrt zu verarbeiten. Daraus können mit einer Disparitätskarte Personen mit geringem Abstand vor dem Fahrzeug detektiert werden. Noch effizienter kann die Entfernung bestimmt werden, wenn die Disparitätskarte zur Nachverarbeitung von Detektionen im Grauwertbild verwendet wird.

2 Arbeiten im Umfeld

2.1 Bilddeskriptoren

Zur Detektion von Objekten in Bildern existiert eine große Bandbreite an Verfahren. Diese Verfahren beinhalten typischerweise einen Deskriptor, um Merkmale von zu detektierenden Objekten aus dem Bild abzuleiten, und einen Klassifikator, um Merkmale verschiedener Objekttypen voneinander zu unterscheiden. Ein Beispiel für ein solches Verfahren ist der HOG-SVM-Detektor (DALAL & TRIGGS 2005). Der Merkmalsvektor des HOG-Deskriptors (engl. *histogram of oriented gradients*) setzt sich aus gewichteten Richtungsangaben der dominierenden Grauwertkanten der rasterförmig angeordneten Bildteile zusammen. Eine zeiteffiziente (KITTI 2016) Erweiterung des HOG-Deskriptors sind ACF-Merkmale (engl. *aggregated channel features*), die mit AdaBoost trainiert werden (DOLLAR et al. 2014). ACF-Merkmale aggregieren sich aus mehreren Kanälen des Bildes. Dazu zählen die LUV Farbkanäle, sechs HOG-Merkmale und die Gradientenstärke.

2.2 Klassifikationsverfahren

Klassifikationsverfahren lassen sich in überwachte und unüberwachte Verfahren unterscheiden. Bei überwachten Verfahren muss beim Anlernen des Klassifikatormodells die tatsächliche Klasse der Trainingsdaten bekannt sein (z. B. Bayes-Klassifikator, FUKUNAGA 1990), während unüberwachte Verfahren aus den Trainingsdaten Cluster bilden (z. B. Expectation Maximization, BILMES 1998). Zur Objektdetektion wird häufig die Support-Vektor-Maschine (JOACHIMS 1999) als überwachtes Klassifikationsverfahren eingesetzt. Sie erlaubt die Bestimmung der Trennebene

zur Unterscheidung von Merkmalsvektoren verschiedener Objekttypen auch für eine große Menge Trainingsdaten in akzeptabler Zeit.

Zur Beurteilung von Klassifikationsverfahren können aus der Zahl der Korrekt-, Falsch- und Fehlklassifikation Precision und Recall berechnet werden (SOKOLOVA & LAPALME 2009). Precision setzt die Zahl der korrekt detektierten Objekte ins Verhältnis zu allen detektierten Objekten und bewertet somit Falschklassifikation. Recall hingegen setzt die Zahl der korrekt detektierten Objekte ins Verhältnis zu allen tatsächlichen Objekten (Ground Truth) und bewertet somit Fehlklassifikation.

2.3 Stereosehen

Algorithmen zur Berechnung von Tiefeninformationen aus Bildern eines Stereokamerasystems sind Stand der Technik (z. B. SCHARSTEIN & SZELISKI 2002). Typischerweise werden zunächst die Bilder beider Kameras mit den Parametern der Stereokamerakalibrierung rektifiziert. Über Punkt-zu-Punkt-Korrespondenzen eines zeitsynchron aufgenommenen Bildpaars kann die Disparitätskarte berechnet werden, aus der Tiefenwerte abgeleitet werden können.

3 Personendetektion in Bildern des Straßenverkehrs bei Nacht

In diesem Beitrag wird ein Personendetektor nach dem HOG-SVM-Ansatz (Kapitel 2) verwendet. Dieser eignet sich zum Beispiel für Bilder, die den Fahrweg eines Fahrzeugs aus der Perspektive einer Kamera am Innenspiegel (vgl. Abb.1) zeigen. Dieser Abschnitt gliedert sich in zwei Teile. Im ersten Teil wird die Vorgehensweise zur Untersuchung der Eignung des Personendetektors für Nachtaufnahmen unterschiedlicher Beleuchtungssituationen beschrieben. Im zweiten Teil wird eine Methode vorgestellt, mit der Detektionen von Personen im Bild anhand der Disparitätskarte auf Plausibilität überprüft werden können. Prinzipiell lassen sich die folgenden Beschreibungen auch auf andere Objekttypen übertragen.

3.1 Spezifische Klassifikatormodelle für jede Beleuchtungssituation

Der Begriff Beleuchtungssituation bezeichnet in diesem Beitrag die Beleuchtung einer Szene (HANEL et al. 2015). Mehrere Beleuchtungsquellen unterschiedlicher Position, Richtung und Stärke tragen zur Beleuchtungssituation bei. Verschiedene Beleuchtungssituationen werden unterschieden („Beleuchtungsklassen“). Für jede Beleuchtungsklasse wird ein separates Klassifikatormodell des Personendetektors angelernt. Zur Detektion von Personen in einem Bild wird dieses rasterförmig abgetastet und jeder entnommene Bildausschnitt mit dem Detektor ausgewertet. Für jeden Bildausschnitt wird zuvor die Beleuchtungsklasse bestimmt (HANEL et al. 2015), um das passende Klassifikatormodell anwenden zu können.

Die Eignung des Personendetektors wird durch einen Vergleich von in Testbildern detektierten Personen mit Ground Truth Daten bestimmt. Der Detektor berechnet dazu für jede detektierte Person die Eckkoordinaten eines umschließenden Rechtecks (Abb. 2a). Jede Detektion weist einen Sicherheitswert auf, der den Abstand des Merkmalsvektors von der Klassengrenze im Merkmalsraum beschreibt (LIBSVM 2016). Liegt der Sicherheitswert über einer festgelegten absoluten Sicherheitsschwelle, wird eine Detektion akzeptiert. Als Ground Truth werden im Bild durch ein umschließendes Rechteck manuell markierte Personen verwendet. Als

Korrektklassifikation wird eine Detektion gewertet, sobald sich mehr als ein festzulegender Anteil der Rechtecksfläche von Detektion und Ground Truth überlappen. Um unterschiedlich groß abgebildete Personen detektieren zu können, wird jedes Testbild in eine Bildpyramide zerlegt.



Abb. 2: Bildausschnitte zum Anlernen eines Personendetektors: a) positiver Bildausschnitt mit Person, b) negativer Bildausschnitt ohne Person.

Zur Beurteilung der Detektionsgüte werden für einen Testdatensatz die mittlere Precision sowie die Precision-Recall-Kurve (PR-Kurve) abgeleitet. Jeder Punkt in dieser Kurve repräsentiert den Precision- und Recall-Wert für eine bestimmte Sicherheitsschwelle. Die Sicherheitsschwelle wird dazu in einem festzulegenden Wertebereich variiert. Je höher die Sicherheitsschwelle, desto weniger Falschklassifikationen, desto mehr Fehlklassifikationen. Die PR-Kurve kann somit genutzt werden, um für die produktive Anwendung des Detektors eine optimale Sicherheitsschwelle zu wählen, bei der Precision und Recall möglichst hoch sind.

3.2 Plausibilitätsprüfung von Detektionen im Bild

Im Folgenden wird eine Methode zur Überprüfung der Plausibilität einer Detektion im Bild mit Hilfe der Disparitätskarte beschrieben. Diese Methode erlaubt es, Falschklassifikationen zu erkennen und zu verwerfen. Dadurch kann für den Detektor eine niedrige Sicherheitsschwelle mit wenigen Fehlklassifikationen gewählt und somit die Detektionsgüte insgesamt gesteigert werden. Voraussetzung zur Plausibilitätsprüfung ist ein Stereokamerasystem, bei dem eines der beiden rektifizierten Bilder zur Personendetektion verwendet wird und die Disparitätskarte abgeleitet werden kann (Abb. 3).

Zur Plausibilitätsprüfung wird für jede zuvor im Bild detektierte Person die metrische Höhe des umschließenden Rechtecks ermittelt. Liegt dieser Wert im Vergleich mit realen Größenangaben von Personen außerhalb eines Akzeptanzintervalls, so wird diese Detektion verworfen. Da Falschdetektionen oft eine unplausible Höhe des umschließenden Rechtecks aufweisen, reduziert sich deren Anzahl.



Abb. 3: Beispielaufnahme (aufgehellt) eines Fahrzeug-Stereokamerasystems in der Nacht. Links: Rektifiziertes Grauwertbild, rechts: Disparitätsbild.

Zunächst werden die Eckkoordinaten des umschließenden Rechtecks vom rektifizierten Stereobild in die Disparitätskarte projiziert. Im Fall einer korrekten Detektion zeigt der Bildausschnitt innerhalb des umschließenden Rechtecks die detektierte Person sowie an jedem Seitenrand Objekte aus dem Bildhintergrund. Um die metrische Größe der Person abzuschätzen, muss deren Entfernung von der Kamera anhand der Disparitätskarte berechnet werden. Dazu wird innerhalb des umschließenden Rechtecks der Median aller Disparitätswerte, die zu den Pixeln der abgebildeten Person gehören, gebildet und anhand der Stereokalibrierparameter in einen Entfernungswert umgerechnet. Für die Person und den Bildhintergrund werden im Histogramm der Disparitäten dazu unterschiedliche, aber dominante Werte angenommen. Durch Bestimmung des markantesten Minimums im Histogramm zwischen diesen Maxima (Peakiness-Test, vgl. HOEGNER et al. 2014) wird eine absolute Schwelle bestimmt, anhand derer zum Bildhintergrund gehörende Pixel erkannt werden können (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Die Höhe des umschließenden Rechtecks wird mit Hilfe des berechneten Entfernungswertes ermittelt und mit Größenangaben realer Personen verglichen. Für die Größenangabe wird eine untere sowie obere Schwelle gewählt, wobei berücksichtigt wird, dass die Höhe des umschließenden Rechtecks ebenfalls den Bildhintergrund beinhaltet. Detektionen mit nicht plausibler metrischer Höhe des umschließenden Rechtecks werden verworfen.

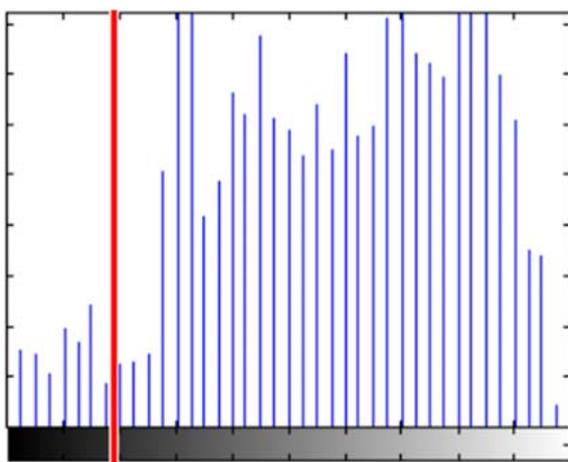


Abb. 4: Histogramm der Disparitätswerte des umschließenden Rechtecks einer Detektion. Disparitäten durch 8bit-Grauwerte codiert auf horizontaler Achse dargestellt. Markantestes Minimum zwischen Disparitätsmaximum von Person und Bild-hintergrund in rot.

4 Experiment

Es wurden am Tag und in der Nacht Messfahrten auf städtischen und ländlichen Straßen durchgeführt. Ein Stereokamerasystem hat Grauwertbilder vom Fahrweg vor dem Fahrzeug durch die Windschutzscheibe aufgezeichnet. Zum Anlernen des Klassifikators werden für jede Beleuchtungsklasse aus den Bildern etwa 10.000 positive und 30.000 negative Bildausschnitte mit beziehungsweise ohne Einzelpersonen extrahiert (Abb. 2). Das ungleiche Verhältnis zwischen positiven und negativen Stichproben ist typisch für Personendetektoren (DALAL & TRIGGS 2005). Positive Bildausschnitte sind durch manuelles Markieren der Personen im Bild vorgegeben, während negative Bildausschnitte zufällig aus Bildern des Fahrzeugkamerasystems ohne abgebildete Personen selektiert wurden. Die Bildausschnitte wurden per Algorithmus absteigend in vier Beleuchtungsklassen (HANEL et al. 2015) eingeteilt: *Tag*, *Nacht-Stadtzentrum* (Ego-Fahrzeug-, Gegenverkehr- und Straßenbeleuchtung), *Nacht-Wohngebiet* (Ego-Fahrzeug- und Straßenbeleuchtung) und *Nacht-Landstraße* (nur Ego-Fahrzeugbeleuchtung).

Die aufgenommenen Stereobilder wurden rektifiziert sowie die Disparitätskarte abgeleitet. Der Personendetektor wurde für jede Beleuchtungsklasse angelernt. Die Eignung des Detektors wird für die Beleuchtungsklasse *Tag* durch 204 Bilder mit 105 abgebildeten Personen, für *Nacht-Wohngebiet* durch 278 Bilder und genauso vielen Personen, für *Nacht-Stadtzentrum* durch 225 Bilder mit 200 Personen und für *Nacht-Landstraße* durch 208 Bilder mit 130 Personen evaluiert. Anschließend wird für jede Detektion die Plausibilitätsprüfung durchgeführt.

5 Ergebnisse und Diskussion

In diesem Abschnitt werden zunächst die Ergebnisse der Personendetektion mit HOG-SVM für die unterschiedlichen Beleuchtungsklassen diskutiert. Grundlage bildet für jede Beleuchtungsklasse die mittlere Precision (Tab. 1) sowie die PR-Kurve (Abb. 5).

Tab. 1: Mittlere Precision des Personendetektors für jede Beleuchtungsklasse vor (mittlere Spalte) und nach (rechte Spalte) der Plausibilitätsprüfung.

Beleuchtungsklasse	Mittlere Precision	Mittlere Precision nach Plausibilitätsprüfung
Wohngebiet	83,80%	87,04%
Stadtzentrum	65,19%	81,89%
Landstraße	65,05%	73,79%
Tag	22,97%	78,82%

Es zeigt sich, dass die mittlere Precision des Personendetektors bei Tagaufnahmen deutlich geringer ist als bei Nachtaufnahmen. Während die Werte der mittlere Precision der Nachtaufnahmen weit über 50% liegen, befindet er sich für Tagaufnahmen weit darunter. Ursache für diesen Unterschied kann sein, dass bei Nachtaufnahmen weniger Falschklassifikationen auftreten. Der durch die Dunkelheit nahezu schwarze Bildhintergrund weicht stärker vom Erscheinungsbild von Personen im Bildvordergrund ab, als der Bildhintergrund bei Tagaufnahmen

mit verschiedenen abgebildeten Objekten. Dadurch sinkt bei Nachtaufnahmen der Sicherheitswert von Detektionen in solchen Bildbereichen, wodurch diese verworfen werden.

Bei Vergleich der Nachtaufnahmen untereinander lässt sich feststellen, dass der Personendetektor für die Beleuchtungsklasse *Nacht-Wohngebiet* eine um 20 Prozentpunkte höhere mittlere Precision aufweist, als für die beiden anderen Nacht-Klassen. In der Klasse *Nacht-Stadtzentrum* variiert das Erscheinungsbild von Personen stärker als bei *Nacht-Wohngebiet*. Während in Wohngebieten die Straßenbeleuchtung regelmäßig ist, variiert die Beleuchtung im Stadtzentrum häufiger (Straßenbeleuchtung, Schaufenster, andere Fahrzeuge). Variierende Beleuchtung beeinflusst das Erscheinungsbild von Personen im Bild, was sich als negativer Einfluss auf die Detektionsgüte äußert. Die Bilder in *Nacht-Landstraße* hingegen sind nur durch den Fahrzeugscheinwerfer beleuchtet, wodurch sich Personen in dunklen Bildteilen befinden können und somit nicht detektierbar sind, was Fehlklassifikationen verursacht.

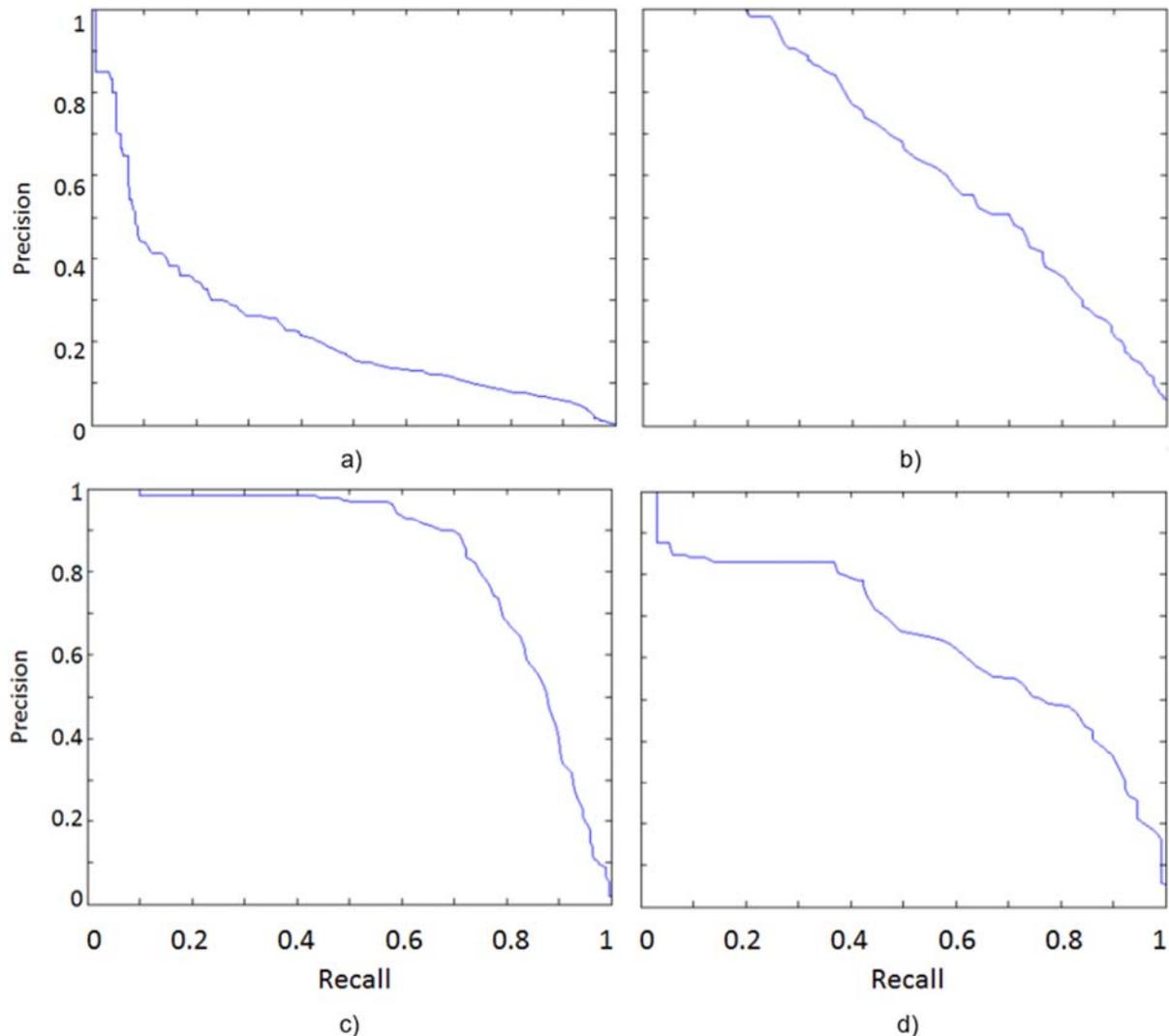


Abb. 5: PR-Kurven für den Personendetektor, a) Tag, b) Stadtzentrum, c) Wohngebiet, d) Landstraße.

Die PR-Kurve verläuft für *Nacht-Wohngebiet* am nächsten an den Asymptoten $Precision = 1$ beziehungsweise $Recall = 1$. Ersteres bedeutet, dass der Detektor für diese Beleuchtungsklasse die beste Precision aufweist, letzteres lässt die gleiche Feststellung für Recall zu. Es gibt somit bei der Wahl der Sicherheitsschwelle des Detektors ein Intervall der Recallwerte (0 bis ca. 0,6), für das keine oder kaum Falschklassifikationen in Kauf genommen werden müssen. Die optimale Sicherheitsschwelle ergibt sich jedoch an der Stelle mit dem größten Abstand der PR-Kurve vom Ursprung ($P \sim 0,8$; $R \sim 0,7$), wobei sowohl Fehl- als auch Falschklassifikationen in Kauf genommen werden. Etwas weniger asymptotisch verläuft die PR-Kurve für die beiden anderen Nacht-Beleuchtungsklassen. Es treten somit mehr Falsch- und Fehlklassifikationen auf. Am schlechtesten ist der Verlauf der PR-Kurve für Tag zu beurteilen, da sie am wenigsten asymptotisch verläuft. Es ist offensichtlich, dass in diesem Fall keine Sicherheitsschwelle gewählt werden kann, bei der nur eine der beiden Fehlerarten auftritt.

Bei den Nachtaufnahmen der Klassen *Wohngebiet* und *Landstraße* muss kritisch betrachtet werden, dass in den Trainings- und Testdaten aufgrund der geringen Passantenfrequenz nachts immer wieder dieselben Personen abgebildet sind. Es ist somit nicht auszuschließen, dass der Personendetektor nur für diese Personen eine hohe Güte liefert. Dieser Argumentation steht entgegen, dass die Güte für *Nacht-Stadtzentrum* vergleichbar ist, in deren Aufnahmen aber viele verschiedene Passanten zu sehen sind. Weiter muss noch angemerkt werden, dass neben der Beleuchtungsklasse auch die Wahl der Parameter des HOG-Deskriptors und des SVM-Klassifikators die Detektionsgüte beeinflussen.

Untersuchungen haben gezeigt, dass zur Personendetektion verschiedene Beleuchtungsklassen für Nachtaufnahmen unterschieden werden müssen. Ein einheitlicher Detektor für alle Nachtaufnahmen liefert schlechtere Resultate als der Detektor für Tagaufnahmen. Zusammengefasst zeigen die vorangegangenen Feststellungen, dass ein HOG-SVM-Personendetektor bei einer ausreichenden Zahl an Beleuchtungsklassen eine hohe Detektionsgüte für Nachtaufnahmen liefern kann.

Im Weiteren werden die Ergebnisse der Plausibilitätsprüfung besprochen. Für alle vier Beleuchtungsklassen steigt die mittlere Precision durch die Plausibilitätsprüfung (Tab. 1). Am deutlichsten ist der Anstieg bei der Klasse *Tag*, am geringsten bei *Nacht-Wohngebiet*. Das zeigt, dass gerade bei geringer mittlerer Precision die Plausibilitätsprüfung eine deutliche Verbesserung der Detektionsgüte bewirkt. Zurückzuführen ist diese Feststellung auf die gesunkene Zahl an Falschdetektionen (Abb. 6 und Abb. 7).

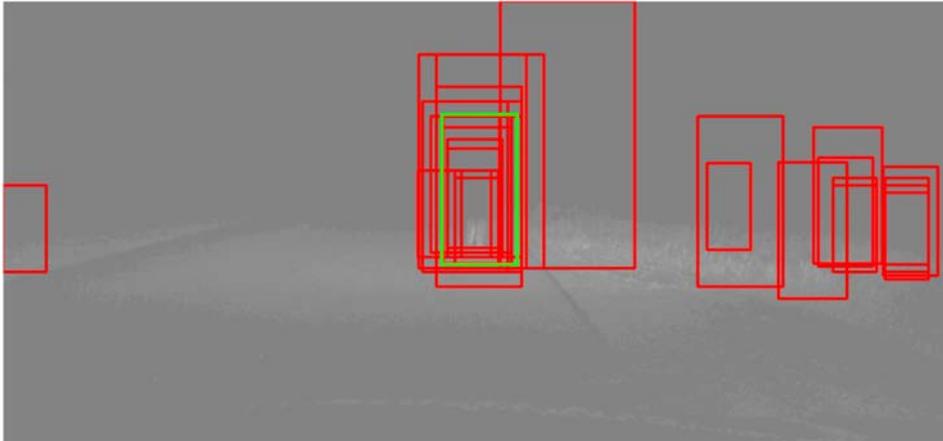


Abb. 6: Detektionen (rot) eines Bildes der Beleuchtungsklasse Nacht-Landstraße. Ground Truth einer Person in grün.

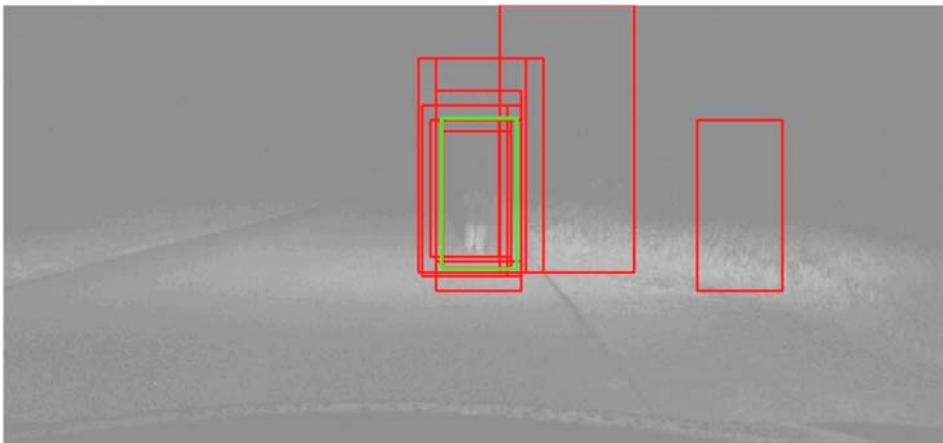


Abb. 7: Detektionen (rot) nach Plausibilitätsprüfung anhand der Disparitätskarte basierend auf den Detektionen aus Abb. 6.

Diese Beobachtung zeigt sich auch an den PR-Kurven. Für Tagaufnahmen verläuft die PR-Kurve (Abb. 8a) nun deutlich asymptotischer als ohne Plausibilitätsprüfung. Selbst für hohe Recall-Werte nimmt die Precision kaum ab. Es ist aber auch ersichtlich, dass immer ein paar wenige Falschdetektionen in Kauf genommen werden müssen (Precision erreicht nie 1). Bei der optimalen Sicherheitsschwelle sind Precision und Recall ($P \sim 0,75$; $R \sim 0,9$) besser als ohne Plausibilitätsprüfung. Eine ähnliche Feststellung lässt sich für *Nacht-Stadtzentrum* treffen (Abb. 8b). Es lässt sich ebenfalls eine Sicherheitsschwelle mit besserer Precision und Recall wählen ($P \sim 0,75$; $R \sim 0,75$). Im Gegensatz zu den Tagaufnahmen treten für diese Beleuchtungsklasse bei niedrigen Recall-Werten überhaupt keine Falschklassifikation auf. Ähnliche Feststellungen lassen sich für *Nacht-Landstraße* treffen. Bei *Nacht-Wohngebiet* sind die Auswirkungen der Plausibilitätsprüfung gering, da bereits vorher eine hohe Detektionsgüte vorgelegen hat. Die Plausibilitätsprüfung erlaubt somit – im Gegensatz zu einer Änderung der Sicherheitsschwelle – eine Reduktion der Falschklassifikationen, ohne zusätzliche Fehlklassifikationen in Kauf nehmen zu müssen.

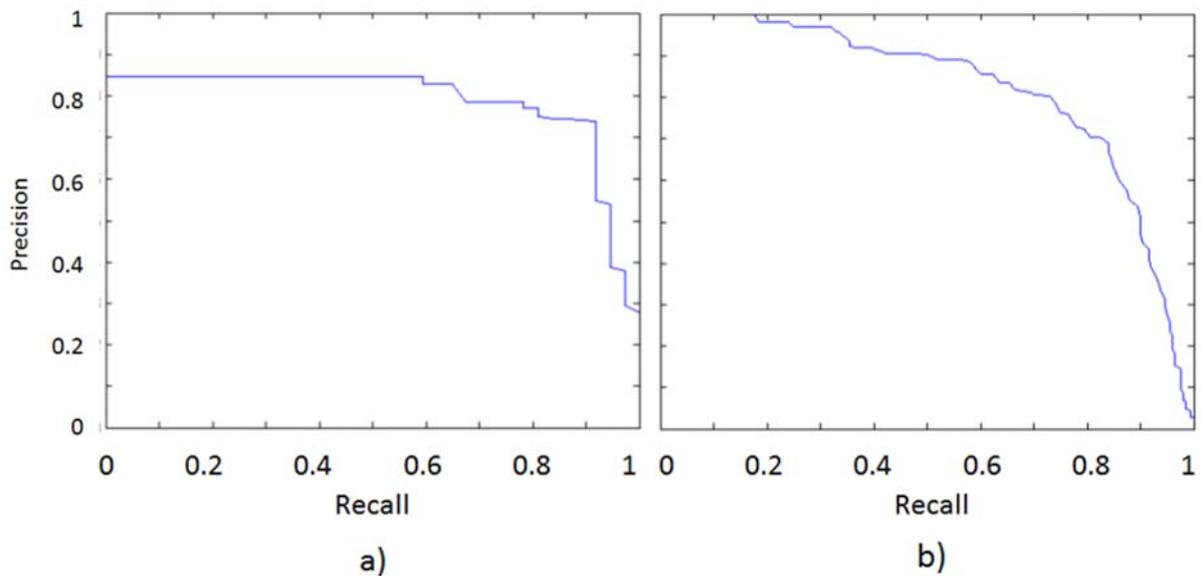


Abb. 8: Precision-Recall-Kurve nach Plausibilitätsprüfung, a) Tag, b) Stadtzentrum.

6 Fazit und Ausblick

Dieser Beitrag gliedert sich in zwei Teile. Zunächst wird ein Verfahren beschrieben, mit dem die Eignung eines HOG-SVM-Personendetektors für Nachtaufnahmen einer Fahrzeugkamera untersucht werden kann. Dazu werden verschiedene Beleuchtungssituationen der Straßenszene unterschieden und für jede Beleuchtungsklasse ein separates Klassifikatormodell angelernt. Die Analyse der Eignung geschieht anhand der mittleren Precision des Detektors sowie der Precision-Recall-Kurve. Im zweiten Teil wird ein Verfahren vorgestellt, mit dem Falschdetektionen in Bildern reduziert werden können. Aus der Disparitätskarte eines Stereokamerasystems wird die metrische Größe einer Detektion bestimmt und mit Größenangaben realer Personen verglichen. Mit einem Stereokamerasystem wurden Bilder vom Fahrweg eines Fahrzeugs bei Tag und Nacht für unterschiedliche Beleuchtungsklassen aufgenommen. Die Ergebnisse zeigen, dass für nächtliche Beleuchtungsklassen eine höhere Detektionsgüte erreicht werden kann als für Beleuchtungsklassen bei Tageslicht. Eine Unterscheidung verschiedener Beleuchtungsklassen bei Nacht ist notwendig, um eine hohe Detektionsgüte sicherzustellen. Die Plausibilitätsprüfung kann insbesondere für die Tagklasse mit zuvor schlechter Detektionsgüte eine deutliche Verbesserung auf das Niveau der Nachtklassen bewirken.

In künftigen Arbeiten kann untersucht werden, ob eine weitergehende Unterteilung der Beleuchtungsklassen die Personendetektion verbessern kann. Insbesondere die schlechte Detektionsgüte der Tagklasse lässt Optimierungsbedarf vermuten. Durch Tracken einer Person über mehrere Bilder einer Bildfolge können ebenfalls bessere Ergebnisse erwartet werden. Ein Stereo-Thermalkamerasystem könnte gerade bei Nacht zu einer verbesserten Disparitätsberechnung und dadurch besserer Plausibilitätsprüfung führen. Überdies kann es zusätzliche temperaturabhängige Merkmale zur Detektion von Personen liefern.

7 Literaturverzeichnis

- BILMES, J., 1998: A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. Technical Report, University of California, Berkeley, Department of Electrical Engineering and Computer Science, Computer Science Division.
- CONTINENTAL, 2011: Mit der zweiten Kamera sieht man besser – die Stereokamera. Website, http://www.continental-corporation.com/www/presseportal_com_de/themen/pressemitteilungen/3_automotive_group/chassis_safety/press_releases/pr_20110504_stereokamera_de.html, abgerufen am 28.04.2016.
- DALAL, N. & TRIGGS, B., 2005: Histograms of oriented gradients for human detection. *IEEE Conference on Computer Vision and Pattern Recognition* **1**, 886-893.
- DOLLAR, P., APPEL, R., BELONGIE, S. & PERONA, P., 2014: Fast Feature Pyramids for Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **36** (8), 1532-1545.
- HANEL, A., 2014: Bestimmen der Beleuchtungsverhältnisse und Optimieren der Fußgängerdetektion aus Bildern eines Fahrzeugkamerasystems. Masterarbeit, Technische Universität München, Ingenieur fakultät Bau Geo Umwelt, Fachgebiet Photogrammetrie und Fernerkundung.
- HANEL, A., KLÖDEN, H., HOEGNER, L. & STILLA, U., 2015: Erkennung von Szenenbeleuchtungen aus Bildern eines Fahrzeugkamerasystems zur Verbesserung der Objektdetektion. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, Band **24**, T. Kersten (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 254-262.
- HOEGNER, L., HANEL, A., WEINMANN, M., JUTZI, B., HINZ, S. & STILLA, U., 2014: Towards people detection from fused time-of-flight and thermal infrared images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (3), 121-126.
- FUKUNAGA, K., 1990: *Introduction to Statistical Pattern Recognition* (2nd Ed.). Academic Press Professional, Inc.
- JOACHIMS, T., 1999: Making large scale SVM learning practical. Technischer Bericht, Universität Dortmund.
- KITTI, 2016: Object Detection Evaluation 2012. Webseite, http://www.cvlibs.net/datasets/kitti/eval_object.php, abgerufen am 28.04.2016.
- LIBSVM, 2016. LIBSVM FAQ. Webseite, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/...faq.html#f4151>, abgerufen am 28.04.2016.
- SCHARSTEIN, D. & SZELISKI, R., 2002: A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision* **47**(1), 7-42.
- SOKOLOVA, M. & LAPALME, G., 2009: A Systematic Analysis of Performance Measures for Classification Tasks. *Inf. Process. Manage.* **45** (4), 427-437.

Utilisation de la corrélation diachronique d'images pour la métrologie de ponts.

MICHEL KASSER¹

Résumé: La corrélation diachronique d'images est une technique de mesure basée sur la recherche de différences géométriques, même minimes, entre des images prises à différentes époques. Ces images sont soit prises depuis des points différents et orthorectifiées, soit simplement prises depuis une même station. Déjà utilisée pour des mesures de déformations tectoniques, ou aussi sur des tests de matériaux en laboratoire, elle permet de déterminer des cisaillements dans l'image avec une précision de l'ordre du centième de pixel, voire mieux, avec un simple appareil photo.

Dans le domaine de la surveillance de ponts, les besoins actuels en solutions d'auscultation très économiques sont grands. La corrélation diachronique d'images, permettant des mesures sans contact, sans pose de repères, et à faible coût, est en mesure de répondre au moins en partie à cette demande.

L'enjeu principal des recherches menées à la HEIG-VD a été d'évaluer quelles sont les possibilités d'adaptations de ces techniques au domaine de la surveillance de tabliers de ponts, et leurs limitations. Les résultats ont révélé que la précision recherchée du centième de pixel peut aussi être atteinte pour de telles mesures de flexion. Finalement des essais ont permis de quantifier les pertes de précision liées à la réfraction sur longues distances qui, dans certains cas, restent très faibles.

1 Introduction

La corrélation d'images a fait son apparition avec l'arrivée des premières images spatiales numériques destinées à être exploitées en stéréo, dans les années 1980 (Spot 1). En corrélation diachronique, la mesure de correspondance entre des images issues de différentes époques permet la détermination d'un champ de déplacements. Son utilisation s'est largement développée dans les milieux de la recherche sur les matériaux pour la réalisation de tests sous contraintes en laboratoire, mais aussi et d'abord pour des études de déformations tectoniques co-sismiques, ou des mouvements de versants instables. Cette technique à l'avantage de permettre des mesures sur toute la surface, sans contact, et ce sur une très large gamme d'échelles de mesure. Il est ainsi possible d'analyser des déplacements sur une surface allant de quelques mm² jusqu'à plusieurs centaines de km² en fonction du type de matériel employé et de la configuration du système mis en place.

Dans le domaine de l'auscultation de ponts, on travaille aujourd'hui généralement avec des outils de topométrie très précis permettant uniquement une surveillance localisée sur des points caractéristiques de l'ouvrage, minutieusement choisis. Ces campagnes sont donc longues et coûteuses en raison des difficultés d'intervention et ne fournissent qu'un ensemble de mesures ponctuelles réparties sur l'ouvrage. Et en général il est difficile, ou en tous cas coûteux, d'atteindre une précision millimétrique. En outre, ces mesures s'étendent sur un temps important (plusieurs

¹ HEIG-VD, Laboratoire G2C, Route de Cheseaux 1, CH-1401 Yverdon-les-bains, Suisse,
E-Mail: michel.kasser@heig-vd.ch

heures), ce qui complique le travail puisque pendant ce temps la température change, et donc les dimensions du pont également.

Les avantages des techniques de corrélation diachronique appliquées à l'auscultation d'ouvrages sont importants. La possibilité de disposer d'un champ de vecteurs sur l'ensemble d'un ouvrage sans contact correspond entièrement aux besoins actuels où l'on cherche à limiter les coûts et réduire la durée des interventions. Aujourd'hui, de nombreux ponts sont vieillissants et devraient faire l'objet d'auscultations régulières. Il en ressort donc une forte demande de surveillance et un intérêt grandissant pour des solutions d'auscultation à très faible coût.

L'enjeu principal de ces recherches est donc d'évaluer quelles sont les possibilités d'adaptations des techniques de corrélation à ce domaine précis de la surveillance de ponts, moyennant des conditions d'environnement très variables. Il convient donc de recenser quelles sont les limites d'utilisation dans les conditions courantes, et en particulier jusqu'à quelle distance du pont on peut acquérir les images.

2 Principes de la corrélation diachronique d'images

La corrélation diachronique d'images est une recherche des déformations entre deux images prises à différentes époques. En considérant un centrage et une orientation identique de la prise de vue aux deux époques, la corrélation permet de déterminer les changements intervenus entre les images. L'image référence correspond à l'état initial et les traitements mesurent les déplacements (disparités) par rapport l'image déformée, dans un état ultérieur.

Pour cela, les algorithmes de corrélation recherchent la correspondance maximale entre des petites portions d'images appelées imagettes. Lorsque le pic de corrélation est identifié, les déformations entre les deux images sont obtenues. La qualité de la corrélation dépend beaucoup de la qualité radiométrique et de la dynamique des images. La présence de bruit a pour effet de réduire la qualité de la corrélation entre les images, et les appareils récents offrent souvent un excellent rapport signal/bruit.

Par ailleurs, il faut que la texture de l'objet photographié soit adaptée, les structures périodiques posent ainsi toujours d'importants problèmes. Par contre, les surfaces anisotropes telles que le béton brut sont généralement satisfaisantes, surtout si la surface est exposée aux intempéries depuis longtemps.

3 Contraintes et limites du système

3.1 Effets thermiques

Les variations dimensionnelles du capteur et de l'optique avec la température rendent impossible la détection de changements ayant un effet isotrope sur l'image : on doit considérer que de telles déformations ne sont pas mesurables avec cette technique. Dans ces conditions, une instabilité de la distance focale, qui engendrerait le même problème, n'a pas à être combattue, ce qui permet de travailler avec tous les appareils photo courants.

3.2 Effets de la distorsion

La distorsion parfois très importante présente dans tous les systèmes d'acquisition d'images est un phénomène et bien connu des photogrammètres. Autrefois ils la corrigeaient en la mesurant au préalable (étalonnage), et maintenant grâce à l'abondance en points de liaison il est classique de résoudre comme des inconnues supplémentaires tous les paramètres d'orientation interne, qui incluent ceux liés à la distorsion.

Celui-ci implique que l'ouvrage ausculté n'est pas représenté par la même portion des deux images. La distorsion étant unique en chaque point de l'image, la mise en correspondance fait apparaître des différences de distorsion, dont l'importance fait disparaître les déformations effectives. Nous avons cherché à savoir s'il était possible de se passer de mesurer cette distorsion dans un cas concret.

La distorsion parfois très importante présente dans tous les systèmes d'acquisition d'images est un phénomène et bien connu des photogrammètres. Autrefois ils la corrigeaient en la mesurant au préalable (étalonnage), et maintenant grâce à l'abondance en points de liaison il est classique de résoudre comme des inconnues supplémentaires tous les paramètres d'orientation interne, qui incluent ceux liés à la distorsion. Pour un appareil courant non étalonné, lorsque le pointage de l'appareil n'est pas parfaitement identique entre les deux prises de vues, cela se traduit certes par une translation entre les images, qui est corrigée lors du traitement, mais les effets dus à la distorsion non corrigée des images sont directement fonction du dépointage. Celui-ci implique que l'ouvrage ausculté n'est pas représenté par la même portion des deux images. La distorsion étant unique en chaque point de l'image, la mise en correspondance fait apparaître des différences de distorsion, dont l'importance masque alors les déformations effectives. Nous avons cherché à savoir s'il était possible de se passer de mesurer cette distorsion dans un cas concret.

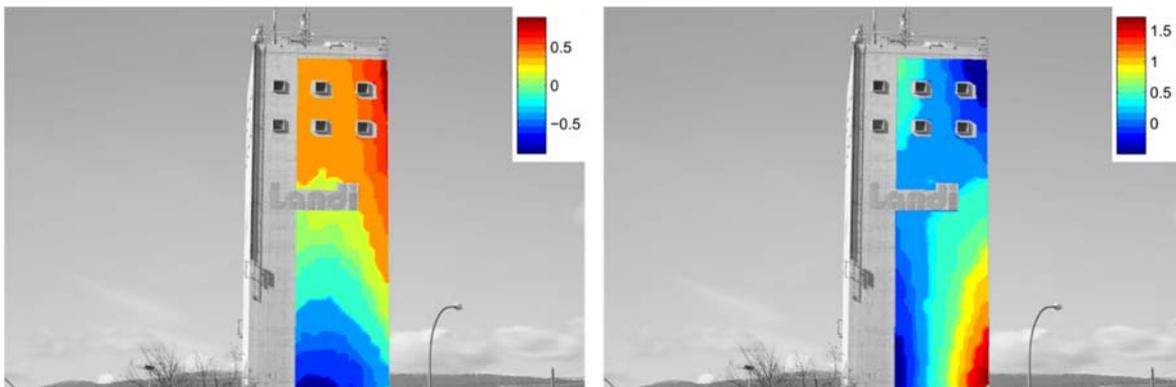


Fig. 1: Effets sur la composante horizontale (à gauche) et verticale (droite) d'un dépointage de 280 pixels, qui engendre un fort effet dû à la distorsion de l'optique. Couleurs exprimées en pixels (1 pixel = 15 mm).

La figure 1 montre les effets présents lorsque le point visé par l'appareil (Canon, focale 50 mm) est modifié de 280 pixels. Ces déplacements erronés atteignent près de 2 px aux extrémités de l'ouvrage. Les tests réalisés ont permis de définir qu'un dépointage de 15 pixels provoque encore des effets allant jusqu'à 0.1 pixel.

Les moyens de mise en place du capteur permettent d'atteindre facilement un dépointage inférieur à 5 px. Sachant que la distorsion augmente avec l'éloignement du centre de l'image, le centième de pixel peut alors être atteint en n'utilisant que leur partie centrale. Mais pour des cas où le

pointage précis est difficile à reproduire, il est finalement préférable de mesurer puis corriger la distorsion, ce qui permet alors aussi de travailler avec des images issues d'appareils différents si besoin est.

3.3 Effets de rotation (mise en station)

En corrélation diachronique, l'appareil doit être centré de la même manière entre les deux états de mesure. En planimétrie, les embases permettent un centrage de l'appareil amplement suffisant pour ces applications. Par contre, pour la mise à l'horizontale, même une très légère inclinaison du capteur influence directement les mesures de disparités entre les images.

La figure 2 montre les résultats de corrélation d'une paire d'images acquises avec une mise à l'horizontale approximative (± 0.2 gons). L'interprétation des disparités de la tour indique une rotation du bâtiment dans le sens horaire. Le défaut d'horizontalité est très visible et, sans corrections, rend l'analyse des déformations impossible.

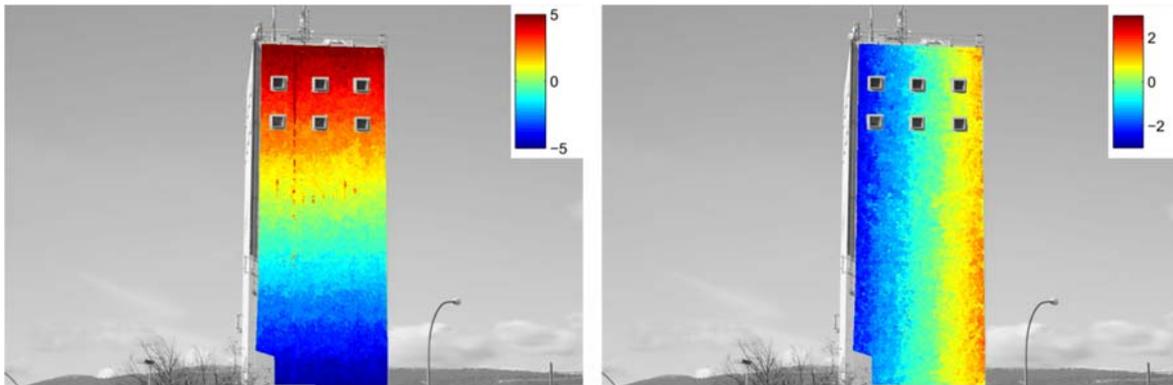


Fig. 2: Composantes horizontale (à gauche) et verticale (droite) de la corrélation entre deux images, avec une rotation de l'appareil entre les deux acquisitions, les couleurs sont exprimées en pixels (1 pixel = 15 mm)

Un point important à noter est que l'on ne corrige que très imparfaitement cet artefact par un calcul normal, car c'est directement la corrélation entre deux imagerie dont l'une a subi une petite rotation qui est dégradée. La solution la plus efficace consiste donc à recalculer, par re-échantillonnage, une des images en supprimant la rotation d'ensemble, et seulement ensuite à appliquer la corrélation diachronique. Dans ces conditions, on parvient à atteindre la précision de 0.01 pixel.

3.4 Dynamique des images

La dynamique des images est un élément essentiel de la qualité de la corrélation. Il faut donc régler l'acquisition d'image pour que l'histogramme soit le plus proche possible de la saturation, sans l'atteindre bien entendu : c'est ainsi que l'image présente le meilleur rapport signal/bruit (Fig. 3). Mais même ainsi, le coefficient de corrélation peut rester insuffisant sur certaines surfaces très unies. La notion de surface unie (radiométrie constante, donc corrélation impossible) est directement liée à la dynamique : avec 12 bits significatifs, on n'observe presque aucune surface unie, là où avec 7 bits on en trouve beaucoup. Cela signifie-t-il qu'on ne peut pas traiter des surfaces unies avec un appareil courant à faible dynamique ? Non, car il est souvent possible,

lorsque l'objet ne se déforme que très lentement, d'effectuer une sommation d'images acquises en rafale.

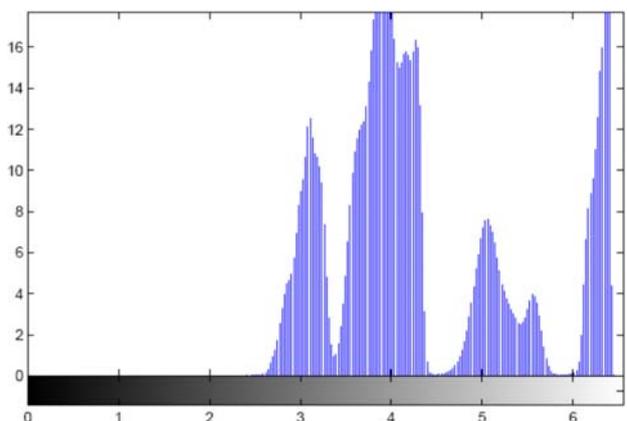


Fig. 3: Un histogramme d'image optimal, s'approchant autant que possible de la saturation sans l'atteindre.

En effet si rien ne bouge pendant les acquisitions successives, pour chaque pixel de l'image on peut sommer les valeurs obtenues dans les images successives. Pour 2 images, on augmente ainsi de 1 bit significatif la dynamique de l'image, ou encore pour $64 = 2^6$ images sommées on l'augmente de 6 bits. Cette technique permet de résoudre à peu près toutes les situations, hormis celles où l'objet est susceptible de bouger pendant les acquisitions (Fig. 4). Dans ce type de traitement d'image, le signal augmente en fonction du nombre N d'images sommées, alors que le bruit, réputé Gaussien, ne croît qu'en $N^{1/2}$, donc le rapport signal/bruit est amélioré lui aussi en $N^{1/2}$.

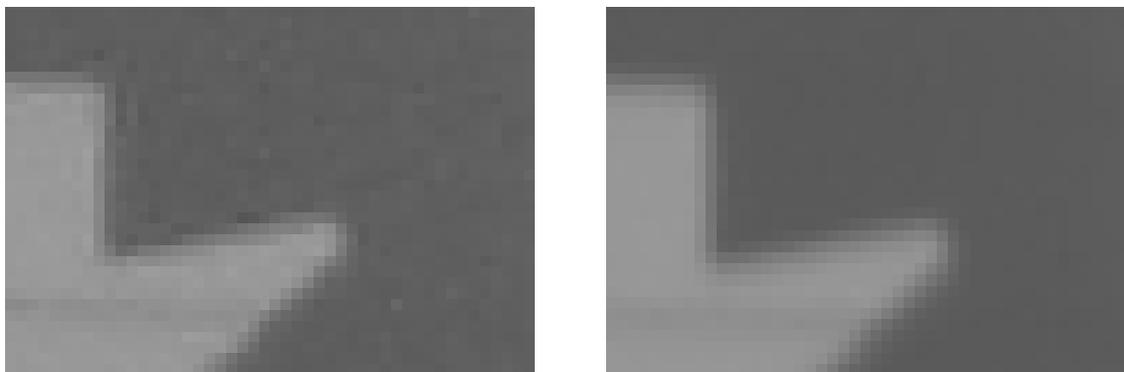


Fig. 4: Exemple de sommation d'images : à gauche une image normale, à droite la somme de 16 images, montrant la réduction considérable du bruit.

Comme exemple, celui d'un mur en béton peint, avec certaines surfaces très unies (Fig. 5).

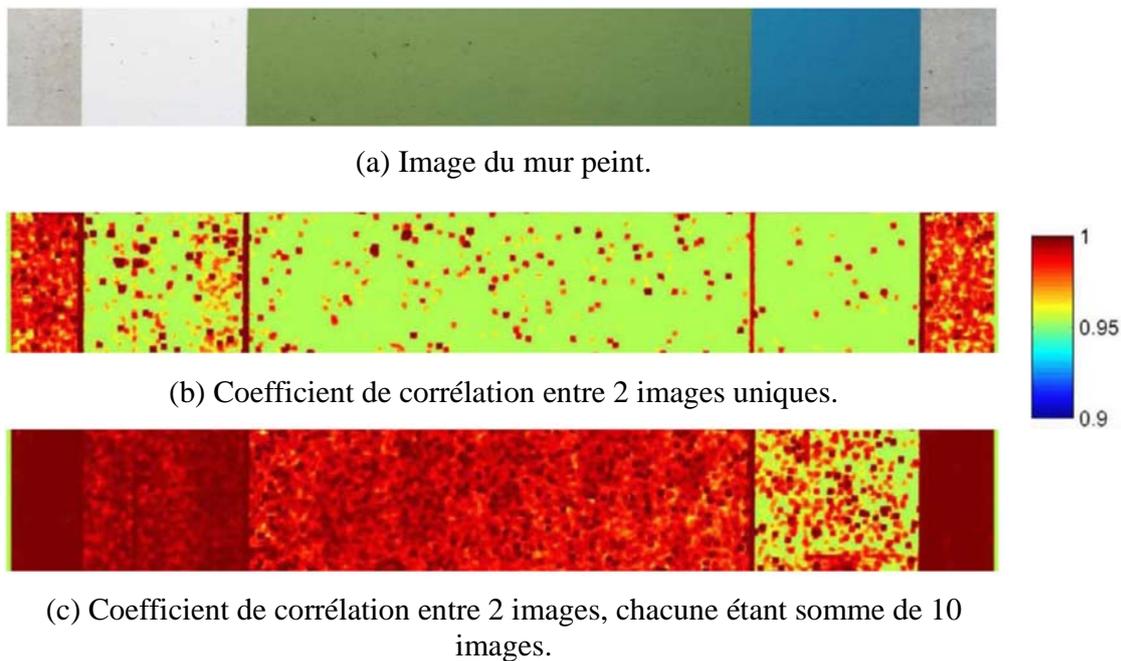


Fig. 5: En (a), l'image du mur, en (b) est représenté le coefficient de corrélation entre deux images successives, avec de larges zones où ce coefficient est inférieur à 0.96, seuil qui correspond à des corrélations satisfaisantes. En (c), avec une sommation de 10 images, le coefficient devient satisfaisant à peu près partout.

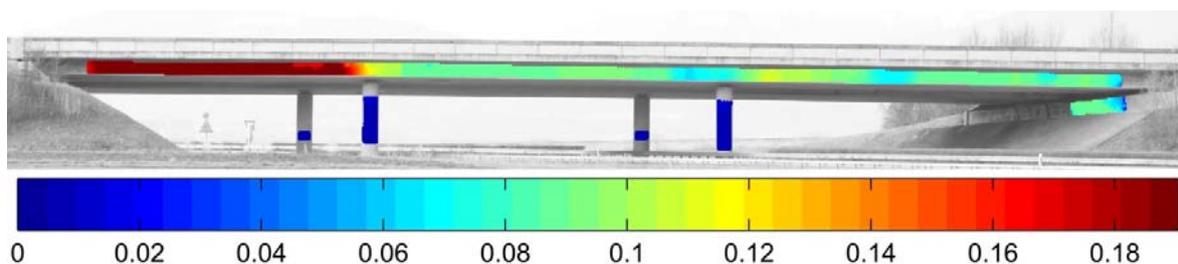
Néanmoins, on observe généralement que, sur les surfaces en béton naturel, le coefficient de corrélation est satisfaisant avec une seule image pour des appareils photo récents, qui ont souvent plus de 10 bits significatifs. C'est ainsi que le pont ci-dessous a permis une corrélation satisfaisante pour des images uniques malgré un éclairage médiocre (contre-jour), ce qui permet de mesurer sa flexion en situation dynamique (Fig. 6).



(a) Acquisition des images.



(b) Exemple d'image traitée.



(c) Résultat de la corrélation sur le tablier et les piles de pont.

Fig. 6: Pont autoroutier et les déplacements verticaux observés à l'occasion du passage d'un véhicule, les couleurs sont exprimées en pixels (1 pixel = 3 cm).

3.5 Effets de la réfraction à grande distance

Afin de répondre à des situations de terrain particulières, il est important de voir s'il est possible de mesurer à des distances importantes, par exemple lorsqu'aucun site proche ne peut être trouvé pour observer le tablier du pont sous un angle satisfaisant. Les effets de la réfraction sont en effet très largement visibles à de telles distances.

Les essais menés ont porté sur une distance de 650 m. Le traitement des images, avec le logiciel Open Source MicMac, a été fait toujours sur le principe suivant : une translation générale strictement verticale est appliquée de manière à annuler le déplacement vertical apparent au niveau des culées, et la comparaison est faite sur toute la surface latérale visible du tablier.

À titre indicatif, avec des conditions normales du coefficient de réfraction k (0.12), le déplacement de l'image est de l'ordre de 3 pixels. Mais on observe couramment qu'à proximité du sol, les valeurs de k peuvent varier dans une très large gamme (de +10 à -10).

Fondamentalement, si cette déviation se fait dans un plan vertical avec une atmosphère bien stratifiée, ce phénomène n'est pas gênant puisqu'il entraîne une translation verticale d'ensemble

de l'image, qui est ensuite évacuée par le calcul. Mais près du sol les gradients thermiques sont plutôt perpendiculaires aux surfaces, et donc pas nécessairement verticaux. Et surtout on peut observer, en périodes de forts échanges de chaleur, des phénomènes de turbulences qui provoquent une forme d'instabilité de l'image, bien connue des topographes. Intuitivement, on peut donc s'attendre à une importante dégradation des résultats de la corrélation diachronique. À noter tout de même que, si le point de vue est bien choisi, la visée peut rester constamment à plusieurs mètres au-dessus du sol (typiquement 5 m pour des ponts sur autoroute, cas de nos essais), ce qui permet de rester loin des zones à réfraction anormale, généralement proches du sol.

Afin de comparer des situations équivalentes, les tests ont été menés sur des situations où aucune charge n'était sur le pont, et nous avons cherché à voir de combien on s'écartait de valeurs nulles sur les déplacements verticaux. Une moyenne glissante a été effectuée sur une fenêtre 100 px x 100 px, soit environ 1 m². Les résultats présentés correspondent à une seule série de mesures, par temps estival chaud donc pas particulièrement favorable. Leur valeur statistique est faible puisque l'opération n'a été menée qu'une seule fois. Mais il faut en retenir la précision obtenue, entre 0.01 et 0.02 px d'écart-type, soit dans ce cas entre 0.1 et 0.2 mm.

Il est donc certain que dans des conditions météorologiques a priori favorables (p. ex. temps nuageux, ou vent soutenu, etc.) ce résultat est reproductible, ce qui représente une performance métrologique remarquable en termes de précision, de simplicité et de coût.

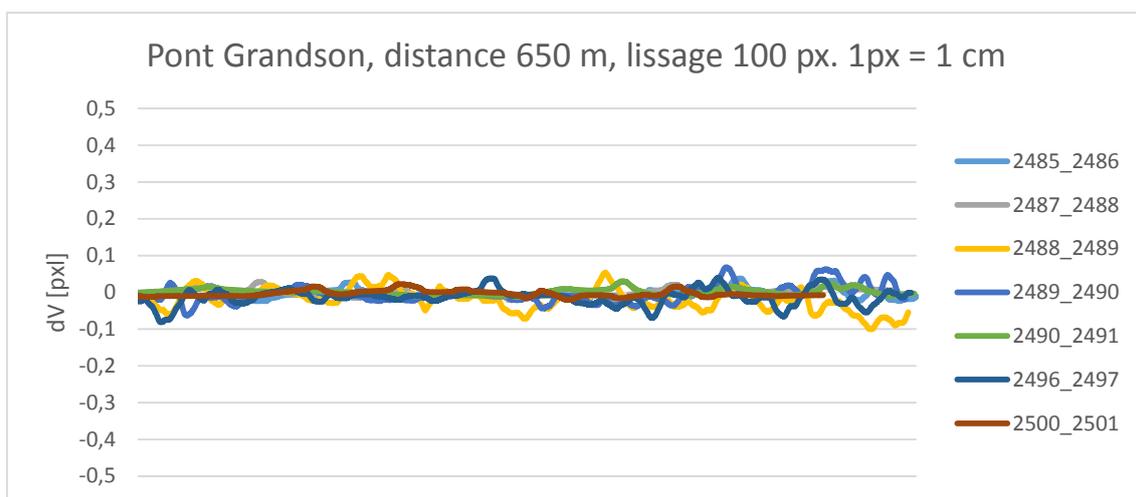


Fig. 7: Résultats de corrélations diachroniques successives sur un pont réputé parfaitement immobile, depuis une distance de 650 m (1 px = 1 cm), avec une moyenne glissante effectuée sur 1 m² de surface apparente du tablier. La déformation mesurée ainsi présente un écart-type de l'ordre de 0.01 px, soit 0.1 mm.

4 Conclusions

Les résultats obtenus permettent de bien définir les performances et les limitations de la corrélation diachronique pour le suivi des déformations de ponts. Au-delà de la finesse des déplacements identifiables, la possibilité de réaliser des mesures sans contact et sans pose de cibles est très économique et d'un grand intérêt. Quelques éléments limitants sont maintenant bien identifiés :

Premier point : on ne sait finalement détecter, avec ce procédé, que des déformations dans le plan perpendiculaire à l'axe optique, et non isotropes à l'intérieur de ce plan. Mais ceci est pleinement satisfaisant pour la plupart des ponts, où l'on va rechercher des flexions, des cisaillements, des ouvertures de fissures, etc., et où aucune déformation à caractère isotrope ne présente généralement d'intérêt. L'observation peut être envisagée même à des distances très importantes de l'ouvrage surveillé, puisqu'on dispose aujourd'hui d'une gamme considérable de focales pour les appareils photo courants, y compris de très longues focales, et puisque nous avons pu voir que la réfraction peut ne pas dégrader les résultats, au moins dans des situations météo favorables.

Par ailleurs, des problèmes liés aux changements d'éclairage pourraient apparaître, mais ce phénomène n'a pas encore été observé durant nos expérimentations. Au pire, ces problèmes pourraient alors être réglés par l'utilisation de prises de vue nocturnes, puisqu'un éclairage artificiel peut être très reproductible.

Nous avons également noté que la partie étudiée ne devait pas être masquée partiellement par un élément aléatoire comme des arbres, sauf à compliquer beaucoup le traitement (en détournant manuellement ces éléments).

En outre, il est généralement recommandé de travailler avec des optiques dont la distorsion est étalonnée. Mais avec des protocoles de mesures assurant une bonne répétitivité du pointage, cette exigence peut être contournée. Ainsi, une simple webcam pourrait-elle devenir un instrument de mesure ultra-précis.

En s'approchant du centième de pixel, la précision des mesures peut atteindre le dixième de millimètre dans de bonnes conditions d'acquisition. La mise en place d'un système de mesure d'une telle précision, à haute vitesse (les appareils courants offrent des rafales allant jusqu'à 10 images/s en pleine résolution, ce qui est considérablement meilleur que l'exploitation d'une video) et à faible coût devrait donc en faire sous peu un outil indispensable dans le monde de l'auscultation.

5 Bibliographie

- COMTE, J. & KASSER, M., 2014: Performances et limitations de la corrélation diachronique d'images pour les ouvrages d'art. *Géomatique Suisse* **11**, 488-492.
- DESEILLIGNY, M.P., 2007: Micmac, un logiciel pour la mise en correspondance automatique d'images dans le contexte géographique. *Bulletin d'Information Scientifique et Technique de l'IGN n°77*.
- VAN PUymbroeck, N., MICHEL, R., BINET, R., AVOUAC, J.P. & TABOURY, J., 2000: Measuring earthquakes from optical satellite images. *Applied Optics* **39** (20), 3486-3494.

Geospatial Augmented Reality als interaktives Erlebnis im Museum am Beispiel der Swissarena App

ROBERT WÜEST¹, MICHAEL ZWICK¹ & STEPHAN NEBIKER¹

Zusammenfassung: Augmented Reality (AR), als aufstrebende Technologie, ermöglicht die Überlagerung von digitalen Inhalten auf einem Live-Kamerabild. Dieser Beitrag untersucht die spezifischen Anwendungsmöglichkeiten von AR im geografischen Kontext unter Verwendung von natürlichen Markern. Dazu gehört die Generierung von optimierten aber dennoch flächendeckenden Kombinationen von Markerbildern zur Steigerung der Robustheit. Aus den Untersuchungen entstand ein Prototyp einer Museums-App für Smartphones und Tablets. In der App werden dem Orthophoto Inhalte mit geografischem Bezug überlagert. Bei den Inhalten handelt es sich teilweise um Live-Daten wie beispielsweise Transponderdaten von Flugzeugen oder aktuelle Wetterdaten. Für die Interaktion mit den Inhalten können sich Anwenderinnen und Anwender der App im Raum bewegen und verschiedene Eingabemöglichkeiten des verwendeten mobilen Geräts nutzen.

1 Einleitung

Am Thema Augmented Reality (AR), der Erweiterung der Realität durch ortsbezogene digitale Inhalte, wurde schon lange vor der Erfindung von Google Glass und ähnlichen Produkten intensiv geforscht und entwickelt (ARTH et al. 2015). Eine der ursprünglichen Definitionen von AR durch AZUMA (1997) als System, welches eine Kombination von Realität und Virtualität, Interaktivität in Echtzeit sowie die Erfassung einer dreidimensionalen Umgebung bietet, hat bis heute ihre Gültigkeit. Aufgrund der hohen Anforderungen an Hardware und Bedienungsergonomie, war die Verwendung von AR bis vor ca. 10 Jahren noch stark auf Festinstallationen und Spezialhardware angewiesen (PAELKE 2007). Mit der Etablierung von Smartphones mit hoher Rechenleistung und integrierter Sensorik ist die Technologie heute weitgehend massenmarktauglich.

In Museen sind multimediale Ergänzungen zu Exponaten und Installationen bereits weit verbreitet. Dabei bietet AR die Möglichkeit, Exponate individuell und interaktiv zu erkunden und dabei neue Formen der Informationsvermittlung zu erschliessen. Im Fokus unserer Untersuchungen steht die Swissarena im Verkehrshaus der Schweiz, dem nationalen Museum und Erlebnisraum für Verkehr und Mobilität. Die Swissarena beherbergt eine eindruckliche, begehbare Orthophotokarte der Schweiz im Massstab 1:20'000. Mit diesem Projekt und der Entwicklung einer AR App für das persönliche Smartphone oder Tablet soll ein völlig neuartiges, interaktives Erlebnis der Swissarena ermöglicht werden.

Wir beginnen den Beitrag mit einer kurzen Übersicht über frühere und aktuelle Forschung und Entwicklungen im Bereich Geospatial AR. Anschliessend stellen wir die Swissarena und ein entsprechendes Testfeld vor. In der Folge präsentieren wir Untersuchungen zur Verwendung eines riesigen Orthophotomosaiks als sogenannt "natürlichen" AR-Marker und zur Evaluation

¹ Fachhochschule Nordwestschweiz Hochschule für Architektur Bau und Geomatik, Institut Vermessung und Geoinformation, Gründenstrasse 40, CH-4132 Muttenz,
E-Mail: [robert.wueest, michael.zwick, stephan.nebiker]@fhnw.ch

geeigneter Inhalte für die AR App. Nach der Vorstellung der Architektur der realisierten App und der Einbindung von Live-Inhalten geben wir einen Ausblick auf zukünftige Weiterentwicklungsmöglichkeiten.

2 Verwandte Arbeiten

Frühe, umfassende Übersichten über Augmented Reality und deren grundlegende Definitionen, Prinzipien und Anwendungen geben AZUMA (1997) und AZUMA et al. (2001). PAELKE (2007), vermittelt einen guten Überblick über erste konkrete AR-Anwendungen im raumbezogenen bzw. geographischen Kontext und identifiziert dabei augmentierte Karten als besonders interessantes Anwendungsfeld. Die frühen Beispiele von projektions-basierten augmentierten Karten (z.B. von REITMAYR et al. 2005) haben nicht zuletzt dank der Augmented Reality Sandbox (REED et al., 2014) eine gewisse Popularität erlangt. Im Rahmen der Projektinitiative GeoAR (Geospatial Augmented Reality) des Instituts Vermessung und Geoinformation (IVGI) der Fachhochschule Nordwestschweiz (FHNW) wurde die Idee der Augmented Maps in den vergangenen Jahren aktiv weiter verfolgt (LOESCH et al. 2015). Dabei wird erfolgreich auf bildbasierte Georeferenzierung unter Verwendung der Karten als natürliche Marker gesetzt. In Abb. 1 wird die Überlagerung einer Landeskarte mit einem texturierten 3D-Landschaftsmodell und mit Live-Daten des öffentlichen Verkehrs demonstriert. Die Arbeiten von LOESCH et al. (2015) bilden die Basis für die weiteren Untersuchungen zum Markertracking in diesem Beitrag. Die Entwicklung der nachfolgend vorgestellten AR App profitierte zudem von der parallel laufenden Entwicklung der App «das Unsichtbare sichtbar machen» für die Römerstadt Augusta Raurica (CHRISTEN et al. 2016). Die neue Besucher-App verwendet unterschiedliche AR-Techniken zur Erkundung der damaligen historischen Stadt. Konkret erlaubt beispielsweise ein interaktives Stadtpanorama den Blick auf eine damalige Strassenszene am heutigen Betrachtungsort (Abb. 2).



Abb. 1: Augmented Maps Prototyp im Rahmen des Forschungsprojekts GeoAR



Abb. 2: Augusta Raurica App mit ortsbezogener Überlagerung des historischen Stadtpanoramas

3 Das Untersuchungsobjekt

3.1 Swissarena – Verkehrshaus der Schweiz Luzern

Das Verkehrshaus der Schweiz befindet sich in Luzern und ist mit jährlich rund 500'000 Eintritten das meistbesuchte Museum in der Schweiz. Die Swissarena ist ein Ausstellungsraum im Verkehrshaus der Schweiz. Das Kernstück der Swissarena ist die rund 200 m² grosse bedruckte Bodenfläche im Massstab 1:20'000. Deren originale, reale Bodenpixelauflösung (Ground Sampling Distance) beträgt 30 cm, woraus eine Bilddimension von 1'333'333 x 800'000 Pixel und eine Datenmenge von rund 1.66 TB resultiert. Der Orthophotodruck der Schweiz und des nahen Auslandes von Mappuls AG / Axon Vibe AG kann von einer Balustrade aus betrachtet und mit speziellen Pantoffeln begangen werden. Teiltransparente Strassen- und Schienengeometrien überlagern das Orthophoto. Dennoch ist die Orientierung für viele Gäste alles andere als einfach. Um sich spielerisch mit der Geografie der Schweiz auseinanderzusetzen existieren tellergrosse Ausschnitte des Orthophotos, die an der korrekten Lage platziert werden sollen. In der Vergangenheit existierten zudem kleine Wagen, die auf einem Display verschiedene Wanderrouen je nach Standort im Raum anzeigten. Die Positionsbestimmung der Wagen erfolgte Location-based mittels Lichtsignalen im Infrarot-Bereich.

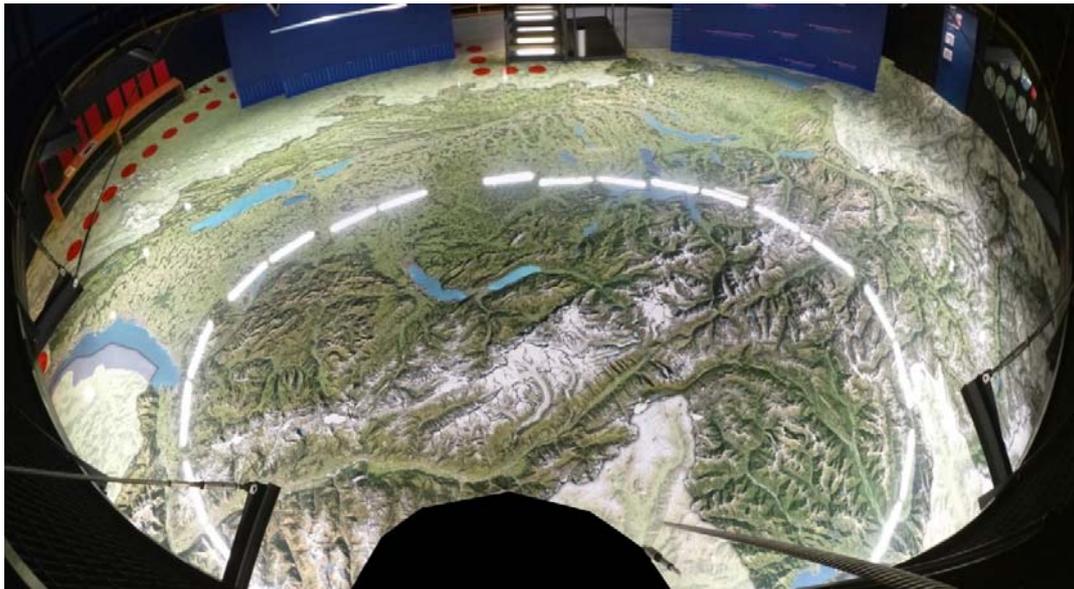


Abb. 3: Ausstellungsraum Swissarena im Verkehrshaus der Schweiz mit Orthophoto als Bodenbelag von der Balustrade aus betrachtet

3.2 Testfeld an der FHNW

Für die nachfolgenden Untersuchungen und für die Entwicklung der swissAREna App wurde an der FHNW ein Testfeld im gleichen Massstab eingerichtet. Dieses Testfeld mit einer Dimension von 2.5 x 2.5 m enthält einen Ausschnitt der Zentralschweiz von 50 x 50 km und basiert auf den identischen Bilddaten wie der Boden in der Swissarena (vgl. Abb. 4). Der Bodenbelag besteht aus einem bedruckten und direkt verklebten Bodenlaminat und weist andere Belichtungsverhältnisse

auf als in der Swissarena, was jedoch aus Sicht der Robustheit der zu realisierenden App durchaus ein erwünschter Parameter ist.



Abb. 4: Swissarena Testfeld an der FHNW mit einem Ausschnitt der Zentralschweiz

4 Landesweites Orthophoto als AR-Tracker

4.1 Ziele und Anforderungen

Das Orthophoto ist das zentrale Ausstellungsobjekt in der Swissarena. Für eine Lokalisierung von mobilen Geräten sollen darauf keine künstlichen Marker angebracht werden. Es wird deshalb angestrebt, Elemente des bestehenden Ausstellungsraumes als natürliche Marker zu nutzen. Wie im GeoAR-Projekt mit Landeskarten von LOESCH et al. (2015) demonstriert, wird in der Swissarena das Orthophoto als natürlicher Marker genutzt. Die Robustheit gegenüber Reflexionen der Leuchtkörper auf der glatten Oberfläche des Schutzlacks (vgl. Abb. 3) aber auch partieller Verdeckungen durch andere Besuchende soll gewährleistet sein. Die App soll über die üblichen Kanäle bezogen und auf gängigen Smartphones und Tablets verwendet werden können.

Das Orthophoto der Swissarena ist mit seinen rund 200 m² und der resultierenden Bilddimensionen und Datenmenge sehr gross. Die Ressourcen (Speicherplatz und Rechenleistung) auf einem üblichen Smartphone sind jedoch beschränkt und bei den Untersuchungen und der Umsetzung zu berücksichtigen. Infolge der Raumanordnung und -ausdehnung in der Swissarena variiert der Abstand vom Marker zur Kamera je nach Standort um mehrere Meter und die Nutzung der App erfolgt über einen grossen Distanz- bzw. Massstabsbereich. Von der Balustrade aus ist ein Grossteil des Orthophotos im Blickfeld, bei der lokalen Verwendung direkt über dem Orthophoto ist jedoch nur ein kleiner Ausschnitt sichtbar. Die Kamera registriert demzufolge unterschiedlich detailreiche Ausschnitte des Markers.

4.2 Lösungskonzept und Untersuchungen

Um das Projektionszentrum der Kamera bestimmen zu können werden bekannte Muster, Marker genannt, aus der Realwelt benötigt. Ein Marker besteht aus einer Bilddatei und der daraus erzeugten Merkmalsdatenbank. In der Merkmalsdatenbank sind markante Bildpunkte beschrieben. Zur Erzeugung der Merkmalsdatenbank wurde Vuforia SDK von PTC Inc. verwendet. Das verwendete Bildmaterial sollte viele verteilte Details mit markanten Punkten und hohen Kontrast aufweisen, die jedoch nicht wiederholend sind. Während der Verwendung einer markerbasierten AR App werden aus dem Kamerabild laufend Merkmale detektiert und mit der Merkmalsdatenbank der bekannten Marker verglichen. Bei einer Übereinstimmung können das Projektionszentrum der Kamera und deren Ausrichtung bestimmt werden. Das Tracking unterstützt den kontinuierlichen Prozess der Positions- und Orientierungsbestimmung mithilfe der Bildverfolgung im Kamerabild.

Um ein stabiles AR-Tracking über eine solch grosse Fläche zu realisieren, bedarf es eines systematischen Ansatzes um einerseits eine komplette Abdeckung der zu trackenden Fläche zu erhalten und andererseits eine Optimierung zu erreichen. Optimierung wird hier vor allem auf die Datenmenge bezogen was die Performance der App negativ beeinflussen könnte. Für die Kombination von im 3D-Raum relativ zueinander orientierter Marker ist von der Vuforia SDK ein sogenannter Multimarker vorgesehen. Hier wird ein solcher Multimarker jedoch zusätzlich verwendet um verschieden detaillierte Marker zu organisieren. Aus dem Orthophoto wurden mit der Quadtree-Struktur Marker als Kacheln für verschiedene Levels of Detail (LOD) generiert wie dies LOESCH et al. (2015) als Weiterentwicklungsvariante für Multimarker beschrieben haben (vgl. Abb. 5).

Durch die Vernachlässigung von überflüssigen Markern (z.B. ausserhalb des Interessenbereichs) und Markern die viele stufenübergreifend redundante Merkmale aufweisen, lässt sich deren Anzahl stark reduzieren. Dazu wurden verschiedene Kombinationen mit einer reduzierten Anzahl Marker erstellt und die daraus resultierende Robustheit und Abdeckung im Ausstellungsraum und auf dem Testfeld untersucht.

4.3 Resultate

Durch die Verwendung von mehreren Markern, die in einem Multimarker zueinander orientiert werden (Verschiebung, Drehung und Skalierung), kann die Robustheit der Markererkennung gesteigert werden. Untersuchungen von ZWICK (2016) haben gezeigt, dass durch die Verwendung von Markern unterschiedlicher Grössen und Auflösungen, sowie deren Integration in einem Multimarker, die Robustheit der Positions- und Orientierungsbestimmung gesteigert werden kann. Werden zu viele Marker verwendet, kann die Rechenleistung an ihre Grenzen stossen, denn für die Merkmale in jedem Bild (Frame) einer Aufnahme wird eine Übereinstimmung in der Merkmalsdatenbank gesucht. Aus den Untersuchungen resultiert ein Multimarker, bestehend aus 51 Markern auf drei verschiedenen LOD-Stufen (vgl. Abb. 6). Zu einer Auswahl an Markern aus den LOD-Stufen 1 und 2 kommt ein einzelner Marker über die ganze Schweiz dazu (LODCH). Dieser zusätzliche Marker ermöglicht auch die Nutzung der App auf einem A3-Ausdruck.

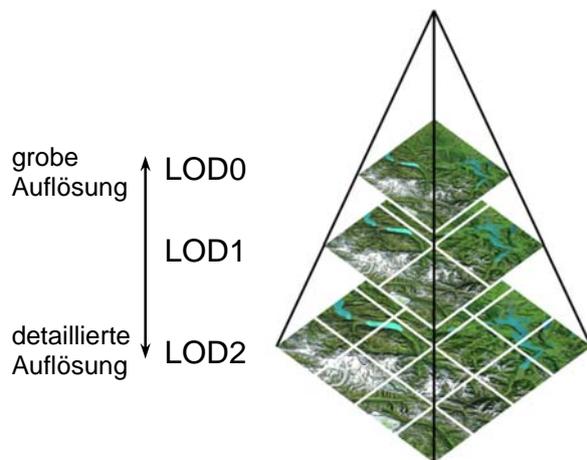


Abb. 5: Multimarker als Bildpyramide mit Quadtree-Kachelstruktur und LOD-Stufen des Orthophotos im Bereich der Zentralschweiz

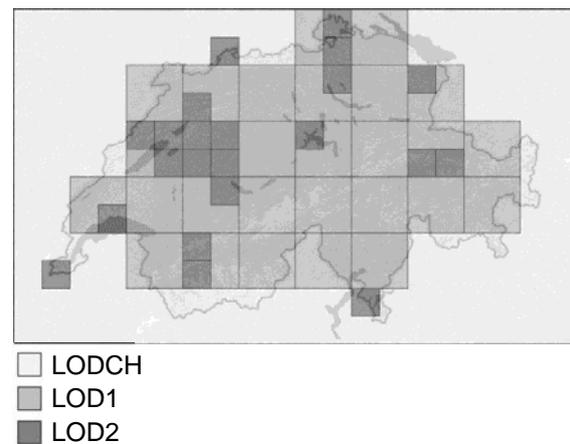


Abb. 6: Markerkonfiguration für den Multimarker bestehend aus drei verschiedenen LOD-Stufen und einer Ausdehnung von 400 x 240 km (Zwick, 2016)

5 Inhalte

5.1 Zielgruppe

Im Verkehrshaus der Schweiz stehen inhaltlich die Kernthemen Verkehr und Mobilität im Vordergrund. So musste bei der Evaluation der zu überlagernden Inhalten dem entsprechend Sorge getragen werden. Die App richtet sich in erster Linie an alle Museumsbesucherinnen und -besucher unabhängig von deren Alter. Einschränkender Faktor dabei ist die Verfügbarkeit eines Smartphones oder Tablets. Die Besucher sollen durch die App angeregt werden forschend die Landkarte der Schweiz zu erkunden. Viele Besucherinnen und Besucher des Verkehrshauses sind Schulklassen (ca. 10%) oder Familien. Den jüngeren Menschen fällt der Zugang zu technologischen Geräten in einem spielerischen Kontext oft leicht, was den Einsatz einer App begünstigt.

5.2 Inhaltskonzept

Die augmentierten Inhalte der App stellen das Wesentliche der App dar. Dabei standen diverse Forderungen im Raum. Einerseits müssen die Daten einen informationellen Mehrwert bieten und andererseits aktuell sein. Der wichtigste Faktor stellt jedoch der räumliche Bezug der Daten dar. Für die Entwicklung der swissARena App wurden diverse Inhalte getestet und danach in thematischen Teilen umgesetzt respektive implementiert. Folgende Daten wurden bei den Untersuchungen in Betracht gezogen:

- Flugverkehr: Transponderdaten mit aktuellen Informationen
- Wetter: Wetterstationsdaten
- Adressen: Geocoding / Reverse-Geocoding
- Öffentlicher Verkehr: Stationen und Fahrpläne
- Motorisierter Verkehr: Aktuelle Stausituation

5.3 Realisierte Inhalte

5.3.1 Adressen & Geocoding

Das Adressmodul bietet die Möglichkeit Orte oder Adressen gezielt zu suchen und führt die Anwenderinnen und Anwender mittels Richtungsindikator an den entsprechenden Ort. Weiter ist mittels Reverse-Geocoding ein Abrufen des Orts im Zentrum des Kamerabildes möglich.

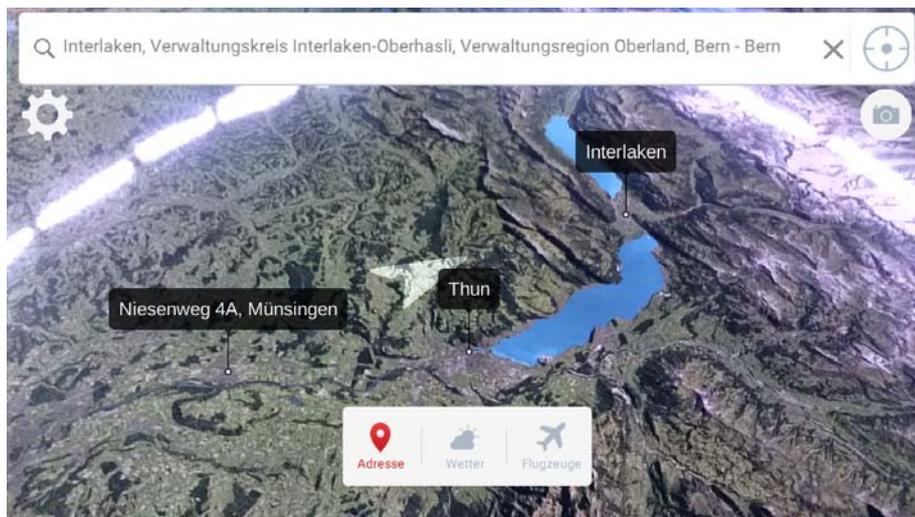


Abb. 7: Adresssuche mit Richtungsindikator (Pfeil im Zentrum)

5.3.2 Wetter

Aus Radarniederschlagsbildern und ortsbezogenen Wetterdaten wird eine Überlagerung der aktuellen Schweizer Wetterverhältnisse erzeugt. Das Wettermodul stellt zwei verschiedene Datensätze dar: Für bestimmte Orte wird ein Wettersymbol und die aktuelle Temperatur aus den Wetterstationsdaten angezeigt. Mit Partikeleffekten wird zudem eine simulierte, jedoch akkurate Wolkendecke aus den Radarniederschlagsbildern dargestellt.



Abb. 8: Wetterdatenüberlagerung mit Wolkensimulation

5.3.3 Flugverkehr

Mittels aktueller Transponderdaten können die Flugzeuge im schweizerischen Luftraum registriert und verfolgt werden. Auf Abb. 9 ist zu sehen wie durch Interaktion mit einem Flugobjekt, dargestellt als vergrössertes 3D-Modell, weitere interessante Informationen zum Flugobjekt abgerufen werden können. Die Position des 3D-Flugzeugs entspricht dabei massstabsgetreu der effektiven Position der Transponderdaten. Weil die Daten in Intervallen abgerufen werden, wird die kontinuierliche Fortbewegung des Flugzeugs durch Interpolation simuliert.



Abb. 9: Flugverkehr über der Zentralschweiz mit Fluginformationen

6 Die swissARena App

6.1 Architektur

Durch die Evaluation verschiedener AR-Frameworks musste eine geeignete Plattform gefunden werden die die genügende Performance und Robustheit des Markertrackings auf mobilen Geräten gewährleistet. Frameworks wie ARToolkit (<http://artoolkit.org>) oder OpenCV (<http://opencv.org>) bieten Lösungen für desktopbasierte Anwendungen. Zur Realisierung von AR auf verschiedenen mobilen Plattformen gleichzeitig (Multiplattform Support für Android, iOS etc.) wurde letztlich auf das Framework Vuforia (<http://www.vuforia.com>), in Kombination mit der 3D-Engine Unity3D (<https://unity3d.com>), zurückgegriffen. Dadurch war es möglich, effizientes AR auf mobilen Geräten zu realisieren und gleichzeitig state-of-the-art 3D-Technologie zu nutzen (vgl. Abb. 10).

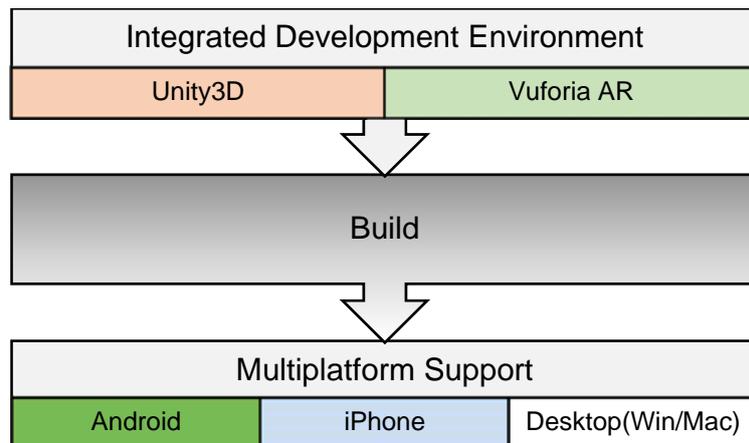


Abb. 10: Entwicklungsarchitektur für AR Apps auf mobilen Geräten

6.2 Geodaten-Einbindung

6.2.1 Datenquellen

Für die umgesetzten drei Inhaltsmodule wurden die folgenden aktuellen und wo möglich frei verfügbaren Datenquellen verwendet:

- **Adressen & Geocoding:** Für das Geocoding beziehungsweise Reverse-Geocoding wird der OpenStreetMap-Dienst Nominatim verwendet (<http://nominatim.openstreetmap.org>). Da wir uns dank der räumlich korrekt platzierten Marker virtuell im Schweizer Bezugsrahmen befinden, können die Koordinaten ohne weitere Transformationen verwendet werden. .
- **Wetter:** Ortsgebundene Wetterdaten können durch OpenWeatherMap (<http://openweathermap.org>) bezogen werden und dienen bei diesem Projekt als Datengrundlage für die Wettersymbole mit den Temperaturangaben. Die Radarbilder werden aus einer kommerziellen Quelle bezogen. Bei beiden Datenquellen handelt es sich um aktuelle Daten mit einem Aktualisierungszyklus von wenigen Minuten.
- **Flugzeuge:** Jedes zivile Luftfahrzeug im Schweizer Luftraum muss über einen sogenannten Transponder verfügen und dessen Daten können mit einem geeigneten Empfänger unverschlüsselt empfangen werden. Die hier verwendeten Daten stammen aus einem Netzwerk von Empfängern und werden freundlicherweise von SwissFlightLive (<http://www.flightlive.net>) zur Verfügung gestellt. Die Daten werden in Zeitabständen von 3-5 Sekunden aktualisiert.

6.2.2 Streaming & Caching

Um die Daten effizient den Benutzerinnen und Benutzer über die App zugänglich zu machen, wurde ein Caching bzw. Streaming-Service realisiert. Dabei soll sichergestellt werden, dass selbst bei einem Ausfall der Rohdatenquelle ein Betrieb der App weiterhin gewährleistet werden kann.

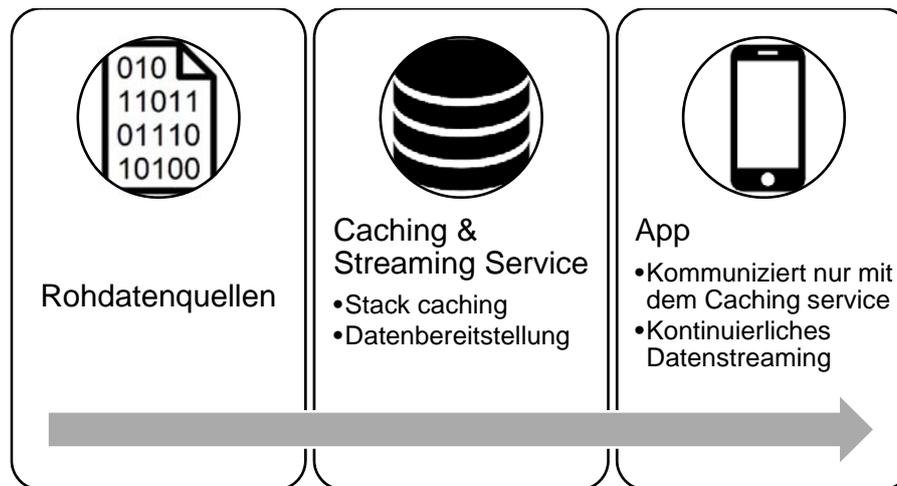


Abb. 11: Caching & Streaming Infrastruktur für die swissARena App

7 Fazit & Ausblick

In diesem Beitrag wurde eine innovative Augmented Reality Anwendung unter Verwendung eines sehr grossen ausgedruckten Orthophotos als natürlicher Marker vorgestellt. Durch den Einsatz eines optimierten Multimarkerkatalogs mit mehreren Skalierungen konnte ein robustes Tracking der gesamten Markerfläche erreicht werden und es wurde gezeigt welche Faktoren bei der Robustheitssteigerung massgebend sind. Dies ist gerade für die Anwendung in der Swissarena von zentraler Bedeutung – einerseits wegen der partiellen Sichtbarkeit des sehr grossen Markers und wegen Verdeckungen durch anderen Museumsbesucherinnen und -besucher andererseits. Im Beitrag wurde die Integration sehr unterschiedlicher raumbezogener Live-Inhalte untersucht und demonstriert. Die Resultate zeigen auf, dass heute anspruchsvolle AR-Anwendungen auf Smartphones und Tablets realisierbar sind, die sowohl hoch genaues, bildbasiertes Tracking als auch komplexe 3D-Inhalte in Echtzeit unterstützen.

Aufbauend auf den vorgestellten Untersuchungen soll die swissARena App im Sommer 2016 veröffentlicht werden. Dabei ist eine laufende Erweiterung um neue Inhalte und Funktionen geplant. Konkret wird der Einsatz dieser Technologie sowohl für Unterrichtszwecke in Schulen selbst, als auch vor Ort im Museum untersucht. Geospatial Augmented Reality weist ein grosses, noch wenig erschlossenes Potential auf, um den intrinsisch motivierten Entdeckungsdrang von Kindern und Jugendlichen für didaktische Zwecke nutzbar machen.

8 Danksagung

An dieser Stelle möchten wir der Firma Axon Vibe AG für die gute Zusammenarbeit und die Bereitstellung der digitalen Orthophotos sowie dem Verkehrshaus der Schweiz für die tatkräftige Unterstützung des Projekts herzlich danken.

9 Literaturverzeichnis

- ARTH, C., GRUBER, L., GRASSETT, R., LANGLOTZ, T., MULLONI, A., SCHMALSTIEG, D. & WAGNER, D., 2015: The History of Mobile Augmented Reality, Technical Report ICG–TR–2015-001, University of Technology, Graz.
- AZUMA, R., 1997: A survey of augmented reality. *Presence: Teleoperators and Virtual Environments* **6** (4), 355-385.
- AZUMA, R., BAILLOT, Y., BEHRINGER, R. & FEINER, S., 2001: Recent advances in augmented reality. *IEEE Computer Graphics and Applications* **21** (6), 34-47.
- CHRISTEN, M., BLASER, S., LOESCH, B. & NEBIKER, S., 2016: Prozedurale Modellierung von historischen 3D-Stadtmodellen und deren Aufbereitung für 3D-Visualisierungen in einer Smartphone-Applikation am Beispiel der römischen Stadt Augusta Raurica. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 25*, T. Kersten (Hrsg.), Dreiländertagung der DGPF, der OVG und der SGPF in Bern, 7.-9. Juni 2016, 217-229.
- LOESCH, B., CHRISTEN, M., WÜEST, R. & NEBIKER, S., 2015: Geospatial Augmented Reality – Lösungsansätze mit natürlichen Markern für die Kartographie und die Geoinformationsvisualisierung im Außenraum. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 24*, T. Kersten (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 16.-18. März 2015 an der Universität zu Köln, Tagungsband auf CD-ROM, 89-97.
- PAELKE, V., 2007: Mixed Reality - Benutzungsschnittstellen für die raumbezogene Informationsverarbeitung. *GIS Geo-Informationssysteme*, **7** (7), 14-21.
- REED, S. E., KREYLOS, O., HSI, S., KELLOGG, L. H., SCHLADOW, G., YIKILMAZ, M. B., SEGALE, H., SILVERMAN, J., YALOWITZ, S. & SATO, E., 2014: Shaping Watersheds Exhibit: An Interactive, Augmented Reality Sandbox for Advancing Earth Science Education. *American Geophysical Union (AGU), Fall Meeting 2014, Abstract No. ED34A-01*.
- REITMAYR, G., EADE, E. & DRUMMOND, T., 2005: Localisation and interaction for augmented maps. *4th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*, 120-129.
- ZWICK, M., 2016: AR App Swissarena – Untersuchungen zur erweiterten interaktiven Nutzung der Swissarena im Verkehrshaus der Schweiz mittels Augmented Reality. Masterthesis, FHNW Fachhochschule Nordwestschweiz.

Potential of Graph Databases in Representing and Enriching Standardized Geodata

AMGAD AGOUB¹, FELIX KUNDE² & MARTIN KADA¹

Abstract: The storage of spatial data in a database usually happens in relational database management system (RDBMS) such as PostgreSQL/PostGIS or Oracle Spatial. Due to the layer-based thematic modeling used in present-day Geographical Information Systems (GIS), the database schemas most often comprise a rather flat structure. However, the trend in spatial models, e.g. by the Open Geospatial Consortium (OGC), in international standards goes for more complex application schemas with well-defined objects, attributes, and relations. Mapping these (most often UML-based) data models into a RDBMS is challenging. Different solutions exist to guarantee a generic mapping, providing a more or less useful database schema to work with. But even these tools can fall short when dynamically supporting multiple user-defined extensions that some OGC standards have started to offer (e.g. CityGML). Therefore, a growing interest for modern NoSQL systems can be noted. They are, by nature, schema-less and often very well suited for web-based applications - albeit their limited spatial functionality, data migrating and testing is often easier as with RDBMS. In particular, graph databases provide a structure that could help dealing with this problem as the underlying data model of nodes and edges can natively represent a conceptual UML diagram. This paper will present a lightweight mapping approach that supports on-the-fly mapping and storage for various OGC standards (i.e. SensorML, CityGML, etc.) into Neo4j, a pure graph-oriented database, and ArangoDB, a recently-developed multi-model database. The chosen DBMS are evaluated in terms of usability and support for spatial data management.

1 Introduction

The Open Geospatial Consortium (OGC) has developed many implementation standards (Simple Features, Open Web Services, Geography Markup Language) that are now widely adopted in commercial and open source software. This has helped to design, plan, and build interoperable spatial data infrastructures all over the world. Within the last 10 years new standards were introduced focusing on domain-specific location-based data models, ranging from sensor properties (SensorML) to urban modeling (CityGML)³. The data models of the Infrastructure for Spatial Information in the European Community (INSPIRE) are also greatly based on the work by the OGC. Therefore, the ability to read, store and write data following these application schemas will be a key to the success of such initiatives.

¹Technische Universität Berlin, Institut für Geodäsie und Geoinformationstechnik, Straße des 17. Juni 135, D-10623 Berlin, E-Mail: amgadagoub@gmail.com, martin.kada@tu-berlin.de

²Beuth Hochschule Berlin, Labor Rechner- und Informationssysteme, Luxemburger Straße 10, D-13353 Berlin, E-Mail: fkunde@beuth-hochschule.de

³<http://www.ogcnetwork.net/gmlprofiles>

2 Limitations for mapping into relational databases

Different solutions have been developed to store and manage standardized data models inside a spatial database. In most cases relational database management systems (RDBMS) are used because of their maturity, their wide acceptance and their capabilities in processing geodata (e.g. Oracle Spatial, PostGIS, SpatiaLite or SQL Server).

Mapping object-oriented OGC data models into compact relational schemas without losing information is a challenging task (“impedance mismatch”). For a generic automatic mapping the class graph of a UML model can be used to define a set of algorithms and transformation rules to map the syntactical structure into an entity-relationship model (ER) (RODRIGUEZ et al. 2011). This can be done on behalf of the XML Schema Definition (XSD) and/or the XML Metadata Interchange (XMI). Such an approach could lead to a relatively complex relational model with a high number of tables (e.g. deegree⁴, Go Loader⁵) or a very generic meta-model (e.g. SupportGIS⁶). Thus, writing queries could require using many (self-)joins either ways which has an effect on the performance. As a consequence, the deegree project provides an additional BLOB-Mode for fast data retrieval.

Finding a good balance between schema design and usability also requires a deep knowledge about the target domain which makes an automatic mapping even harder. The 3D City Database project⁷ applies a semi-automatic mapping, where the UML class diagram of the OGC standard CityGML is first simplified manually with respect to performance and later use cases before converting it into an entity relationship model. Despite the advantage of having a more query-friendly database schema, an additional manual mapping step reduces the flexibility in serving arbitrary OGC schemas. The CityGML standard itself allows for extending and creating feature classes via the Application Domain Extension (ADE) mechanism (VAN DEN BRINK et al. 2014), which is not yet supported by the 3D City Database at the moment.

3 NoSQL databases as an alternative?

NoSQL ("not only SQL") has become a synonym for a big family of either new database systems that have been developed to fulfill modern business requirements of the internet age or even for older DBMSs that do not store their data in tabular relations. A common characteristic by many NoSQL systems is a high flexibility for storing very heterogeneous data. This feature raises the question if using a NoSQL database might overcome the limitations described in chapter 2.

3.1 Document-oriented databases

Data that is described using an OGC standard is written to an instance document of the UML information model which is encoded in XML. These documents can directly be imported in a document-oriented database that is able to store any kind of documents in the same encoding. KOCH (2015) evaluated BaseX, a native XML database, as a possible candidate for storing

⁴<http://download.deegree.org/documentation/3.3.13/html/featurestores.html>

⁵<http://www.snowflakesoftware.com/products/goloader/>

⁶<http://www.cpa-software.de/index.php?do=bas&do2=db&lang=d> (in german)

⁷<http://www.3dcitydb.net>

CityGML data. BaseX provides a flexible schema control and is very fast on imports and exports. However, performance of analytical spatial queries has been considerably slower compared to an RDBMS (PostGIS).

MAO et al. (2014) created a 3D city model management that can scale well on cluster-based architectures. Therefore, they used MongoDB, a popular document-oriented database that manages data in collections of binary JSON documents. The conversion (unmarshalling) of CityGML files is done through a java binding library⁸. The database is comprised of three collections in order to serve three different kinds of queries (metadata, thematic and spatial features, 3D for visualization).

3.2 Graph databases

The underlying data model of graph databases follows the property-graph model consisting of nodes and edges which can natively represent a conceptual UML diagram⁹. FALKOWSKI & EBERT (2009) discussed a graph-based schema for integrated models of urban data which are represented as CityGML. In his opinion explicit graph representations can be used for processing semantics, geometry, topology and appearance of the city model. DELFOSSE et al. (2012) presented two methods of mapping a UML class diagram to a graph model. The first approach suggests an implicit mapping, where the data is stored without semantic schema information. The second approach includes the semantics into the graph data model as additional nodes which are connected to each instance of the corresponding feature class ("Instance of relationship").

3.3 Polyglot persistence and multi-model databases

Each type of NoSQL databases is suitable for specific data and use cases. Sometimes it can be necessary to use more than one database to encompass all user requirements. This strategy is called 'polyglot persistence'. BOYD (2015) presented a back end coupling MongoDB for data storage with Neo4j for shortest path queries¹⁰. A similar concept is thinkable for OGC standards where objects would form semantically classified collections of documents and their relationships (linkage, recursion, versioning etc.) would be expressed through graphs.

Orchestrating different database servers often creates an overhead in maintenance, software development and data throughput. Multi-model databases combine different storage models in one software to overcome these issues. Two modern DBMSs combining the document and the graph model are OrientDB and ArangoDB. However, both systems are still lacking a sophisticated spatial support. In 2013, a GraphGIS extension was introduced for OrientDB (HOUBIE 2013), but the project has been inactive since then.

4 Mapping concept

This section proposes a concept of mapping the schema definition of OGC standards to a property-graph model in order to store OGC-compliant instance documents into a graph database. The applied transformation rules are based on the theoretical work of DELFOSSE et al. (2012) with some minor changes. Fig. 1 provides a general overview of how a UML class diagram is mapped:

⁸<https://github.com/citygml4j/citygml4j>

⁹<http://lambdazen.blogspot.de/2014/01/from-entity-relationship-to-property.html>

¹⁰<http://neo4j.com/blog/polyglot-persistence-mongodb-wanderu-case-study/>

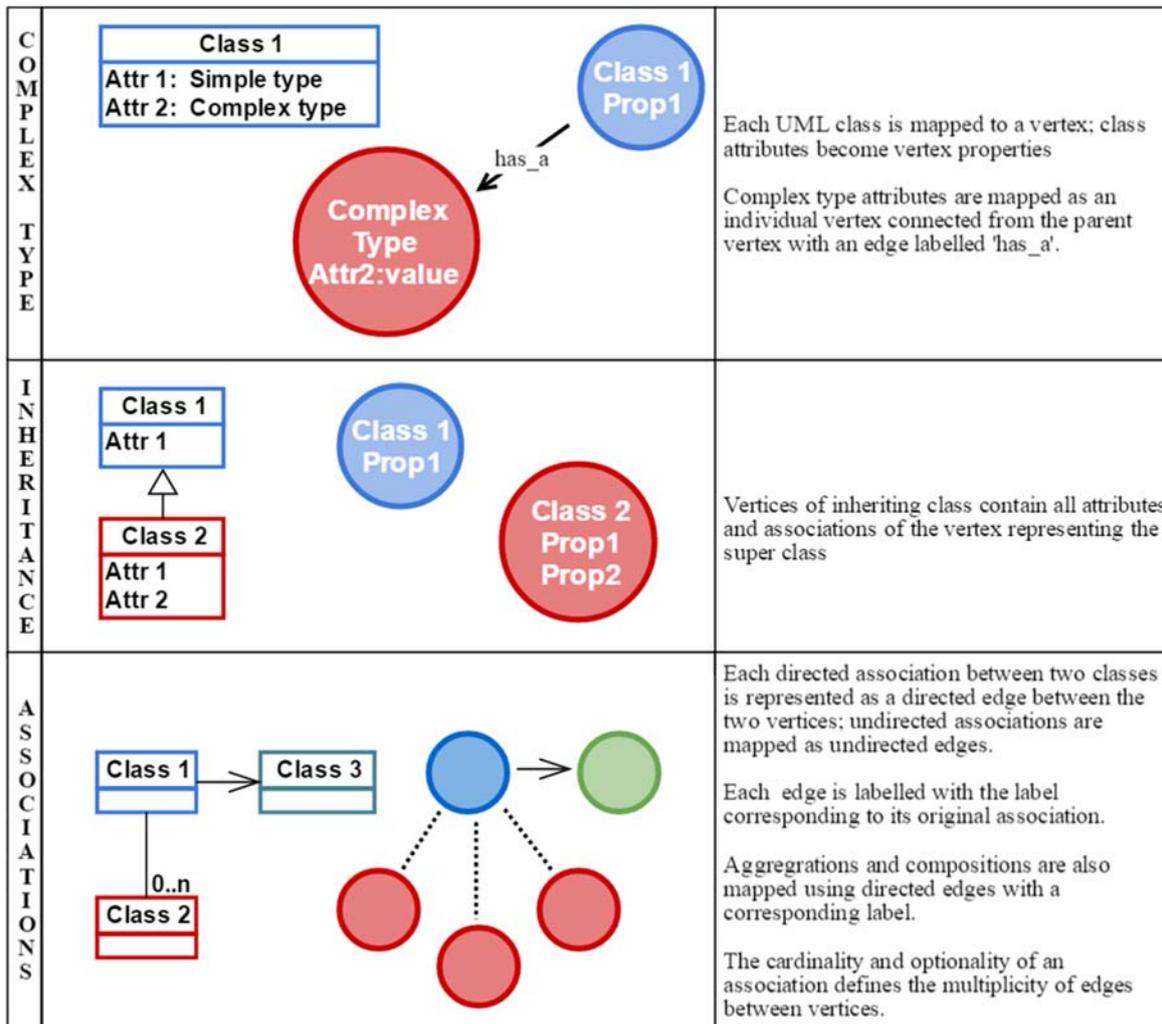


Fig. 1: Overview of the mapping context.

4.1 Generic import of OGC XML documents

A machine readable class graph of the UML model can be parsed from an XSD file. Fortunately, most of the OGC standards are already mapped to XSD files by the OGC itself and can be downloaded from its official home page: <http://schemas.opengis.net/> (2016). The download process can be automated while parsing an XML file if it contains a reference to the schema location in its root element¹¹.

An import process begins with binding the elements of an XSD schema file to a set of classes that represent the schema in the programming language used for the importer software. These classes define the mapping rules (see Fig. 1) which are further applied for unmarshalling the XML data into objects of the used binding. The mapping of XML data types should be type-safe to guarantee consistency within the target database.

¹¹<https://www.w3.org/TR/xmlschema-1/>

A challenge when mapping from an XSD file is to differentiate between features, relationships and non-primitive attributes in the later graph model because all of them are encoded as complex types (`xs:complexType`) and create the same structure of edges and vertices based on the transformation rules. For this reason, associations are mapped to connecting vertices between two features instead of having only an edge. For imports into a multi-model database this step requires additional software logic to create a working structure.

4.2 Writing bind objects to graph data models

The objects from the XML binding can be prepared for imports into either a pure graph database or hybrid graph DB that has the ability to use other data storage models in conjunction (see multi-model databases in chapter 3). A unique identifier is generated and added to each vertex as a property. For depending child nodes this identifier is used to add a reference property. It is a common strategy for graph databases to collect many vertices in bigger batches to reduce processing time and network traffic. The edges are written afterwards on behalf of the reference properties.

4.3 Importer architecture

The introduced mapping and import concept has been realized as a client-server application and is written in JavaScript. Although powerful XML binding frameworks exist for languages like Java (e.g. JAXB – Java Architecture for XML Binding), JavaScript had been chosen to create a lightweight prototype that is easy to deploy on every operating system and machine or device. The server application is built with Node.js, a runtime environment built on Google's V8 JavaScript engine. Node.js has a very active community that provides many different kinds of plugins and application programming interfaces (API) including also drivers for graph databases.

The importer core depends mainly on two Node.js packages: `Jsonix`¹² and `OGC Schemas`¹³. `Jsonix` is used to unmarshal XML into JSON objects (JavaScript Object Notation) and can be seen as the JavaScript equivalent to JAXB. `OGC Schemas` provides JAXB and `Jsonix` bindings for all OGC standards. However, it is not an official OGC project. The classes of these bindings provide the context for converting the schema elements of the XML file in corresponding JSON objects. This conversion step is type-safe because it depends on the file's schema and not a generic JSON parser. Two databases were chosen to test the mapping approaches explained in this chapter. As for a pure graph DBMS `Neo4j`¹⁴ was selected, because it offers a spatial plugin which would be a must-have feature for being a real alternative to relational geodatabase systems. For a multi-modal database the rather young project `ArangoDB`¹⁵ was tested, because it offers an additional JavaScript layer (`Foxx`) for creating additional software logic to keep the database schema consistent and OGC-compliant. `Foxx` integrates well into the proposed architecture.

¹²<https://github.com/highsource/jsonix>

¹³<https://github.com/highsource/ogc-schemas>

¹⁴<http://neo4j.com/>

¹⁵<https://www.arangodb.com/>

5 Use cases and results

5.1 Mapping to Neo4j

Vertices were imported into Neo4j and then edges were created with a query written in Cypher. Cypher has been developed by the Neo4j developers and is the default query language of Neo4j. It is declarative by nature and very similar to SQL. Queries to retrieve features were comparably shorter than their equivalents in RDBMS. The following code is an SQL query to retrieve the geometric and appearance properties of a single building from the 3D City Database database schema.

```
SELECT sg.*, tp.*, sd.*, a.* FROM building b
JOIN thematic_surface ts ON ts.building_id = b.id
JOIN surface_geometry sg ON sg.root_id = ts.lod2_multi_surface_id
LEFT JOIN textureparam tp ON tp.surface_geometry_id = sg.id
LEFT JOIN surface_data sd ON sd.id = tp.surface_data_id
LEFT JOIN appear_to_surface_data ats ON ats.surface_data_id = sd.id
LEFT JOIN appearance a ON a.id = ats.appearance_id
WHERE b.id = ?
```

The equivalent query in Cypher is as follows and yields the following result as shown in Figure 2.

```
Match (B:Building {B.id:id})-[:REL*]->(m)
return B,m
```

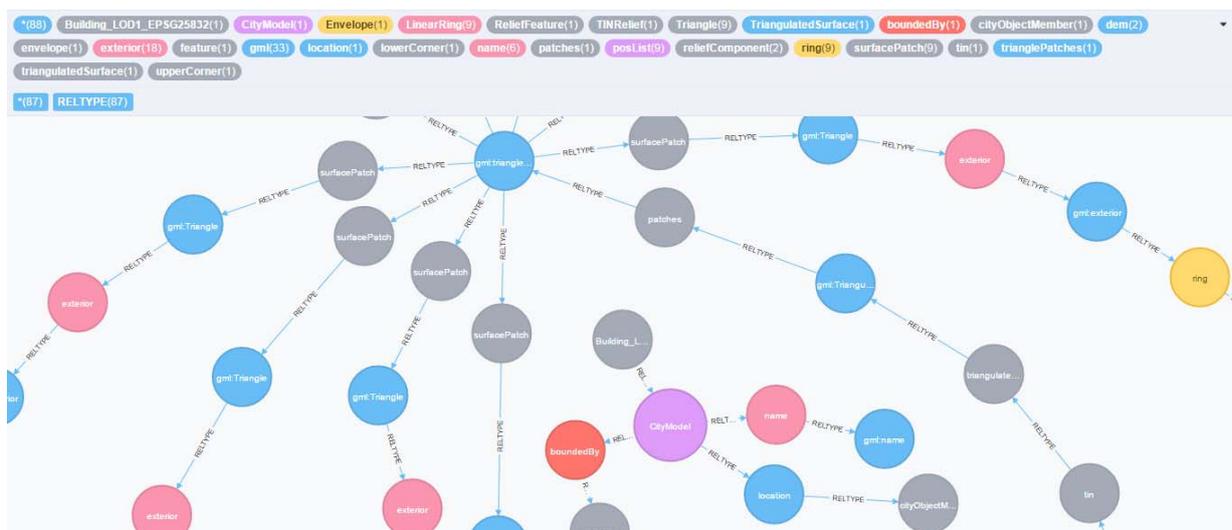


Fig. 2: CityGML dataset of a 3D building model ('Schloss Charlottenburg') in LOD 2 stored in Neo4j.

Some limitations were noted when using this mapping approach. Neo4j currently does not have support for GML geometry types out of the box. Therefore, GML data is stored in long chains of nodes as shown in Figure 3. If a GML-compliant geometry type would be integrated in Neo4j the resulting structure would be reduced to a single node or even to a node property with type GML.



Fig. 3: GML polygon represented as a chain of nodes in Neo4j database which corresponds to its original XML tree structure.

Neo4j offers a native support for a variety of spatial capabilities and can be extended further with the 'Neo4j-spatial' extension¹⁶. Neo4j-spatial provides the possibility to store geometries as a node property with the type WKB (Well Known Binary). However, they are currently limited to 2D representation and the only supported coordinate reference system (CRS) is WGS84 (World Geodetic System 1984) (BAAS 2012). This would not be sufficient for many domain-specific data models of the OGC, e.g. CityGML, which targets mostly 3D use cases. Coordinate transformations would always be mandatory if the CRS is not WGS84, which generates an overhead for imports and exports.

Another limitation refers to the object hierarchies of the XML document tree which can only be mapped to parent-child relationships in Neo4j. This does not gather the full potential of graph databases because relationships are not fully exploited in this case. Further development into this direction can leverage the usage of relationships in the semantic context.

5.2 Mapping to ArangoDB

ArangoDB can natively store a nested JSON object as a data entry inside a collection. Therefore, there is no need to disassemble the resulting JSON objects. Thus, the stored data would simply inherit the tree structure of the XML data. This makes indexing and querying considerably harder in contrast to simple scalar (flat) documents.

The schema-free nature of ArangoDB enables the storage of data of arbitrary schema definitions and attributes without validation. But the database can be extended seamlessly through the Foxx framework mentioned in chapter 4.3. It can be used to create a schema validation for each added schema which can be applied through a http request. This layer was tested on a basic JSON schema as a proof of concept. Test data could be validated successfully against the schema. The proposed architecture for mapping and importing OGC documents on-the-fly to ArangoDB is shown in Figure 4. The unmarshalling to JSON would be similar to the Neo4j example.

¹⁶<https://github.com/neo4j-contrib/spatial>

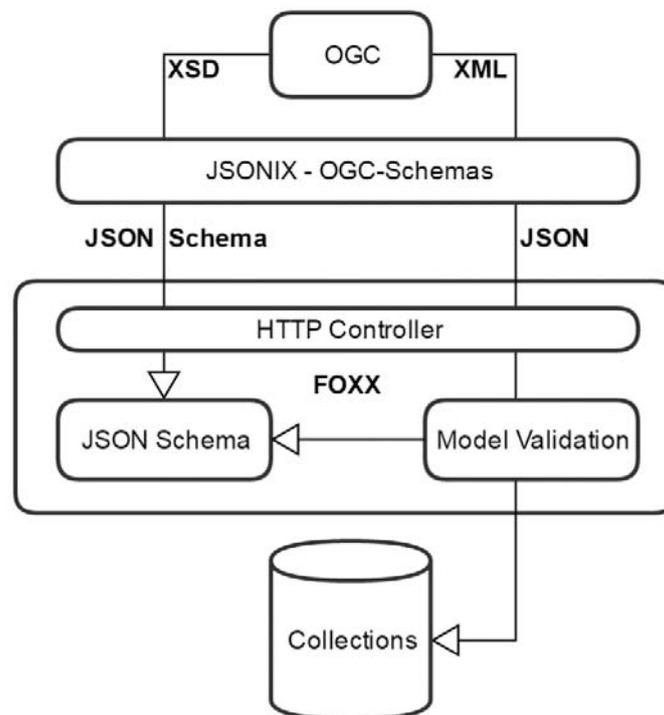


Fig. 4: Proposed mapping approach for importing OGC data into ArangoDB.

Several advantages should be noted when using the multi-model approach. First of all, each feature is mapped as an individual entry (document) in the database. Therefore, it is possible to retrieve the feature by calling its unique identifier yielding the following query as an equivalent for the aforementioned SQL query:

```

For x in Building
Filter x.id == id
Return x
  
```

Secondly, in many cases real world objects can be represented in multiple files formats. For example, a city model can be described using CityGML, but for web visualization formats like glTF or COLLADA are better suitable. It would not make much sense to store these files in the database but their metadata can be linked to the corresponding documents for the OGC base data. Currently, ArangoDB has more spatial limitations than Neo4j. It supports only points as a geometric representation, only one reference system (WGS84) and a few spatial query operations¹⁷.

6 Conclusion

In this paper, we presented a prototype that should prove the suitability of graph databases for storing and managing OGC standard documents. The schema-less design facilitates to store data from different domain-specific data models and any user-defined application extensions to these

¹⁷<https://docs.arangodb.com/SimpleQueries/GeoQueries.html>

models in one database. However, it is still possible to create a validation layer in between to guarantee standard compliance and data consistency. Especially, multi-model databases make it easy to create additional conformance requirements that data imports need to pass.

The spatial support for the tested DBMS (and many NoSQL databases in general) still falls short compared to object-relational databases. At the moment they might only be used for storage and simple thematic and spatial queries rather than advanced location intelligence (LI) tasks. But with a worldwide growing GIS market and GIS community there is enough potential to generate demands to extend these software products. An easy-to-use and easy-to-deploy data mapping and management solution like the developed example also helps to lower the boundaries for working with OGC standard data. The work presented some query examples that were much simpler than their relational counterparts.

Future work might focus on improving the mapping rules for multi-model databases. Exports from the graph model to OGC XML documents is not covered. At the moment the JavaScript code is not designed to consume massive data sets, as this research is a proof of concept.

7 References

- BAAS, B., 2012: NoSQL spatial: Neo4j versus PosGIS. Master thesis, Delft University of Technology.
- BOYD, R., 2015: Neo4j Blog - Polyglot Persistence Case Study, 2015: Wanderu + Neo4j + MongoDB - Neo4j Graph Database. Neo4j Graph Database. Available at: <http://neo4j.com/blog/polyglot-persistence-mongodb-wanderu-case-study/> [Accessed 28 Apr. 2016].
- DELFOSSÉ, V., BILLEN, R. & LECLERCQ, P., 2012: UML as a schema candidate for graph databases. NoSQL Matters.
- FALKOWSKI, K. & EBERT, J., 2009: Graph-based urban object model processing. City Models, Roads and Traffic (CMRT'09): Object Extraction for 3D City Models, Road Databases and Traffic Monitoring-Concepts, Algorithms and Evaluation, Paris, France, 9.
- HOUBIE, F., 2013: GraphGIS, bringing spatial functionalities in NoSQL graph databases. FOSS4G – OSGeo's Global Conference for Open Source Geospatial Software, Nottingham, UK, 17-21 September.
- KOCH, S., 2015: Storage and Querying of CityGML Models in BaseX - Evaluation of the Usage of a Native XML Database System for 3D City Models. Master Thesis, Technische Universität Berlin.
- 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.
- RODRÍGUEZ, A., FERNÁNDEZ-MEDINA, E., TRUJILLO, J. & PIATTINI, M., 2011: Secure business process model specification through a UML 2.0 activity diagram profile. Decision Support Systems **51** (3), 446-465.
- 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.

Prozedurale Modellierung von historischen 3D-Stadtmodellen und deren Aufbereitung für 3D-Visualisierungen in einer Smartphone-Applikation am Beispiel der römischen Stadt Augusta Raurica

MARTIN CHRISTEN¹, STEFAN BLASER¹, BENJAMIN LOESCH² & STEPHAN NEBIKER¹

Zusammenfassung: Ein texturiertes 3D-Stadtmodell auf einem Smartphone zu visualisieren, ist immer noch eine grosse Herausforderung, da die Grafikleistung auf diesen Geräten zwar immer leistungsstärker wird, jedoch nicht zu vergleichen ist mit herkömmlichen Desktops oder Laptops. In diesem Beitrag wird ein mögliches Verfahren am Beispiel des historischen 3D-Stadtmodells von Augusta Raurica gezeigt.

1 Einleitung

Augusta Raurica in der Nähe der heutigen Stadt Basel war eine der grössten römischen Städte nördlich der Alpen. Mit einer Augmented Reality App soll den zukünftigen Besucherinnen und Besuchern ein neuer Einblick in die ehemalige Römerstadt ermöglicht werden. An mehreren Stationen sollen dazu verschiedene interaktive Inhalte zum aktuellen Standort abgerufen werden können. Dazu gehören unter anderem detailliert rekonstruierte 3D-Modelle antiker Monumente als Augmented Reality oder gerenderte Panoramaansichten des ehemaligen Strassenraums als Virtual Reality. Die App sollte entweder vorgängig oder vor Ort auf das eigene Smartphone heruntergeladen werden können. Neben den Stationsinhalten sollte die App sowohl eine aktuelle, als auch eine historische 3D-Übersichtskarte zur Orientierung im Gelände bieten.

Dieser Beitrag zeigt wie das flächendeckende texturierte 3D-Stadtmodell von Augusta Raurica erstellt wurde. Die prozedurale Modellierung eignete sich am besten dafür. Als geometrische Grundlage dienten einerseits die aktuellen Grabungsdaten aus dem archäologischen GIS sowie andererseits existierende physikalische 3D-Modelle der Römerstadt, welche vorgängig photographisch erfasst wurden. In die regelbasierte 3D-Modellierung flossen zudem die aktuellsten archäologischen Erkenntnisse mit ein.

Das resultierende historische 3D-Stadtmodell umfasst verschiedene LOD-Stufen. So wurde eine Repräsentation mit einem geringeren Detaillierungsgrad für die Erstellung eines grossen Bronze-modells ausserhalb des Museums verwendet (SCHAUB 2014). Aus einer texturierten Repräsentation wurden mit serverseitigem Rendering 3D-Schrägansichten für die historische Übersichtskarte prozessiert. Die Augmented Reality- und Virtual Reality-Anwendungen beinhalten Teile der detailliertesten Repräsentation des digitalen 3D-Stadtmodells.

¹ Fachhochschule Nordwestschweiz Hochschule für Architektur Bau und Geomatik, Institut Vermessung und Geoinformation, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: [martin.christen, stefan.blaser, stephan.nebiker]@fhnw.ch

² iNovitas AG, Oberrohrdorferstrasse 1c, CH-5405 Baden-Dättwil, E-Mail: benjamin.loesch@inovitas.ch

Dieser Beitrag erläutert einerseits wie die mobile Applikation aufgebaut ist, andererseits wie das 3D-Stadtmodell prozedural modelliert und die 3D-Kartenansicht mit vorprozessierten Schrägan-sichten realisiert wurden. Die Augmented Reality App von Augusta Raurica dient als exemplarisches Beispiel dafür, wie für mobile Endgeräte effizient aufbereitete 3D-Daten zukünftig einem breiten Publikum zugänglich gemacht werden können.

2 Die Römerstadt Augusta Raurica

Augusta Raurica war eine der grössten römischen Städte nördlich der Alpen. Sie liegt am südlichen Rheinufer der heutigen Gemeinden Augst und Kaiseraugst, ca. 15 km östlich von Basel. Die Stadt wurde um 15 v. Chr. gegründet und entwickelte sich in 200 Jahren zu einer regionalen Metropole und einem blühenden Handelsplatz mit rund 15'000 Einwohnerinnen und Einwohnern. Die pulsierende Stadt bot alles, was zum damaligen städtischen Leben gehörte: u.a. ein grosses Theater für bis zu 30'000 Personen (Abb. 1), mehrere Tempel und Heiligtümer, ein eindruckliches Forum, ein Amphitheater sowie eine mächtige Zentraltherme.

Das heutige Augusta Raurica ist einerseits eines der beliebtesten Ausflugsziele in der Nordwestschweiz mit einem Römermuseum und mehr als zwanzig Monumenten und Fundstellen. Dazu gehören das besterhaltene römische Theater nördlich der Alpen, ein komplett eingerichtetes Römerhaus oder etwa eine gewaltige Kastellmauer. Andererseits ist es eine wichtige archäologische Forschungsstätte mit einem wissenschaftlichen Team von rund 60 Mitarbeitenden. Dieses Team hat unter anderem ein umfassendes Geoinformationssystem aufgebaut (BROMBACH 2010). Die archäologischen Daten der Römerstadt sind grösstenteils über das Geodatenportal des Kantons Basel-Landschaft frei zugänglich (<http://geoview.bl.ch/>).



Abb. 1: Das restaurierte Theater von Augusta Raurica (AUGUSTA RAURICA 2016)

3 Das neue 3D-Stadtmodell

3.1 Stadtmodellierungsansätze und Anwendungen in der Archäologie

Die explizite und die prozedurale Modellierung beschreiben zwei verschiedene Ansätze, welche im Rahmen der digitalen 3D-Stadtmodellierung in der Archäologie eingesetzt wurden. Um historische Monumente für die Kulturgütererhaltung zu dokumentieren, fand primär die explizite 3D-Modellierung mit 3D-CAD-Software Anwendung. Die manuelle Digitalisierung von 3D-Polygonen und die Anpassung der so erzeugten 3D-Modelle an den jeweils neusten archäo-

logischen Erkenntnisstand erwiesen sich als sehr zeitaufwändig. Durch die explizite Definition der 3D-Geometrie resultierten jedoch Modelle ohne geometrische Redundanzen und folglich mit geringen Speichervolumina. Mit dem 3D-GIS DILAS stand eine Verwaltungs- und Visualisierungslösung für explizit definierte 3D-Modelle zur Verfügung, welche unter anderem für das frühere 3D-Stadtmodell von Augusta Raurica eingesetzt wurde (WÜST, NEBIKER & LANDOLT 2004). Der prozedurale Ansatz der 3D-Stadtmodellierung wird beispielsweise in PARISH & MÜLLER (2001) abgehandelt. Die urbane Umgebung, bestehend aus Strassennetzwerk und Grundstücken, wird automatisch generiert. Mit einer eigens dafür entwickelten und Computer Generated Architecture (CGA) genannten Programmiersprache lässt sich der Siedlungsraum, inklusive Gebäude, regelbasiert modellieren. Die prozedurale Modellierung wurde schon mehrfach im Zusammenhang der Archäologie und Kulturgütererhaltung eingesetzt. MÜLLER et al. (2006) fokussierten sich mit der 3D-Rekonstruktion der Puuc-Gebäude in Xkipché auf die detaillierte Modellierung einzelner archäologischer Monumente für deren Dokumentation, während dem in SALDANA et al. (2013) die Modellierung ganzer Strassenzüge des antiken Roms als Grundlage für Computerspiele im Vordergrund standen. In HAEGLER et al. (2009) wurde das Prinzip der prozeduralen Modellierung, insbesondere der CGA-Regeldefinition, an verschiedenen Beispielen der Archäologie und der Kulturgütererhaltung diskutiert. DYLLA et al. (2009) zeigen mit dem Projekt Rome Reborn 2.0 auf, dass mithilfe prozeduraler Modellierung, ein homogenes 3D-Städtemodell von Rom erstellt werden konnte.

3.2 Prozedurale Modellierung von Augusta Raurica

Das erstellte 3D-Stadtmodell von Augusta Raurica wurde technisch stark an das 3D-Stadtmodell der ehemaligen römischen Stadt Pompeii (HAEGLER et al. 2009) angelehnt. Für beide Stadtmodelle wurde die Software ESRI CityEngine mit dem in PARISH & MÜLLER (2001) vorgestellten prozeduralen Modellierungsansatz für 3D-Städte eingesetzt. Beide Modelle wurden mit mehreren Detaillierungsgraden (LOD) realisiert und konnten Lücken im Modell mit Wahrscheinlichkeiten schliessen.

Das Modell von Pompeii war vom bekannten antiken Strassennetzwerk ausgehend aufgebaut worden, indem die Grundrisse gewöhnlicher Stadtbauten nach römischem Siedlungsmuster automatisch aus den Freiflächen generiert wurden. Im Gegensatz dazu, waren die Grundrisse beim Modell von Augusta Raurica bereits von früheren physikalischen Modellen gegeben und konnten als Polygone importiert werden. Die CGA-Regeln für die Gebäude wurden somit direkt den importierten Gebäudegrundrissen zugeordnet. Ein weiterer Unterschied zwischen den beiden Modellen besteht im Einbezug der Topographie. Währendem das Modell von Pompeii auf einer flachen Ebene liegt, wurde beim Modell von Augusta Raurica die Höhenlage der Geometrien an die DTM-Höhen angeglichen. Ein wichtiger Grund dafür, war die geplante spätere Nutzung für genaue ortsbezogene Augmented Reality Anwendungen.

Die Angleichung von Gebäudegrundrissen an DTM-Höhen erfolgte in der Annahme, dass es sich dabei stets um horizontale Ebenen handeln würde. Aufgrund dessen kam es innerhalb der für Augusta Raurica charakteristischen Säulendurchgänge, den sogenannten Portiken, zu unerwünschten Abstufungen (Abb. 2, links). Mit einem Python-Skript konnten die Abstufungen automatisiert beseitigt werden, indem die Höhen von Punkten mit identischer Lage aneinander angepasst wurden (Abb. 2, rechts).

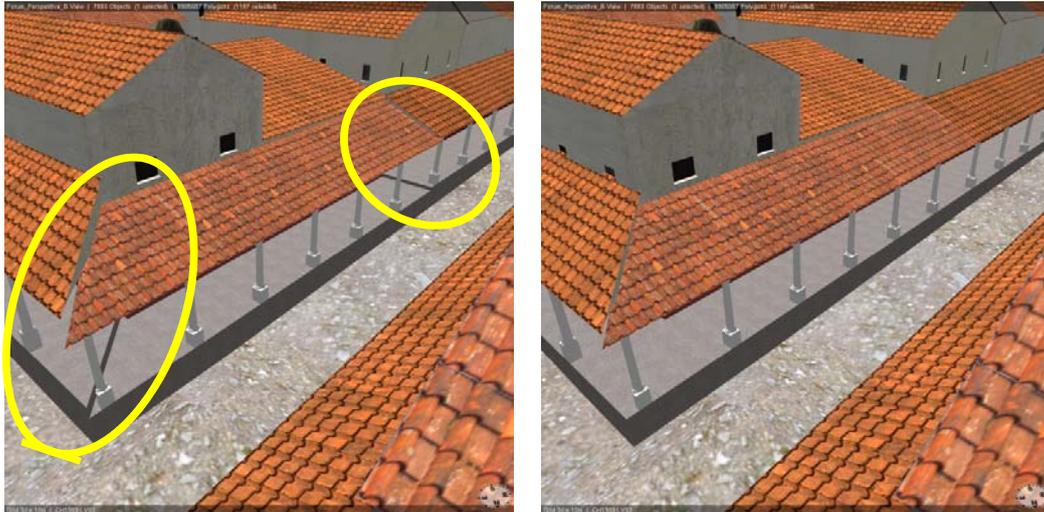


Abb. 2: Säulendurchgänge mit unerwünschten Abstufungen (links) und das Ergebnis nach der automatischen Beseitigung (rechts).

Grossflächige Geometrien, wie beispielsweise Plätze (Abb. 3, links), liessen sich kaum passend an das DTM angleichen, da die DTM-Höhen auch innerhalb der Flächen variierten. Eine Möglichkeit wäre die Unterteilung der Flächen gewesen. Je nach Objekttypus hätte eine weitere Unterteilung in Kombination mit den CGA-Regeln, welche den Teilflächen separat zugeordnet worden wären, zu unerwünschten Mustern geführt. Dies wurde beim Strassen- und Gewässernetzwerk umgangen, indem die betreffenden texturierten Geometrien mit GDAL direkt in die DTM-Textur eingebrannt wurden (Abb. 3, rechts).



Abb. 3: Vergleich Strassennetzwerk mit grossem Platz als 2D-Geometrie (links) und als Objekte welche direkt in die Geländetextur gebrannt wurden (rechts).

3.3 Detaillierungsgrade

Das prozedural modellierte 3D-Stadtmodell von Augusta Raurica umfasst drei verschiedene Detaillierungsgrade. Der niedrigste Detaillierungsgrad LOD1 (Abb. 4, links) repräsentiert

Augusta Raurica als eingefärbtes Strukturmodell. Dabei wurden die aus den GIS-Daten importierten Gebäudegrundrisse extrudiert und zusätzlich mit Dachgeometrien versehen. Sowohl die Gebäudehöhe als auch die Ausprägung des Dachs (Dachform, Dachwinkel und Dachausrichtung) wurden parametrisch definiert, welche sich bei Bedarf dynamisch anpassen liessen. Die LOD1-Repräsentation von Augusta Raurica wurde unter anderem für die Erstellung des Bronze-modells (SCHAUB 2014) verwendet. Der mittlere Detaillierungsgrad LOD2 umfasst eine texturierte Repräsentation des 3D-Modells (Abb. 4, Mitte). Die Dachflächen wurden in Quadrate mit einer Seitenlänge von 1.80 m, was ungefähr der Seitenlänge von vier römischen Ziegeln entspricht, unterteilt. Nach zufälligem Muster wurden die Quadrate mit Texturen aus einer, aus 14 Dachtexturen bestehender Bibliothek versehen. Dadurch konnte eine realitätsnahe und uniforme römische Dachlandschaft ohne störende repetitive Muster erstellt werden. Die Gebäudefassaden wurden zusätzlich zur LOD1-Repräsentation parametrisch ausmodelliert. Die Parameter von Fenster und Türen nahmen standardmässig zufällige Werte innerhalb eines definierten Wertebereichs an und liessen sich bei Bedarf pro Etage und Fassade separat anpassen. Die LOD2-Repräsentation diente zur Erstellung der 3D-Schrägansichten von Augusta Raurica. Der höchste Detaillierungsgrad LOD3 enthält, im Gegensatz zum LOD2, weitere ausmodellierte Details, wie einzelne Ziegelsteine oder ausgearbeitete Fenster mit Fensterglas, welches physikalisch korrekt auf dessen Beleuchtung reagiert (Abb. 4, rechts). Der LOD3 eignete sich für die Visualisierung und Betrachtung einzelner Bauten und Monumente aus der Nähe.



Abb. 4: Vergleich der drei Detaillierungsgrade LOD1 (links), LOD2 (Mitte), LOD3 (rechts) am Beispiel eines römischen Stadtquartiers.

4 Das 3D-Stadtmodell als Basis für interaktive Anwendungen

Das flächendeckende Stadtmodell von Augusta Raurica bildet eine 3D-Geodatenbasis der historischen Römerstadt. Die Daten lassen sich erst mit geeigneten Anwendungen und Produkten für eine breite Öffentlichkeit nutzbar machen. Denkbar dafür wären Webanwendungen oder Applikationen für mobile Endgeräte, womit das Modell, dank geeigneter Navigations- und Visualisierungskonzepte, auf einfache und zugleich attraktive Art und Weise betrachtet werden könnte. Die grossen Herausforderungen solcher Anwendungen liegen in den zurzeit begrenzten Rechen- und Speicherkapazitäten von mobilen Endgeräten. Schon die Darstellung eines 3D-Stadtmodells in der Grösse von Augusta Raurica erweist sich aufgrund der genannten Limitierungen als schwierig. Um das erzeugte 3D-Stadtmodell in mobilen Anwendungen nutzen zu können, wur-

den daraus verschiedene Produkte bzw. 3D-Anwendungen abgeleitet, die in der Folge vorgestellt werden.

4.1 Panoramaansicht im Strassenraum

Eine realisierte Anwendung zur interaktiven Betrachtung des 3D-Modells besteht aus 360°-Panoramen von ausgewählten Standorten. Für die spätere Darstellung sind vorgerenderte Ansichten des Modells notwendig, welche aus Bilddaten bestehen. Für die Visualisierung von 360°-Panoramen werden in der Computergraphik verbreitet Cube Maps (GREENE 1986) eingesetzt. Die 360°-Ansicht wird dabei als Würfel angenähert (Abb. 5, links), wobei sechs entsprechend ausgerichtete perspektivische Renderings auf die Würfelflächenseiten zu liegen kommen. Der Würfel kann beliebig um die im Würfelzentrum liegende virtuelle Kamera gedreht werden. Aufgrund des perspektivischen Renderings wären bei der Cube Map ungewünschte Verzerrungen entstanden. Durch die Verwendung eines Zylinders anstelle eines Würfels liessen sich die Verzerrungen in der horizontalen Richtung vermeiden. Dabei kam ein 360°-Panoramabild auf die Innenfläche des Zylindermantels zu liegen. Da sowohl für die Unter- als auch die Oberseite keine Renderings erstellt wurden, liess sich der Zylinder bei der Betrachtung lediglich um die Zylinderachse drehen. Beim Rendering des 360°-Panoramabilds wurde die Zylinderprojektion mit 36 perspektivisch gerenderten Streifen angenähert (Abb. 5, rechts). Die 36 streifenförmigen Renderings wurden in CityEngine mithilfe eines Python-Skripts automatisiert erstellt und anschliessend mit der Python Image Library (PIL) zu einem Panoramabild zusammengefügt.

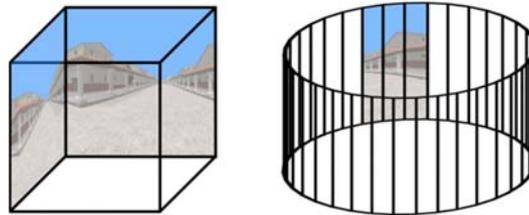


Abb. 5: Schematische Darstellungen des Panoramarenderings für eine Cube Map (links) und für einen Zylinder (rechts).

Das erstellte Panoramabild wurde mittels Bildbearbeitung zusätzlich künstlerisch aufgewertet, indem Römerinnen und Römer, sowie diverse kleinere Objekte in die Szene eingefügt worden waren (Abb. 6). Dieser Schritt könnte künftig automatisiert werden, indem die zusätzlichen Elemente als 3D-Objekte direkt in das Modell eingefügt würden.



Abb. 6: Mit CityEngine gerendertes, künstlerisch aufgewertetes Panoramabild einer Strassenszene in Augusta Raurica.

4.2 Aufbereitung detaillierter 3D-Architekturmodelle

Die Darstellung von detaillierten 3D-Architekturmodellen mit markerbasiertem Augmented Reality wurde als weitere Anwendung realisiert. Beim markerbasierten Augmented Reality werden Features im Live-Kamerabild des mobilen Endgeräts detektiert und mit Features eines bereits bestehenden Markers verglichen. Mit den übereinstimmenden Features können die äusseren Orientierungsparameter der Kamera in Echtzeit errechnet werden. Dank der laufend bekannten äusseren Orientierung, können 3D-Inhalte, welche sich im Markerkoordinatensystem befinden, lagekorrekt ins Kamerabild eingeblendet werden (LOESCH et al. 2015). Die zu visualisierenden detaillierten 3D-Architekturmodelle wurden aus dem mit CityEngine erstellten prozeduralen Stadtmodell von Augusta Raurica exportiert. Die exportierten 3D-Modelle wiesen redundante Geometrien, wie doppelt vorhandene Punkte und unnötige Unterteilungen innerhalb ebener Flächen auf, welche auf die prozedurale Modellierung zurückzuführen waren. Da die Redundanzen einerseits unnötigen Speicherbedarf und andererseits höheren Rechenaufwand für die Darstellung verursacht hätten, wurden die exportierten Modelle skriptbasiert mit der 3D-Graphiksoftware Blender optimiert. Mit dem Löschen doppelter Stützpunkte, sowie innenliegender Flächen, konnte der Speicherbedarf für die Geometrien um über die Hälfte verringert werden.

4.3 Cloud-basiertes Rendering der 3D-Schrägansichten

Für die 3D-Visualisierung grosser 3D-Geodatasets in Smartphone Apps kann aus Gründen von Performance und App-Grösse ein bildbasierter Ansatz verwendet werden, welcher mit dem 3dmaps.ch Framework realisiert wurde (CHRISTEN 2016). Es werden verschiedene Ansichten eines 3D-Stadtmodells inklusive Orthophoto und Geländemodell als Cloud-Dienst gerechnet, gecached und zum Client gestreamt. Das Rendering unterstützt verschiedene Zoom-Level. Die berechneten Bilder folgen dem Prinzip des G-Buffers (SAITO 1990). Der erste Bildtyp ist eine Farbkarte, welche die Farben – ohne Beleuchtung – der 3D-Ansicht enthält. Der zweite Bildtyp ist eine Normalmap, welche die Normalen in RGB kodiert enthält. Mit diesen zwei Bildtypen ist es beispielsweise möglich eine dynamische Beleuchtung zu rechnen. Ein dritter Bildtyp enthält die ID der Objekte, kodiert als Farbe. Ein vierter Bildtyp speichert die Tiefeninformation als Bild ab. Die 32bit Gleitkommazahl wird in RGBA kodiert, und kann auch dekodiert werden (CHRISTEN 2016). Mit der Tiefenkarte kann die 3D-Position jedes Pixels auf dem Bildschirm ermittelt werden.

Die 3dmaps.ch Web-Applikation besteht aus verschiedenen Python-Skripts, welche von einem Webinterface gestartet werden können. Das Webinterface wurde mittels Flask (FLASK 2015; LERNER 2014) entwickelt und benutzt MongoDB, eine Open Source Dokument-Datenbank, welche insbesondere für die Skalierung entwickelt wurde (MONGODB 2015). Das Web-Interface ist optional, die Python-Skripts können auch direkt aus der Konsole gestartet werden. Die gesamte Server-Architektur folgt der “View Extension” des OGC 3D Portrayal Service (OPEN GEOSPATIAL CONSORTIUM 2015). Zusätzlich wurde ein Caching hinzugefügt, damit dieselbe Ansicht nur einmal gerechnet wird.

Zunächst wird das Skript „createScene.py“ gestartet, um eine neue 3D-Szene zu definieren. Eine 3D-Szene besteht normalerweise aus 3D-Geometrien, Texturen und Materialien. Eine 3D-Szene kann auch Punktwolken enthalten. Sobald die Szene definiert ist, können die Daten hinzugefügt werden. Dies geschieht mit dem Skript „addFilesToScene.py“. Orthophotos und Geländemodelle

können zusammen auch zu einem 3D-Modell konvertiert werden. Ist die Szene komplett definiert, so kann dieses optimiert werden, beispielsweise durch Entfernung doppelter Flächen, Punkte und Materialien. Eine weitere Optimierung ist die Erstellung einer räumlichen Datenstruktur, um die Renderingzeit signifikant zu verkürzen. Eine der populärsten räumlichen Struktur für diesen Zweck ist die BVH - bounding volume hierarchy (BARRINGER & AKENINE-MÖLLER 2014). Die BVH wird mit allen 3D-Objekten, inklusive Punktwolken und Terrain (NEBIKER et al. 2010), berechnet. Zudem werden Texturen in ein Format konvertiert, welches Mehrfachauflösung und direkten Pixelzugriff unterstützt. Sobald die Szene fertig vorbereitet ist, können Render-Ansichten konfiguriert werden. Die Definition umfasst die Bildtypen (Farbkarte, Normalenkarte, Tiefenkarte, Id-Karte), die Kameraposition und die Blickrichtung. Die Renderingdefinition geschieht mit dem Python-Script „createRendering.py“. Beispiele für verschiedene Blickrichtungen sind in der Abb. 4 zu sehen.

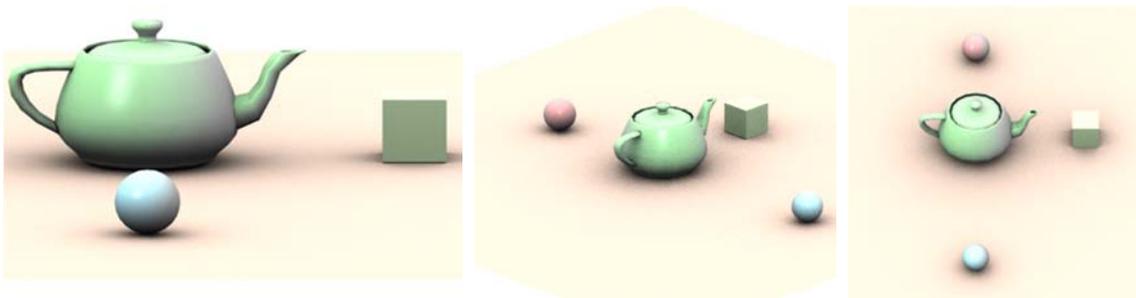


Abb. 7: Verschiedene Blickwinkel auf dieselbe Szene.

Die Ansichten werden als Kacheln in einer Quadtree-Struktur gerechnet, wie in Abb. 7 gezeigt. Die Objekte werden dabei in Parallelprojektion gerendert. Der Vorteil dabei ist, dass ein gleichmässiger Level of Detail mit einem konstanten Massstab über den ganzen Bildschirm dargestellt wird. Ein weiterer Vorteil der Quadtree-Struktur ist, dass nur die unterste Zoom-Stufe gerechnet werden müssen und die restlichen Kacheln über Bildverarbeitung erstellt werden. Somit kann das volle Potential der Beschleunigungs-Datenstruktur (BVH) ausgenutzt werden.

Das Rendering der untersten Stufe geschieht über Ray-Tracing. Dies ermöglicht die Verwendung von nahezu beliebig viel Geometrie und Texturen, da bei der Bildgenerierung jede Operation pixelweise ausgeführt wird. Für das Ray Tracing kann Renderman (RENDERMAN 2015) oder POV-Ray verwendet werden (POVRAY 2015). Die Kacheln, welche gerendert werden, sind als Render-Jobs in der MongoDB gespeichert und können in der Cloud von einem beliebigen Node gerechnet werden. Es können bei Bedarf jederzeit neue Rendering-Nodes hochgefahren werden. Die resultierenden Bilder werden wiederum direkt in der Cloud gespeichert. Das 3D-Modell von Augusta Raurica wurde auf 6 Zoomstufen vorgerechnet und ist in Abb. 8 zu sehen.



Abb. 8: 3D-Ansicht von Augusta Raurica mittels vorgerechneter Schrägansichten

5 Smartphone-Applikation

In Zusammenarbeit mit dem Archäologen-Team von Augusta Raurica wurde die App „unsichtbares sichtbar machen“ für Android und iOS entwickelt. Die App soll beim Besuch des historischen Geländes zusätzliche Informationen liefern. Da heute grosse Teile der antiken Römerstadt verschwunden sind, soll auch mittels Augmented Reality ein Eindruck der einstigen Situation vermittelt werden. Zusätzlich soll die historische 3D-Ansicht angezeigt werden.

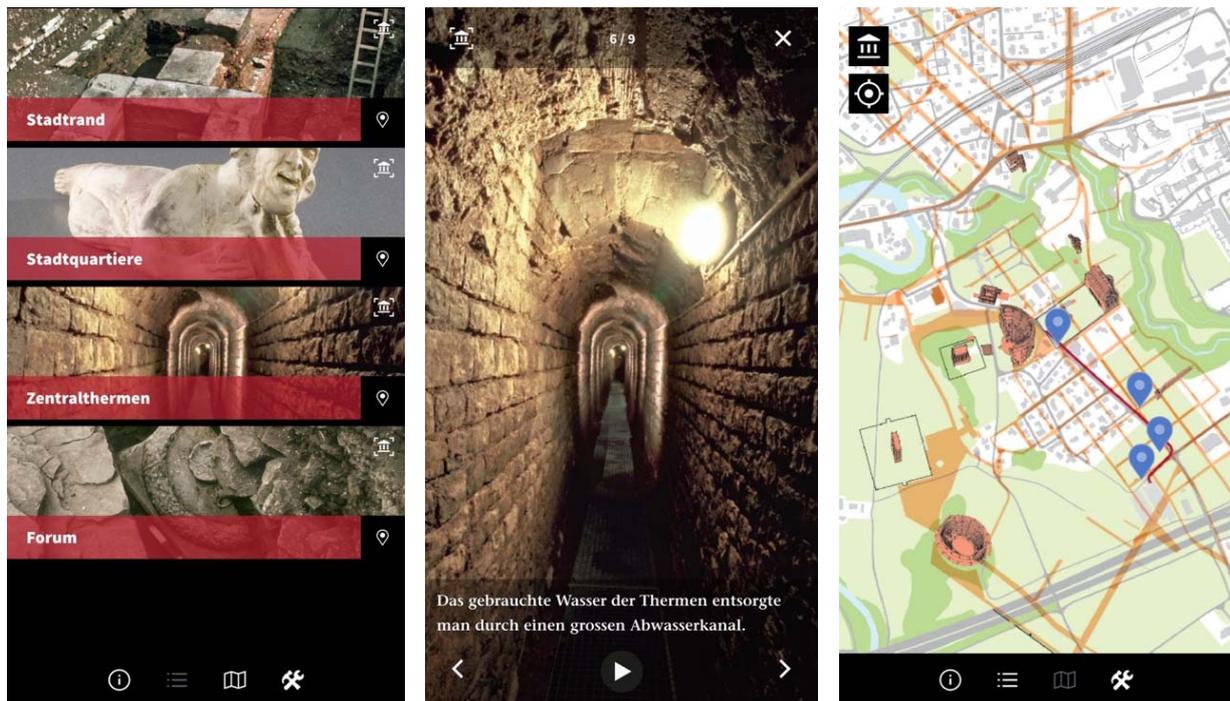


Abb. 9: Das Hauptmenu der App (links) und Beispielinhalt mit Audio-Guide (Mitte) und Kartenansicht mit aktueller Position und den 4 Posten (rechts).

Die App hat 4 verschiedene Posten, welche zusätzliche Informationen in Form von Bildern und Text liefern. Die Informationen können auch vorgelesen werden, somit dient die App auch als

Audio-Guide. Die Posten werden über das Hauptmenu ausgewählt. Die aktuelle Position sowie die Postenpositionen werden auf einer 2D Karte dargestellt (Abb. 9).

Die App soll auch von Schulklassen bedient werden, daher ergab sich die Einschränkung, dass für die Bedienung der App keine Internet-Verbindung erforderlich sein darf. Für den Download der App steht am Museumseingang WLAN zur Verfügung. Apps haben für Android eine Maximalgrösse von 100 MB (GLICK 2015). Das texturierte 3D-Modell besteht aus ca. 1 GB an Geometrie und Texturdaten und kann nicht direkt mit der App mitgeliefert werden oder während der Benutzung heruntergeladen werden. Daher wurde der Ansatz mit vorgerechneten Schrägansichten gewählt (siehe Kapitel 4.3). Dabei kann das historische 3D-Modell mit nur ca. 10 MB komplett dargestellt werden und erfordert praktisch keine Rechenressourcen des Smartphones. Die heutige Position wird mit den Ortungsdiensten des Smartphones zusätzlich angezeigt (Abb. 10).

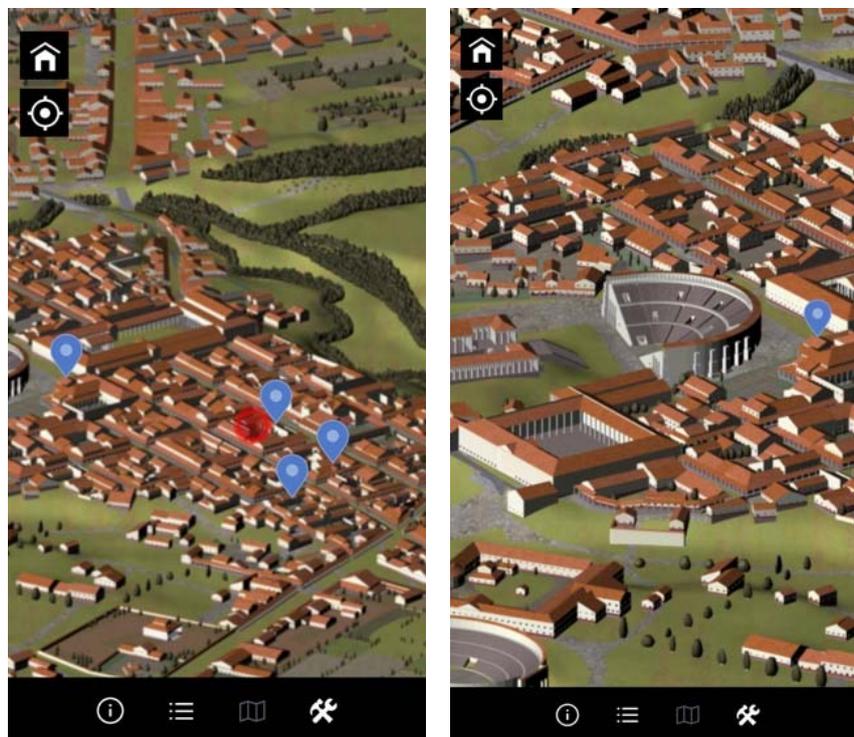


Abb. 10: 2D Karte mit überlagerten historischen Daten (links) und 3D Karte erstellt mit der vorgerechneten Schrägansicht.

Eine weitere Eigenschaft der App ist, dass an jedem Posten mittels Augmented Reality ein Blick in die Vergangenheit ermöglicht wird. Dabei wird ein rekonstruiertes, texturiertes 3D Modell überlagert. Als natürliche Marker dienten Archäologische Rekonstruktionszeichnungen der Monumente, welche auf Hinweistafeln im Gelände angebracht wurden (Abb. 11). Die Augmented Reality-Anwendung wurde um eine zusätzliche Navigationsfunktionalität erweitert, mit welcher das zu überlagernde 3D-Architekturmodell mit Wischgesten innerhalb eines zulässigen Bereichs im Raum gedreht werden kann.

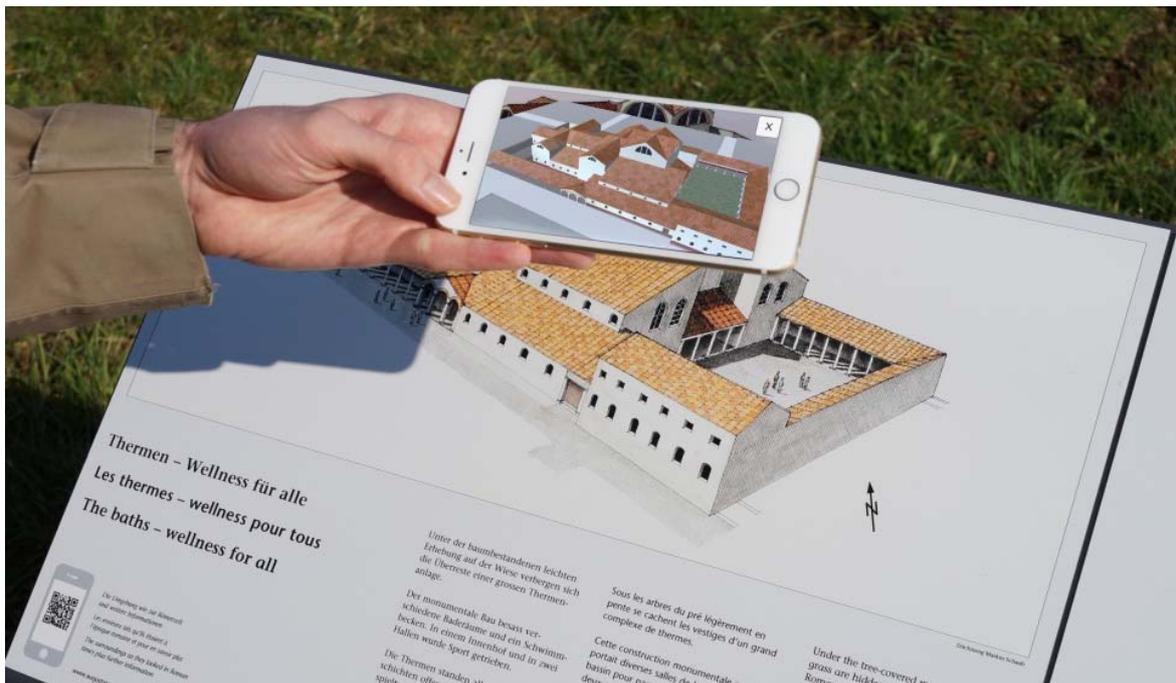


Abb. 11: Detailliertes 3D-Architekturmodell der Zentralthermen als Augmented Reality-Anwendung. Die Rekonstruktionszeichnung auf der Hinweistafel im Gelände dient als Marker.

6 Diskussion und Ausblick

Im vorliegenden Beitrag wurde die prozedurale Generierung eines texturierten historischen 3D-Stadtmodells am Beispiel der Römerstadt Augusta Raurica aufgezeigt. Das Stadtmodell umfasst drei unterschiedliche Detaillierungsstufen (LOD), zwischen welchen mittels Parameteranpassung gewechselt werden konnte. Damit lässt sich das 3D-Stadtmodell sowohl für Anwendungen mit geringem Detaillierungsgrad, wie im Beispiel des realisierten grossen Bronzomodells, oder aber auch für Augmented Reality Anwendungen mit teilweise sehr hohen Anforderungen an die Realitätstreue verwenden. Am Beispiel dreier Repräsentationsformen wurde aufgezeigt, wie ein sehr grosses 3D-Stadtmodell in einem mobilen Kontext und für ein breites Publikum nutzbar gemacht werden kann. Vorgerenderte interaktive 360°-Panoramen von ausgewählten Standorten wurden erzeugt, um Einblicke in den altrömischen Strassenraum zu ermöglichen. Im Rahmen einer zweiten Anwendung wurde der Speicherbedarf von detaillierten prozeduralen 3D-Architekturmodellen verlustfrei halbiert, sodass die Modelle auf mobilen Endgeräten visualisiert werden können. Schliesslich wurde aus dem Stadtmodell mit 3dmaps.ch ein serverseitiger Bilddienst generiert, der performante Betrachtung aus unterschiedlichen Perspektiven ermöglicht. Die untersuchten Augmented Reality Konzepte wurden erfolgreich in die neue Besucher-App "Unsichtbares sichtbar machen" (<http://www.augustaurica.ch/app>) der Römerstadt Augusta Raurica integriert.

Die vorgestellten Konzepte und Technologien sollen zukünftig auf andere 3D-Stadtmodelle übertragen werden, um AR-Anwendungen etwa im Tourismus oder im Schulunterricht zu realisieren. Die vorgestellten bildbasierten Rendering-Ansätze und die entsprechenden neuen OGC

Standards bieten zudem ein grosses Potential für inhaltliche, funktionale und semantische Erweiterungen. Dazu gehören beispielsweise die zukünftige Integration von 3D-Punktwolken oder etwas die Nutzung der Tiefeninformation oder zusätzlicher thematischer Layer.

7 Dank

Die Autoren danken der Römerstadt Augusta Raurica und ihren kompetenten wissenschaftlichen Mitarbeitenden für die langjährige Projektzusammenarbeit und für das Vertrauen, insbesondere im Zusammenhang mit dem kürzlich erfolgreich realisierten Projekt "Unsichtbares sichtbar machen". Besonderen Dank sei Markus Schaub, Leiter der Publikumsgrabungen und wissenschaftlicher Zeichner der Römerstadt Augusta Raurica, ausgesprochen. Ohne die Einbringung seines fundierten archäologischen Fachwissens und seiner zeichnerisch-gestalterischen Erfahrung und der sehr engagierten, unkomplizierten Zusammenarbeit wäre die Realisierung des 3D-Modells in dieser Ausprägung undenkbar gewesen.

8 Literaturverzeichnis

- AUGUSTA RAURICA, 2016: Webseite von Augusta Raurica, <http://www.augustaurica.ch>.
- BARRINGER, R. & AKENINE-MÖLLER, T., 2014: Dynamic Ray Stream Traversal. *ACM Transactions on Graphics* **33** (4), Article 151.
- BROMBACH, U., 2010: Fünf Jahre Geografisches Informationssystem in Augusta Raurica. *Jahresberichte aus Augst und Kaiseraugst* **31**, 193-205.
- CHRISTEN, M., 2016: Visualisierung komplexer 3D-Geodaten im (mobilen) Webbrowser mittels Cloud-prozessierten Schrägansichten. *Geoinformationssysteme 2016 – Beiträge zur 3. Münchner GI-Runde 2016*, Kolbe, T.H., Bill, R. & Donaubaue, A. (Hrsg.), Wichmann, 68-76.
- DYLLA, K., FRISCHER, B., MÜLLER, P., ULMER, A. & HAEGLER, S., 2009: Rome Reborn 2.0: A Case Study of Virtual City Reconstruction Using Procedural Modeling Techniques. *Proceedings of the CAA Conference 2009*, Williamsburg, Virginia, USA, 62-66.
- FLASK, 2015: Flask Documentation / User's Guide, <http://flask.pocoo.org/docs/0.10/> (letzter Zugriff: 23.4.2015).
- GLICK K., 2015: Support for 100MB APKs on Google Play. <http://android-developers.blogspot.ch/2015/09/support-for-100mb-apks-on-google-play.html> (letzter Zugriff: 12.4.2016).
- GREENE N., 1986: Environment Mapping and Other Applications of World Projections. *IEEE Computer Graphics and Applications* **6** (11), 21-29.
- HAEGLER, S., MÜLLER, P. & VAN GOOL, L., 2009: Procedural Modeling for Digital Cultural Heritage. *EURASIP Journal on Image and Video Processing*, Volume 2009, 7.
- LOESCH, B., CHRISTEN, M., WÜEST, R. & NEBIKER S., 2015: Geospatial Augmented Reality – Lösungsansätze mit natürlichen Markern für die Kartographie und die Geoinformationsvisualisierung im Aussenraum. *Publikationen der DGPF, Band 24*, T. Kersten (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 16.-18. März 2015 an der Universität zu Köln, Tagungsband auf CD-ROM, 89-97.

- LERNER, M.R., 2014: At the forge: flask. Issue of Linux Journal, July 2014, Belltown Media, Houston, TX.
- MONGODB, 2015: The MongoDB 3.0 Manual, <http://docs.mongodb.org/manual/> (letzter Zugriff: 20.11.2015).
- MÜLLER, P., VEREENOOGHE, T., WONKA, P., PAAP, I. & VAN GOOL, L.J., 2006: Procedural 3D Reconstruction of Puuc Buildings in Xkipché. Proceedings of the 7th International Symposium on Virtual Reality, Archeology and Cultural Heritage VAST, 139-146.
- NEBIKER, S., BLEISCH, S. & CHRISTEN, M. 2010: Rich point clouds in virtual globes – A new paradigm in city modeling? Computers, Environment and Urban Systems **6** (34), 508-517.
- OPEN GEOSPATIAL CONSORTIUM, 2015: OGC seeks public comment on candidate 3D Portrayal Service Standard, <http://www.opengeospatial.org/pressroom/pressreleases/2165> (letzter Zugriff: 19.11.2015).
- PARISH, Y.I.H. & MÜLLER P., 2001: Procedural modeling of cities. Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, 301-308.
- POVRAY, 2015: Persistence of Vision Renderer „POV-Ray“. <http://povray.org> (letzter Zugriff: 20.11.2015).
- RENDERMAN, 2015: Pixar Renderman Pro Server. <http://renderman.pixar.com> (letzter Zugriff: 20.4.2015).
- SAITO, T. & TAKAHASHI, T., 1990: Comprehensible rendering of 3-D shapes. ACM SIGGRAPH Computer Graphics **24** (4), 197-206.
- SALDAÑA, M., & JOHANSON, C., 2013: Procedural Modeling for Rapid-Prototyping of Multiple Building Phases. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **40** (5/W1), 205-210.
- SCHAUB, M., 2014: Das Bronzemedell von Augusta Raurica: vom Stadtplan zum Stadtmodell. Augusta Raurica Magazin, 2014(1), 7-9.
- WÜST, T., NEBIKER, S. & LANDOLT, R., 2004: Applying the 3D DIS DILAS to Archaeology and Cultural Heritage Projects – Requirements and First Results. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (Part 5), 407-412.

Korrektur der Signaldämpfung und Analyse zur Bestimmbarkeit der Gewässertrübung in Laserbathymetrie-Daten

KATJA RICHTER¹, PATRICK WESTFELD¹, HANS-GERD MAAS¹ & ROBERT WEIß²

Zusammenfassung: Seit einigen Jahren werden Laserbathymetriesysteme erfolgreich zur Erfassung der Topografie von Gewässerbettflächen eingesetzt. Die Leistungsfähigkeit des Messverfahrens wird allerdings durch die Gewässertrübung eingeschränkt. Beim Durchgang des Laserpulses durch die Wassersäule wird die Intensität des emittierten und reflektierten Laserpulses durch Absorptionsprozesse im Wasser verringert. In Abhängigkeit von der Gewässertrübung und -tiefe wird die Signalintensität im unteren Bereich der Wassersäule geringer. Bei stärkerer Gewässertrübung oder in tieferen Bereichen kann das am Gewässerboden reflektierte Echo mit klassischen Auswertemethoden nicht mehr detektiert werden. Eine Kompensation des Einflusses der Signaldämpfung kann hier – je nach Detektionsverfahren – eine Möglichkeit zur Erhöhung des Anteils detektierter Bodenpunkte darstellen.

Darüber hinaus wird ein erster Ansatz vorgestellt, aus dem Verlauf der aufgezeichneten Signale pulsweise integrale Trübungsparameter für das Gewässer abzuleiten. Diese flächenhafte Bestimmung der Gewässertrübung kann eine sehr interessante Alternative zu konventionellen Verfahren darstellen, bei denen Trübungsparameter nur stichprobenhaft erfasst werden.

1 Einleitung

Zu den Aufgaben der Wasser- und Schifffahrtsverwaltung gehört neben der Verkehrssicherung auch ein aktives Wasserstraßenmanagement, das eine optimale Bewirtschaftung der Bundeswasserstraßen sicherstellt. Eine Grundlage hierfür bilden genaue Informationen über die Geometrie der Wasserstraßen, die in Form digitaler Geländemodelle mit Wasserlauf (DGM-W) bereitgestellt wird. Eine besondere Herausforderung stellen dabei die Flachwasserbereiche dar, die mit konventionellen Methoden nur sehr aufwendig zu erfassen sind. Mit Laserbathymetriesystemen kann die Geometrie der Wasserwechselzonen und anderer Flachwasserbereiche wesentlich wirtschaftlicher vermessen werden. Untersuchungen der Bundesanstalt für Gewässerkunde in Kooperation mit dem Wasser- und Schifffahrtsamt Dresden zeigten allerdings, dass Standardauswertemethoden zur Ableitung der Gewässergeometrie nur eingeschränkt geeignet sind, da das Gewässerbodenecho nur in wenigen Fällen detektiert wurde (WEIß 2015). Datengrundlage sind die im Rahmen eines Pilotprojekts mit den bathymetrischen ALS Systemen AHAB Chiroptera und RIEGL VQ820G durchgeführten Befliegungen der Bundeswasserstraße Elbe im Bereich Klöden (Abbildung 1). Im Projektgebiet befinden sich neben der Elbe einige Altarme, die ebenfalls für die Untersuchungen interessant sind.

¹ TU Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstraße 10, D-01069 Dresden; E-Mail: [katja.richter1, patrick.westfeld, hans-gerd.maas]@tu-dresden.de

² Bundesanstalt für Gewässerkunde, Am Mainzer Tor 1, D-56068 Koblenz

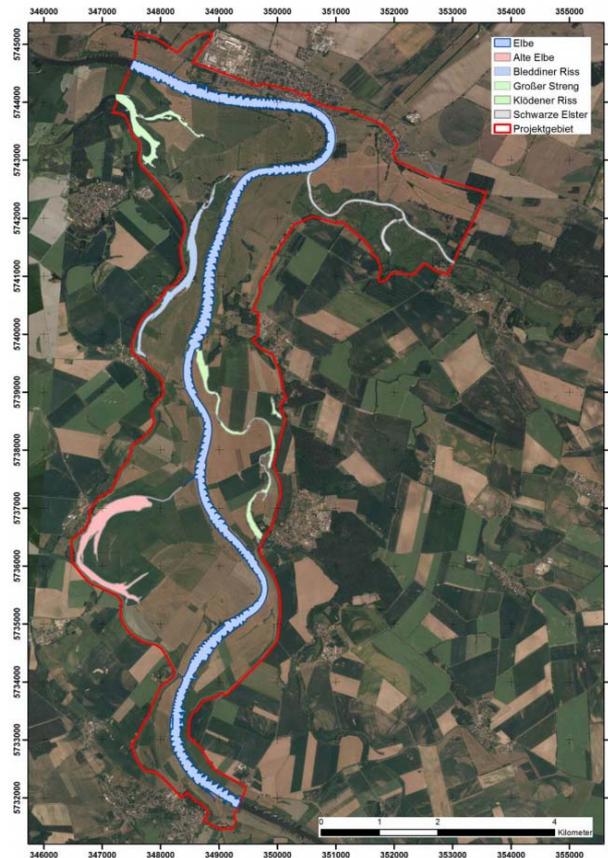


Abb. 1: Untersuchungsgebiet (Elbe km 184-203) zwischen Mauken und Elster/Wittenberg

Um das Potential der erfassten Bathymetriedaten voll auszuschöpfen, müssen neue Auswertemethoden entwickelt werden. Die verbesserte Interpretation der aufgezeichneten Waveform-Daten erfordert ein genaues Verständnis der Interaktion zwischen elektromagnetischer Strahlung und Wassersäule. Die Ausbreitung des emittierten Laserpulses innerhalb der Wassersäule ist durch Absorptions- und Streuprozesse an Wassermolekülen sowie an verschiedenen Partikeln und Schwebstoffen charakterisiert (GUENTHER et al. 2000). Während bei der Absorption ein Teil der Energie verloren geht, gelangt durch die Streuung ein Teil der Energie zurück zum Empfänger (Volume Backscatter). Bei den Streuprozessen wird in Einfach- und Mehrfachstreuung unterschieden (DONEUS et al. 2013). Die Mehrfachstreuung ist abhängig von der Gewässertrübung und beeinflusst vor allem die Amplitude des Signals, weniger ihren zeitlichen Verlauf (HARSDORF et al. 1999). Wird der Messprozess als ideal betrachtet, ergibt sich das gemessene Signal P_r aus der Faltung von emittiertem Signal P_e und differentiellem Rückstreuquerschnitt je Volumeneinheit σ (HARSDORF et al. 1999). Die Mehrfachstreuung wird in diesem mathematischen Modell vernachlässigt. Der differentielle Rückstreuquerschnitt charakterisiert die Eigenschaften des vom Laserpuls beleuchteten Mediums und kann mit geeigneten Entfaltungstechniken aus emittiertem und reflektiertem Signal rekonstruiert werden. Er stellt die Grundlage für die in Kapitel 2 beschriebene Modellierung der Signaldämpfung und anschließende Ableitung zusätzlicher Gewässerbodenpunkte dar.

Neben der Gewässergeometrie ist auch die Bestimmung von Trübungsparametern von Interesse. Die Gewässertrübung wird konventionell durch Schwebstoffmessungen und Sichttiefenmessungen regelmäßig an definierten Positionen erfasst. Die gemessenen Werte sind jedoch nicht für das gesamte Gewässer repräsentativ, da zum Teil sehr große lokale Trübungsvariationen auftreten. Außerdem hängen die Messergebnisse stark von subjektiven Einflussfaktoren auf den Messvorgang ab. Eine Ableitung der Trübungsparameter aus den Laserbathymetriedaten ermöglicht demgegenüber eine objektiviertere flächenhafte Analyse der Gewässertrübung. Die entwickelten Methoden und Ergebnisse werden in Kapitel 3 dargestellt.

2 Modellierung der Signaldämpfung

Ursache der Signaldämpfung ist die Abschwächung der Intensität des emittierten Laserpulses durch Absorptionsprozesse beim Durchgang des Laserpulses durch die Wassersäule. Dadurch ist die Signalintensität im unteren Bereich der Wassersäule kleiner. Diese Abschwächung kann vereinfacht mit der Gleichung für die Strahlungsdämpfung (Gleichung 1) beschrieben werden, die auf dem Lambert-Beer-Gesetz basiert.

$$I_1 = I_0 \cdot e^{-\varepsilon \cdot d} \quad \text{Gleichung 1}$$

Dabei ist I_0 die Intensität des einfallenden Laserpulses und I_1 die Pulsintensität nach Passieren der Wassersäule. Mit ε und d werden der Absorptionskoeffizient und die Länge der durchdrungenen Wassersäule beschrieben. Der Absorptionskoeffizient ist von der Wellenlänge λ des Lasers abhängig. Nach Bestimmung des Absorptionskoeffizienten aus dem Signalverlauf kann die Signaldämpfung korrigiert werden. Der Vergleich zwischen theoretischem Modell und realem Signal (Abbildung 2) zeigt jedoch, dass die Signaldämpfung in der Realität nicht kontinuierlich verläuft.

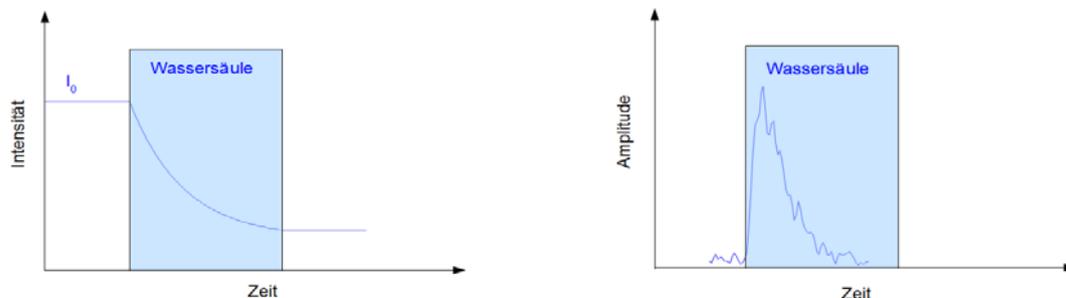


Abb. 2: Vergleich Modell Signaldämpfung (links) und reale Signaldämpfung (rechts)

Durch Analyse der jeweiligen Pulshistorie kann für jede Waveform eine individuelle Dämpfungskorrektur direkt aus dem differentiellen Rückstreuquerschnitt abgeleitet werden. Grundidee des Korrekturverfahrens ist das schrittweise Anheben der Waveform-Intensitätswerte in den unteren Bereichen der Wassersäule mit Hilfe geeigneter Korrekturfaktoren. Die Korrekturfaktoren c_i sind vom Anteil p_i der Pulsintensität abhängig, der im oberen Bereich der Wassersäule absorbiert wurde (Gleichung 2).

$$c_i = \frac{1}{(1-p_i)} \quad \text{Gleichung 2}$$

Um diesen Anteil p_i zu berechnen, werden als Referenzwert Informationen über die an der Wasseroberfläche eintreffende Energie benötigt. Diese Informationen lassen sich nicht aus der digitalisierten emittierten Waveform ableiten, da der Zusammenhang zwischen digitalisierter Amplitude und tatsächlich emittierter Energie unbekannt ist. Daher wurden zwei in vorangegangenen Studien entwickelte Methoden getestet, mit denen der Referenzwert mit statistischen Mitteln aus dem Datensatz selbst abgeleitet werden kann (RICHTER et al. 2015). Der Anteil p_i berechnet sich aus dem Verhältnis der Amplitude des differentiellen Rückstreuquerschnittes an der Abtaststelle A_i zum Referenzwert Ref (Gleichung 3).

$$p_i = \frac{A_i}{Ref} \quad \text{Gleichung 3}$$

Zur Beurteilung der Dämpfungskorrektur können die relativen Korrekturfaktoren c_i als Funktion der Zeit visualisiert werden. Für das auf dem Lambert-Beer-Gesetz basierende theoretische Modell hat die Korrekturfunktion einen streng exponentiellen Verlauf (Abb. 3, links). Die aus dem Rückstreuquerschnitt abgeleitete Korrekturfunktion (Abb. 3, rechts) hat grundsätzlich einen ähnlichen, jedoch an die individuelle Pulshistorie angepassten Verlauf.

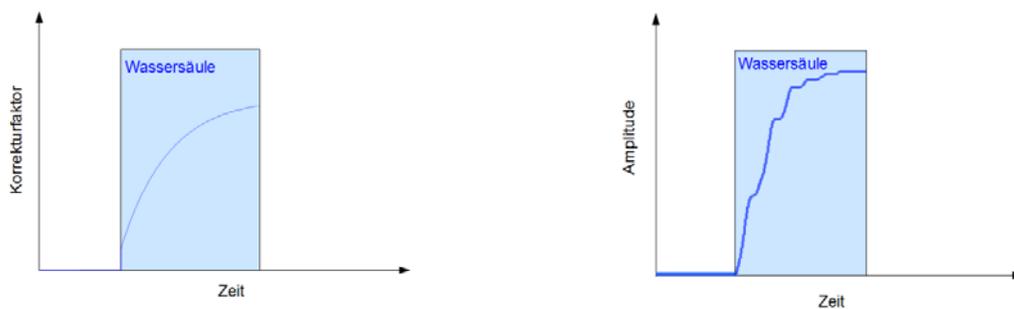


Abb. 3: Vergleich Korrekturfunktion Modell (links) und reale Korrekturfunktion (rechts)

2.1 Referenzwertbeschaffung

2.1.1 Automatische Methode

Die automatische Methode zur Ermittlung des Referenzwertes basiert auf der Annahme, dass das an der Wasseroberfläche eintreffende Signal dem maximalen rückgestreuten Signal entspricht. Der Referenzwert wird aus der Analyse der Integrale der differentiellen Rückstreuquerschnitte abgeleitet. Die Größe des Integrals hängt direkt mit der Gewässertrübung und damit der Signaldämpfung zusammen. In einem trüben Gewässer wird das Signal stark gedämpft und das Integral fällt klein aus. In einem klaren Gewässer fällt die Dämpfung geringer aus und das Integral hat einen entsprechend größeren Wert. Weitere Einflussfaktoren sind die ebenfalls von der Gewässertrübung abhängige Reflektivität des Wassers sowie das Vorhandensein von Gewässerbodenreflektionen im Signal. Abbildung 4 zeigt die Histogramme aller im Untersuchungsgebiet auftretenden Intergale. Um statistische Fehler zu reduzieren, wird der Referenzwert im Abstand von 2σ zum Mittelwert bestimmt. Es ergeben sich der Wert $Ref=59,0$.

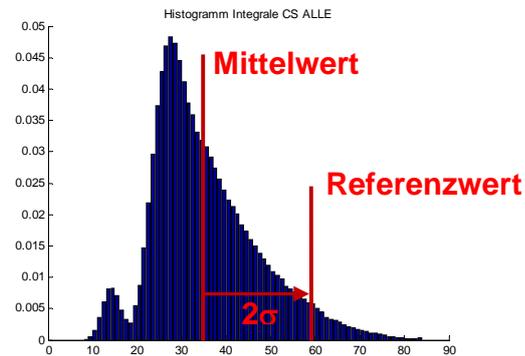


Abb. 4: Histogramm Integrale

Ein großer Vorteil der Methode ist die komplette Automatisierung. Allerdings führt die getroffene Annahme tendenziell zu einer Unterschätzung des Referenzwertes. Da die maximalen rückgestreuten Signale ebenfalls von Dämpfungseffekten beeinflusst sind, fällt der ermittelte Referenzwert tendenziell etwas kleiner aus als der wahre Wert. Außerdem können im Signal enthaltene Reflektionen am Gewässerboden das Verfahren beeinflussen. Reflektivitätsunterschiede zwischen den unterschiedlich trüben Gewässern werden vernachlässigt.

2.1.2 Interaktive Methode

Die Interaktive Methode stellt ein alternatives Verfahren dar, mit dem sich Referenzwerte unabhängig von gedämpften Signalen ableiten lassen. Sie basiert auf der Annahme, dass das an der Wasseroberfläche eintreffende Signal nach einer Reflektivitätsanpassung der Reflektion an einer Referenzfläche entspricht. Um den Referenzwert zu bestimmen, werden zunächst im gesamten Untersuchungsgebiet geeignete Referenzflächen mit bekannter Reflektivität gesucht. Im Datensatz sind mehrere urbane Bereiche mit zahlreichen Asphaltstraßen vorhanden, die sich als Referenzfläche anbieten. Abbildung 5 zeigt eine der insgesamt 17 verwendeten Referenzflächen.

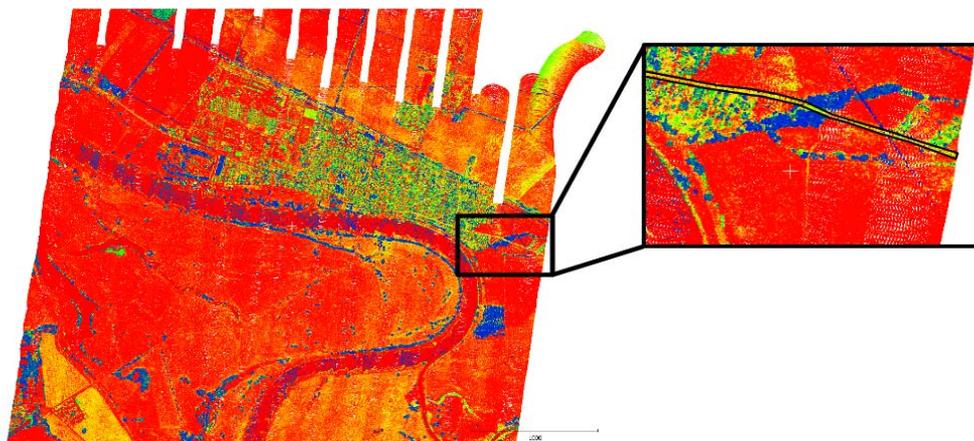


Abb. 5: Beispiel Referenzfläche

Die Analyse der Integrale der Rückstreuquerschnitte aller Waveforms der Referenzflächen ergibt einen Mittelwert $M=104,7$ für die Reflektion an Asphalt. Durch eine Reflektivitätsanpassung (Gleichung 4) werden aus diesen Mittelwerten die gesuchten Referenzwerte ermittelt. Die relative Reflektivität für Asphalt ($\rho_{Asphalt}$) und Wasser (ρ_{Wasser}) können der Literatur entnommen werden (Pfennigbauer, 2011 und Liew, 2001). Es ergibt sich der Referenzwert $Ref=69,8$.

$$Ref = MW \cdot \frac{\rho_{Wasser}}{\rho_{Asphalt}} \quad \text{Gleichung 4}$$

Die interaktive Methode zur Referenzwertbestimmung basiert wie die automatische Methode auf der Annahme, dass alle Gewässer im Untersuchungsgebiet die gleiche Reflektivität aufweisen und ρ_{Wasser} als gemeinsamer Wert angesetzt werden kann. Beide Methoden stützen sich gegenseitig, da sie unabhängig voneinander ähnlich Ergebnisse liefern. Die mit der automatischen Methode ermittelten Werte fallen aus den oben genannten Gründen allerdings tendenziell zu klein aus. Für die Dämpfungskorrektur werden daher die mit der interaktiven Methode bestimmten Referenzwerte verwendet. Der mit der interaktiven Methode verbundene Aufwand kann im Vergleich zum Aufwand konventioneller Methoden zur Bestimmung der Gewässertrübung an diskreten Punkten als gering bezeichnet werden.

2.2 Ergebnisse

Die Ergebnisse der entwickelten Dämpfungskorrekturmethode werden an Hand eines Beispiel-Signals präsentiert. In Abbildung 6 (oben) ist eine reflektierte Waveform im Untersuchungsgebiet dargestellt. Obwohl ein Peak im Bereich des Gewässerbodens erkennbar ist, wurde in der Standardauswertung durch den Systemhersteller für dieses Signal kein Bodenpunkt detektiert. Bereits im rekonstruierten differentiellen Rückstreuquerschnitt (Abbildung 6, unten, blau) ist der Bodenpeak besser zu erkennen und unterscheidet sich nach der Dämpfungskorrektur (Abbildung 6, unten, grün) deutlicher von anderen Peaks. Ein unerwünschter Nebeneffekt der Entfaltung ist die Verstärkung des Signalrauschens. Dadurch können beim rekonstruierten differentiellen Rückstreuquerschnitt auch negative Amplituden A_i auftreten, deren Verarbeitung natürlich nicht sinnvoll ist. Stattdessen wird $p_i=0$ gesetzt, so dass an diesen Abtaststellen keine Dämpfungskorrektur durchgeführt wird.

Im nächsten Schritt werden die Peaks im Rückstreuquerschnitt detektiert und Koordinaten für Wasseroberfläche und Gewässerboden abgeleitet. Dabei stellt die Unterscheidung zwischen echten Gewässerbodenreflektionen und durch die Entfaltung verstärktem Rauschen eine Herausforderung dar, die eine Weiterentwicklung der bestehenden Algorithmen erfordert. Das Potential der entwickelten Methode zeigt sich bereits bei der Betrachtung eines Gewässerprofils im Untersuchungsgebiet. In Abbildung 7 (oben) ist ein Ausschnitt der mit Standardmethoden aus dem reflektierten Signal abgeleiteten Punktwolke dargestellt. Die Farbcodierung entspricht der Amplitude der detektierten Peaks. Während die Wasseroberfläche mit Ausnahme der Randbereiche komplett erfasst ist, wird der Gewässerboden nur durch wenige Punkte im Flachwasserbereich repräsentiert. Die Ableitung der Gewässerbodenpunkte aus den rekonstruierten Rückstreuquerschnitten (Abbildung 7, unten) liefert deutlich mehr Bodenpunkte für die gesamte Gewässersohle. Für die geplante Validierung stehen Daten eines Fächerecholotes zur Verfügung.

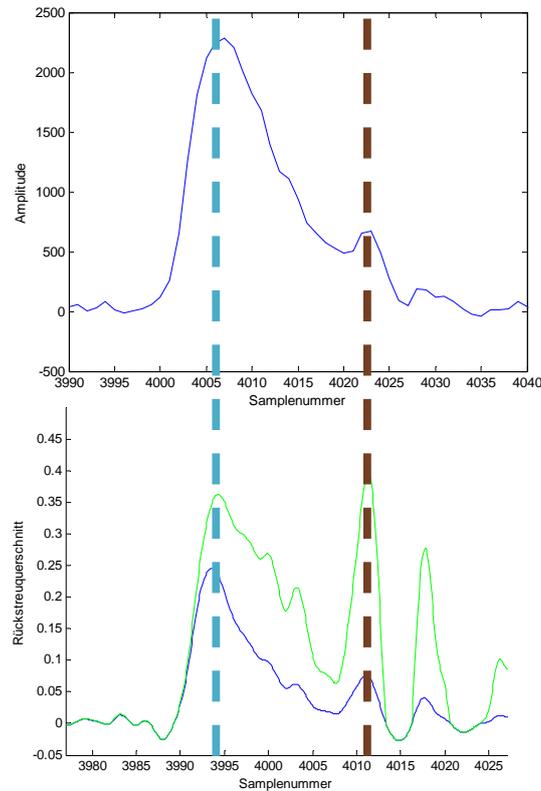


Abb. 6: Vergleich reflektiertes Signal (oben), differentieller Rückstreuquerschnitt (unten, blau) und dämpfungskorrigierter differentieller Rückstreuquerschnitt (unten, grün)

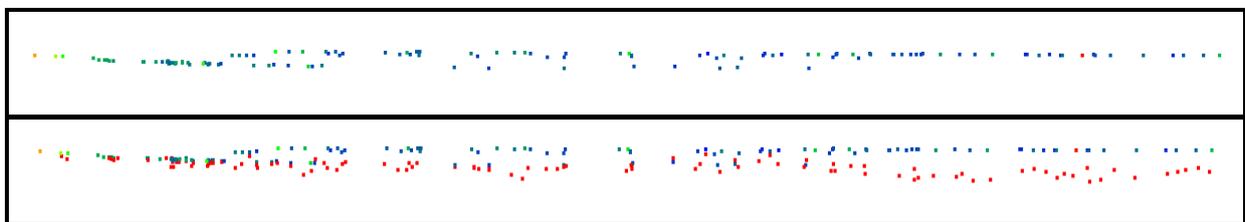


Abb. 7: Gewässerprofil, oben: Punktwolke Standardauswertung, unten: Punktwolke Standardauswertung überlagert mit zusätzlich aus diff. Rückstreuquerschnitt abgeleiteten Gewässerbodenpunkten (rot)

3 Ableitung Trübungsparameter

3.1 Methode

Das Lambert-Beer-Gesetz (Gleichung 1) kann auch eine Grundlage für die Ableitung von Trübungsparametern bilden, wobei der Absorptionskoeffizient ε ein (integrales) Maß für die Gewässertrübung darstellt. Um den Absorptionskoeffizienten zu bestimmen, wird eine Exponentialfunktion in das Signal eingepasst. Da der exponentiell abfallende Trend in der reflektierten Waveform deutlicher ausgeprägt ist als im rekonstruierten differentiellen Rückstreuquerschnitt, wird die Funktionsapproximation direkt am Ausgangssignal durchgeführt

und liefert Schätzwerte für die Parameter I_0 und ε . Abbildung 8 zeigt die Einpassung der Exponentialfunktion in das Signal der Wassersäule.

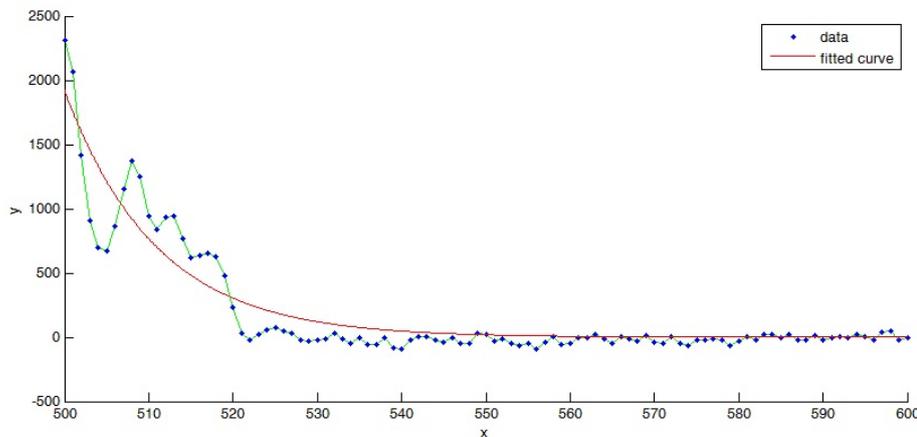


Abb. 8: Approximation Exponentialfunktion (rot) in Signal der Wassersäule (grün)

3.2 Ergebnisse

Die Gewässertrübung wurde für jede Waveform im Untersuchungsgebiet abgeleitet. Abbildung 9 zeigt eine Visualisierung der für den Altarm Bleddiner Riss ermittelten Absorptionsparameter in einer farbcodierten Darstellung. Die Koordinaten entsprechen dem an der Wasseroberfläche detektierten Punkt, die Farbcodierung dem ermittelten Absorptionsparameter, dabei entspricht blau einer geringen und rot einer starken Gewässertrübung. Der ermittelte Absorptionsparameter ist dabei zunächst ein rechnerischer Wert ohne konkrete physikalische Bedeutung und ist dementsprechend dimensionslos. Für die weitere Verwendung des Parameters für limnologische Analysen ist er anhand von Referenzmessungen zu skalieren (vgl. Kap. 3.3).

3.3 Validierung

Für die Validierung der entwickelten Methode stehen Referenzwerte in Form von lokalen Sichttiefenmessungen an unterschiedlichen Standorten im Bereich der Elbe und der Altarme zur Verfügung. Die Sichttiefenmessung mittels Secchischeibe liefert metrische Werte (SECCHI, 1866). Im Gegensatz dazu ergeben sich aus der Ableitung der Absorptionsparameter aus der bathymetrischen ALS-Befliegung dimensionslose Koeffizienten. Die Messergebnisse lassen sich somit nicht direkt vergleichen. Für die Validierung des entwickelten Messverfahrens muss daher geprüft werden, ob zwischen den Messwerten der beiden Methoden ein stetiger Zusammenhang besteht. Um diesen Zusammenhang nachzuweisen, wurde eine Korrelationsanalyse durchgeführt, in der die Ergebnisse der Sichttiefenmessung mit den mittleren Absorptionskoeffizienten der einzelnen Gewässer im Untersuchungsgebiet verglichen wurden. Der berechnete Korrelationskoeffizient ist ein Maß für die Stärke eines eventuell vorhandenen linearen Zusammenhangs. Da er sich mit $r=0,9994$ dem Wert 1 nähert, liegt mit großer Wahrscheinlichkeit ein funktionaler Zusammenhang zwischen Sichttiefe und Absorptionskoeffizienten vor. Die statistische Signifikanz des vermuteten Zusammenhangs wurde mit einem t-Test geprüft.

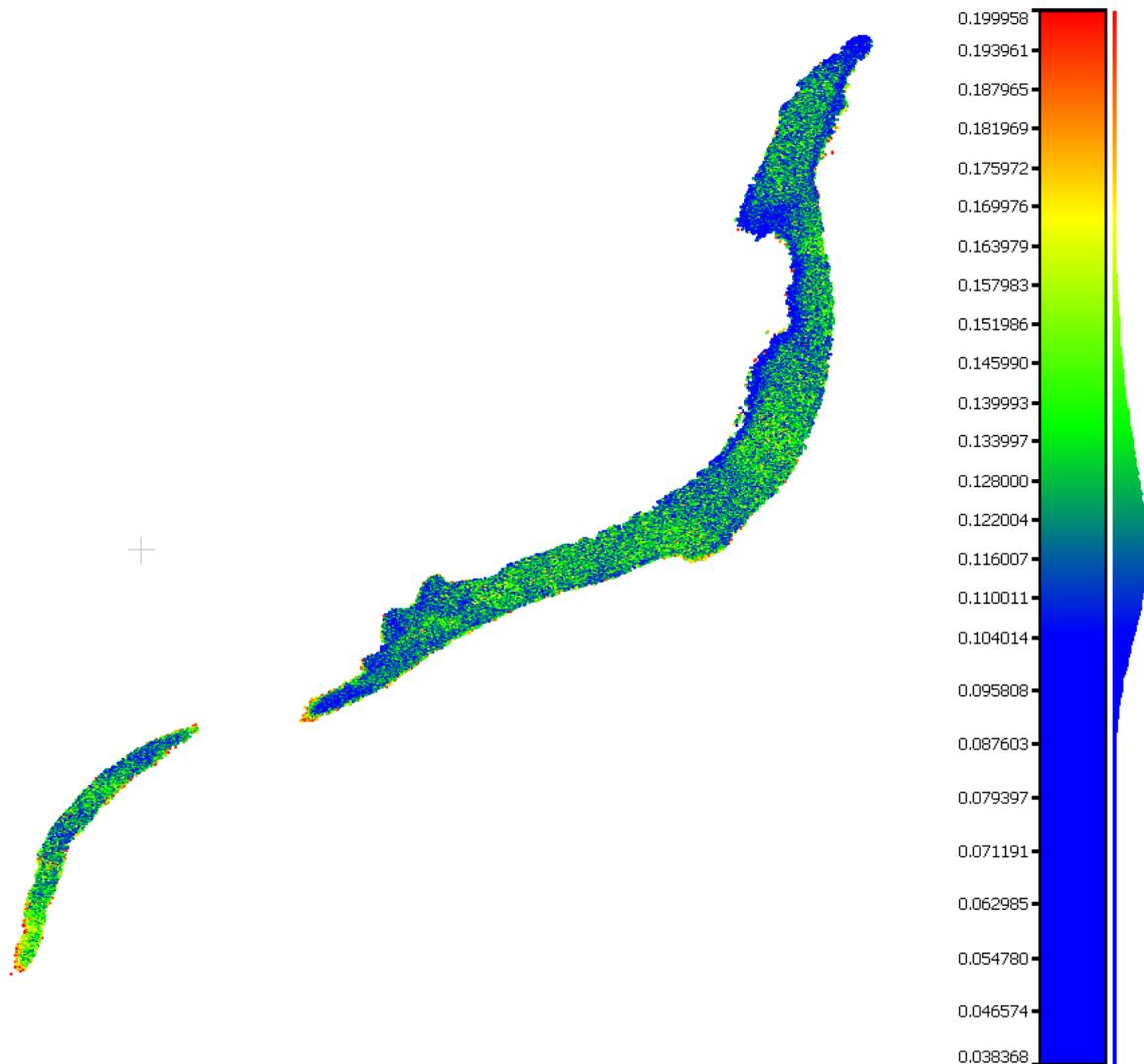


Abb. 9: Visualisierung Absorptions-parameter Altarm Bleddiner Riss

4 Zusammenfassung und Ausblick

Die entwickelte Methode zur Modellierung und Korrektur der Signaldämpfung in Laserbathymetrie-Daten bildet die Grundlage für eine verbesserte Extraktion der Gewässerbodentopographie. Auf dieser Basis können verbesserte Algorithmen zur Ableitung zusätzlicher Bodenpunkte aus den dämpfungskorrigierten Rückstreuquerschnitten entwickelt werden, welche anhand von Referenzdaten zu validieren sind. Für die Beurteilung der Genauigkeit der Gewässerbodenpunkte müssen auch lokale Welleneffekte berücksichtigt werden. Anders als in vielen Anwendungen der Laserbathymetrie angenommen, können bei entsprechenden Genauigkeitsanforderungen auch kleinere Wellenmuster nicht vernachlässigt werden, da sie zu einem signifikanten Lage- und Tiefenversatz am Gewässerboden führen (WESTFELD et al. 2016).

Die flächenhafte Bestimmung integraler Gewässertrübungsparameter durch die Analyse des Signalverlaufs der Laserbathymetriewaveformdaten stellt einen großen Fortschritt zu konventionellen Verfahren dar, bei denen Trübungsparameter nur stichprobenhaft vor Ort erfasst werden. In dieser ersten Untersuchung konnte eine hohe Korrelation zwischen aus den Waveforms abgeleiteten Trübungsparametern und durch konventionelle Secchischeibenmessungen ermittelten Referenzwerten festgestellt werden. Das Verfahren ist nicht nur für die Untersuchung von Binnengewässern sondern auch für meereskundliche Fragestellungen wie beispielsweise die Untersuchung von Meeresplankton (SULLIVAN et al. 2012) relevant.

5 Literaturverzeichnis

- DONEUS, M., DONEUS, N., BRIESE, C., PREGESBAUER, M., MANDLBURGER, G. & VERHOEVEN, G., 2003: Airborne laser bathymetry - detecting and recording submerged archaeological sites from the air. *Journal of Archaeological Science* **40** (4), 2136-2151.
- GUENTHER, G.C., CUNNINGHAM, A.G., LAROCQUE, P.E. & REID, D.J., 2000: Meeting the accuracy challenge in airborne lidar bathymetry. *Proceedings of EARSeL-SIG-Workshop LIDAR, Dresden/FRG*. 2000.
- HARSDORF, S. & REUTER, R., 1999: Laser remote sensing in highly turbid waters: validity of the lidar equation. *Industrial Lasers and Inspection (EUROPTO Series)*, International Society for Optics and Photonics, 369-377.
- LIEW, S.C., 2001: Optical Remote Sensing. Tutorial, <http://www.crisp.nus.edu.sg/~research/tutorial/optical.htm>
- PFENNIGBAUER, M. & ULLRICH, A., 2011: Multi-wavelength airborne laser scanning. *Proceedings of the International Lidar Mapping Forum, ILMF, New Orleans*.
- RICHTER, K., BLASKOW, B., STELLING, N. & MAAS, H.-G., 2015: Reference value provision schemes for attenuation correction of full-waveform airborne laser scanner data. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* (2015), II-3/W5, 65-72.
- SECCHI A., 1866: Esperienze per determinare la trasparenza del mare. Sul moto ondoso del mare e su le correnti di esso specialmente su quelle littorali, A. Ciadi (ed.), Rome, 258-288.
- SULLIVAN, J., TWARDOWSKI, M., DONAGHAY, P., RINES, J., MCFARLAND, M., TALAPATRA, S., KATZ, J., CHURNSIDE, J., WEIDEMANN, A., 2012: Biological thin layers: history, ecological significance and consequences to oceanographic sensing systems. *SPIE Defense, Security, and Sensing, International Society for Optics and Photonics*.
- WEIß, R., 2015: Erprobung der Laserbathymetrie an Bundeswasserstraßen im Binnenbereich. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 24, T. Kersten (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 16.-18. März 2015 an der Universität zu Köln, Tagungsband auf CD-ROM, 27-34*.
- WESTFELD, P., RICHTER, K. MAAS, H.-G., WEIß, R., 2016: Analysis of the Effect of Wave Patterns on Refraction in Airborne LiDAR Bathymetry. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **41** (1/W2).

Baumartenklassifikation mittels multi-temporalen Full-Waveform Laserscanning-Daten

MARKUS HOLLAUS¹, NORBERT PFEIFER¹, GOTTFRIED MANDLBURGER¹ & PATRICK INSINNA²

Zusammenfassung: Flugzeuggetragene Laserscanning- (ALS) Daten haben sich im letzten Jahrzehnt als geeignete Datenquelle zur dreidimensionalen Beschreibung von Waldbeständen als auch für die Ableitung von Forstparametern etabliert. In diesem Beitrag werden multi-temporale Full-Waveform (FWF) ALS-Daten für die Unterscheidung von Laub- und Nadelwald analysiert. Auf Einzelbaum- sowie auf Rasterniveau werden die Unterschiede in der vertikalen Punktverteilung als auch die Veränderung der FWF-Attribute als Grundlage für eine Baumartenklassifizierung herangezogen. Speziell werden auch Parameter vorgestellt, die als Differenzen oder Verhältnisse berechnet werden. Damit kann die Übertragbarkeit zwischen verschiedenen, auch nicht kalibrierten Missionen erleichtert werden. Die Analysen wurden im Fürstentum Liechtenstein durchgeführt und zeigen das große Potential multi-temporaler FWF-ALS-Daten für die Unterscheidung von Nadel- und Laubwald.

1 Einleitung

Flugzeuggetragene Laserscanning Daten haben sich im letzten Jahrzehnt als geeignete Datenquelle zur dreidimensionalen Beschreibung von Waldbeständen (z.B. Waldfläche, Struktur) als auch für die Ableitung von Forstparametern (z.B. Baumhöhen, Holzvorrat) etabliert (HOLLAUS et al. 2009a; HYYPPÄ et al. 2008; LEITERER et al. 2013; LIM et al. 2003; NÆSSET 2007; NÆSSET et al. 2004). Speziell in Berggebieten kommt dem flugzeuggetragenen Laserscanning (engl. Airborne Laserscanning – ALS) als aktive Fernerkundungsmethode eine besondere Bedeutung zu, da einerseits Höheninformation des überschrmtten, oftmals stark reliefierten Geländes als auch der Kronenoberfläche erhalten werden und andererseits topographische Einflüsse (z.B. Abschattung) auf ein Minimum reduziert sind. Aus den ALS Daten lassen sich neben dem hochgenauen digitalen Geländemodell (DGM) Baumhöhen direkt ableiten und stellen somit eine fundamentale Eingangsgröße für die Berechnung vieler Forstparameter (z.B. Bandeshöhen, Stammvolumen, Biomasse) dar. Die Nützlichkeit von ALS Daten für die Forstwirtschaft wurde im letzten Jahrzehnt im Rahmen von wissenschaftlichen Studien und Anwendungsbeispielen eindeutig dargestellt und belegt. So belegen beispielsweise Studien von NÆSSET (2014; 2007) oder KANGAS et al. (2014) die erfolgreiche Einbindung von ALS Daten in operationelle Waldinventuren.

Für die Ableitung forstlicher Parametern aus ALS Daten ist es oftmals sinnvoll eine Stratifizierung in Laub- und Nadelwald und/oder in locker und dicht bestockte Bestände vorzunehmen. Für die Klassifizierung in Laub- und Nadelwald hat sich speziell der Mehrwert

¹ TU Wien, Department für Geodäsie und Geoinformation; Gußhausstraße 27-29, 1040 Wien
E-Mail: [Markus.Hollaus; Norbert.Pfeifer]@geo.tuwien.ac.at

² Landesverwaltung Fürstentum Liechtenstein, Amt für Umwelt, Dr. Grass-Straße 12, Postfach 684, 9490 Vaduz; E-Mail: Patrick.Insinna@llv.li

von Full-Waveform (FWF) ALS-Daten als hilfreich erwiesen (BRANDTBERG 2007; HOLLAUS et al. 2009b; KIM et al. 2009; REITBERGER et al. 2006). Neben der höheren Anzahl gemessener Echos pro Schuss und dem verringerten Mindestabstand aufeinanderfolgender Echos, liefert das FWF-ALS besonders zusätzliche Attribute pro Echo (z.B. Echobreite, Amplitude, Rückstreuquerschnitt).

Ziel des vorliegenden Beitrages ist es, die Möglichkeiten von multi-temporalen (belaubt und laublos) FWF-ALS-Daten für die Unterscheidung von Laub- und Nadelwald zu analysieren. Die Untersuchung analysiert i) den Unterschied zwischen einem Einzelbaum-basierten und einem flächigen Ansatz, ii) Parameter der (geometrischen) Punktverteilung und Parameter aus den FWF-Attributen, und iii) deren Änderung aufgrund der Phänologie. Die Untersuchungen werden im Fürstentum Liechtenstein durchgeführt.

Im folgenden Kapitel 2 werden sowohl das Untersuchungsgebiet als auch die verwendeten Daten beschrieben. Im Kapitel 3 werden die angewandten Methoden zur Vorprozessierung der ALS-Daten und zur Klassifizierung von Laub- und Nadelwald beschrieben. Im Kapitel 4 werden Ergebnisse präsentiert und diskutiert. Abschließend werden in Kapitel 5 eine Zusammenfassung und ein Ausblick gegeben.

2 Untersuchungsgebiet und Daten

2.1 Untersuchungsgebiet

Das Untersuchungsgebiet umfasst das Fürstentum Liechtenstein mit einer Fläche von ca. 160 km². Das Untersuchungsgebiet ist geprägt durch den Hauptsiedlungsraum im Rheintal im Westen und der Hochgebirgslandschaft des Rätikon im Osten. Die Landesfläche lässt sich in ca. 11% Siedlungsraum, 33% landwirtschaftlicher Nutzflächen, ca. 41% Wald und ca. 15% unproduktiver Flächen einteilen. Basierend auf der dritten Erhebung der Landeswaldinventur befinden sich 42% des Waldes unter 1000 m Seehöhe und 58% über 1000 m Seehöhe. Der mittlere Holzvorrat liegt bei 409 m³/ha und setzt sich aus 77% Nadelholz und 23% Laubholz zusammen. Hinsichtlich des Holzvorrates sind die dominierenden Baumarten die Fichte mit 54% und die Buche mit 11%. In Abb. 1 ist ein aus ALS Daten abgeleitete Kronenhöhenmodell für das gesamte Untersuchungsgebiet mit den dazugehörigen ALS Flugtrajektorien abgebildet.

2.2 Daten

2.2.1 ALS-Daten

Die Landesfläche wurde mit zwei ALS Kampagnen am 15. April sowie am 26. August 2015 von der Fa. TopScan GmbH befliegen. Die Aprilbefliegung deckt die Tallagen bis 1000 m Seehöhe ab und erfolgte unter laubfreien Bedingungen. Die Augustbefliegung umfasst die Gebirgszonen mit belaubten Vegetationsstand und schneefreier Topographie. Der Überlappungsbereich der beiden Befliegungen beschränkt sich auf wenige ALS Streifen (vergl. Abb. 1). Beide Befliegungen erfolgten mit einem Full-Waveform ALS Scanner (LiteMapper 7800 – Riegl LMS-Q780) mit einer Mindestpunktdichte von 4 First-Pulse Punkten pro m². Die mittlere Flughöhe über Grund betrug 800 m mit einem Scanwinkel von $\pm 30^\circ$. Die ALS-Daten wurden in bereits georeferenzierter Form als klassifizierte Streifendaten zur Verfügung gestellt.

2.2.2 Waldinventurdaten

Vom Amt für Umwelt der Landesverwaltung Fürstentum Liechtenstein wurden 54 voll kluppierte ca. 700 m² große Probekreisflächen zur Verfügung gestellt. Für jede Probefläche wurden die Positionen der Einzelbäume, deren Brusthöhendurchmesser, Baumhöhen als auch Baumarten erhoben. Der mittlere Holzvorrat der 54 Stichproben liegt bei 267 m³/ha mit einer Standardabweichung von 190 m³/ha. Die mittlere Baumhöhe der 998 Probestämme beträgt 17.8 m (Std.Abw.=6.7 m), der mittlere Brusthöhendurchmesser (BHD) 30.3 cm (Std.Abw.=15.9 cm). Für die Validierung der aus den ALS-Daten generierten Laub-Nadelwaldkarte kann der Grundflächen gewichtete Nadel- / Laubholzanteil der Stichproben herangezogen werden.

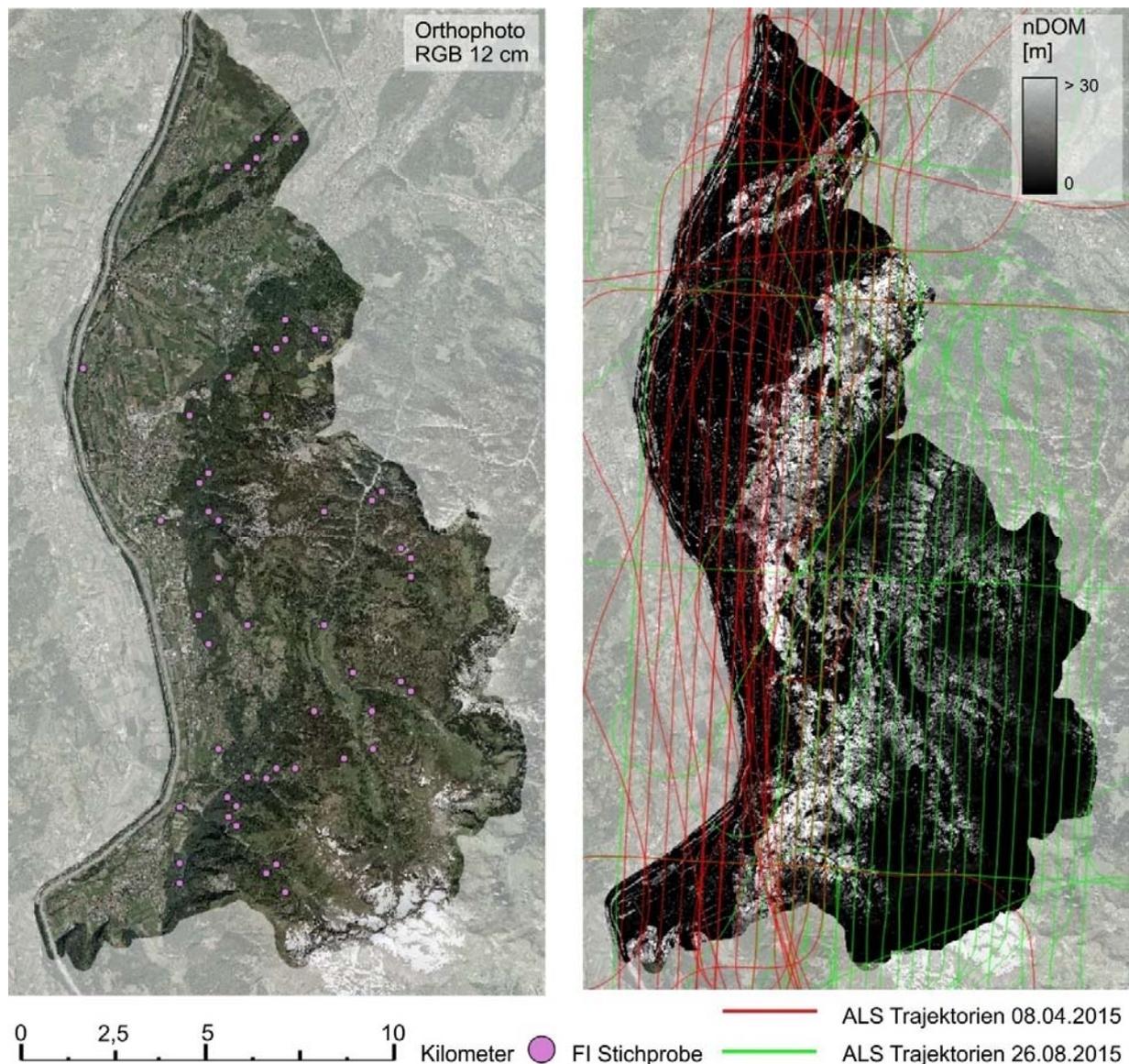


Abb. 1: Untersuchungsgebiet Fürstentum Liechtenstein. Abgebildet ist das RGB Orthophoto überlagert mit den Waldinventurstichproben (links) sowie das Kronenhöhenmodell überlagert mit den ALS Flugtrajektorien (rechts).

3 Methoden

Ausgehend von den georeferenzierten Punktwolken wurde zunächst die Qualität der Georeferenzierung mittels Streifendifferenzen überprüft. Da die mittlere Streifendifferenz von Boden und Dachflächen größtenteils $< \pm 20$ cm beträgt wurde auf eine zusätzliche rigorose Streifenausgleichung verzichtet.

In den nachfolgenden Unterkapiteln werden die angewandten Methoden für die Ableitung der topographischen Modelle, der Waldabgrenzung und auch der Klassifizierung in Nadel- und Laubwald beschreiben.

3.1 Berechnung der topographischen Modelle

Das Geländemodell (DTM) wird aus den klassifizierten Bodenpunkten mittels Delaunay-Triangulierung berechnet und als Raster mit $1 \times 1 \text{ m}^2$ Zellgröße exportiert.

Für die Ableitung des digitalen Oberflächenmodells (DOM), oftmals auch als Kronenmodelle bezeichnet, werden alle Laserpunkte der Klassen Boden, niedrige-, mittlere- und hohe Vegetation herangezogen. Um eine bestmögliche Vergleichbarkeit der Oberflächenmodelle vom April- und Augustflug zu erzielen, werden diese durch eine einheitliche, landbedeckungsabhängige Berechnungsmethode (HOLLAUS et al. 2010) erstellt. Diese Methode verfolgt einen kombinierten Berechnungsansatz, der abhängig von der Oberflächenrauigkeit das DOM mit dem höchsten Punkt pro Rasterzelle bzw. mit lokal in die Punktwolke eingepassten, gleitenden Schrägebenen (engl. moving planes) zusammensetzt. Die „moving planes“ Interpolation wird dabei auf einen ausgedünnten Datensatz (höchste Punkte innerhalb von $0.5 \times 0.5 \text{ m}^2$ Zellen) angewandt um sicher zu stellen, dass das interpolierte Modell bestmöglich die Oberfläche der Baumkronen beschreibt. Diese Berechnungsmethode führt speziell in bewaldeten Gebieten im Vergleich zu einer reinen „moving planes“ Interpolation zu einer verbesserten Beschreibung der Baumkronen als auch der Waldlücken. Des Weiteren hat dieser Ansatz speziell bei Datensätzen mit geringer Punktdichte den Vorteil, dass für Bildelement, in denen keine Punkte liegen, durch die „moving planes“ Interpolation mit Hilfe der benachbarten Punkte eine Höheninformation geschätzt werden kann.

Das normalisierte digitale Oberflächenmodell (nDOM) wird durch Subtraktion des DTMs vom DOM berechnet und hat so wie das DTM und das DOM eine Pixelgröße von $1 \times 1 \text{ m}^2$.

3.2 Waldabgrenzung

Um das Untersuchungsgebiet nur auf die Waldflächen einzugrenzen, wird eine Waldabgrenzung mit Hilfe der Methode von EYSN et al. (2012) durchgeführt. Dieser Ansatz erlaubt es, Mindestwerte für Baumhöhen, Überschirmungsgrad, Breite und Fläche zu berücksichtigen. Der Überschirmungsgrad wird dabei für jeweils drei benachbarte Bäume ermittelt (= Fläche der konvexen Hülle / Summe der Kronenflächen) (EYSN et al. 2011). Die Position der Einzelbäume wird mittels lokalem Maximafilter anhand des nDOMs berechnet, die Kronenradien werden mittels lokal angepasster Funktionen in Abhängigkeit der Baumhöhe und der Seehöhe geschätzt. Für die vorliegenden Untersuchungen wurden die in Tab. 1 zusammengefassten Mindestwerte angewandt. Die für die Baumartenklassifizierung erforderlichen Analysen werden nur innerhalb der abgegrenzten Waldfläche durchgeführt.

Tab. 1: Angewandte Parameter für die Waldabgrenzung

Parameter	Baumhöhe	Fläche	Breite	Überschirmungsgrad
Mindestwerte	≥3.0 m	≥250 m ²	≥5.0 m	≥10%

3.3 Baumartenklassifizierung

Aufgrund der Tatsache, dass die ALS-Befliegungen der Tallagen unter laublosen Bedingungen (April) und jene der Gebirgsregionen unter belaubten Zustand (August) erfolgten, werden die Eingangsdaten für die Klassifizierung für beide Datensätze getrennt berechnet.

Generell orientiert sich der Klassifikationsansatz an den geometrischen und radiometrischen Rückstreuungseigenschaften des Laserlichts und der vertikalen Punktverteilung für Nadel- und Laubwald sowie deren Veränderungen zwischen belaubten und unbelaubten Zustand.

Tab. 2: Berechnete FWF ALS Parameter. Die Berechnungen werden für Rasterzellen als auch Einzelbaumkronen durchgeführt.

Parameter	Beschreibung
EW _{veg-top} [ns]	Durchschnittliche Echobreite aller Vegetations-Echos der obersten Kronenschicht (DOM - 1.0 m)
EW _{diff-veg-top} [ns]	EW _{veg-top} (August) - EW _{veg-top} (April)
A _{ratio-last} [%]	100 * (Summe der Amplituden aller last Echos in der Bodenschicht (nDOM < 0.3 m)) / (Summe der Amplituden aller last Echos)
A _{ratio-first} [%]	100 * (Summe der Amplituden aller first Echos in der Bodenschicht (nDOM < 0.3 m)) / (Summe der Amplituden aller first Echos)
A _{ratio-veg-ground} [%]	100 * (Summe der Amplituden aller Vegetationsechos (nDOM > 1.5 m)) / (Summe der Amplituden aller Bodenechos (nDOM < 0.3 m))
R _{first} [dB]	Durchschnittliche Reflektivität aller first Echos
R _{last} [dB]	Durchschnittliche Reflektivität aller last Echos
R _{all} [dB]	Durchschnittliche Reflektivität aller Echos
R _{veg} [dB]	Durchschnittliche Reflektivität aller Vegetations-Echos (nDOM > 1.5 m)
R _{ground} [dB]	Durchschnittliche Reflektivität aller Boden-Echos (nDOM < 0.3 m)
R _{veg-top} [dB]	Durchschnittliche Reflektivität aller Vegetations-Echos der obersten Kronenschicht (DOM - 1.0 m)
R _{diff-veg-top} [dB]	R _{veg-top} (August) - R _{veg-top} (April)
E _{ratio-first} [%]	100 * (Summe der Energie (= Amplitude * Echobreite) aller first Echos in der Bodenschicht (nDOM < 0.3 m)) / (Summe der Energie aller first Echos)
E _{ratio-last} [%]	100 * (Summe der Energie (= Amplitude * Echobreite) aller last Echos in der Bodenschicht (nDOM < 0.3 m)) / (Summe der Energie aller last Echos)
E _{ratio-veg-ground} [%]	100 * (Summe der Energie aller Vegetationsechos (nDOM > 1.5 m)) / (Summe der Energie aller Bodenechos (nDOM < 0.3 m))
HQ ₉₀ [m]	90-stes Höhenquantil aller Vegetationsechos (nDOM > 1.5 m)
HQ ₅₀ [m]	50-stes Höhenquantil aller Vegetationsechos (nDOM > 1.5 m)
HQ _{ratio-90-50} [%]	100 * HQ ₅₀ / HQ ₉₀
HQ _{diff-50} [m]	HQ ₅₀ (August) - HQ ₅₀ (April)
D _{vegtop-veg} [%]	100 * (Anzahl der Echos in der obersten Kronenschicht (DOM - 1.0 m)) / (Anzahl aller Vegetationsechos (nDOM > 1.5 m))
D _{veg-ground} [%]	100 * (Anzahl der Echos in der Vegetationsschicht (nDOM > 1.5 m)) / (Anzahl aller Bodenechos (nDOM < 0.3 m))
D _{diff-veg-ground} [%]	D _{veg-ground} (August) - D _{veg-ground} (April)

Die Klassifizierung wird sowohl auf Rasterniveau mit einer Zellgröße von $9 \times 9 \text{ m}^2$ als auch auf Einzelbaumniveau durchgeführt. Die hierfür erforderlichen Einzelbäume werden anhand der 3D Punktwolke mittels lokaler Maximasuche (Suchradius = 1.5 m) detektiert (siehe auch Abschnitt 3.2, dort jedoch im Raster). Für die Analysen auf Einzelbaumniveau werden die ALS-Punkte rund um die gefundenen Einzelbaumpositionen in einem Suchzylinder mit einem Radius von 2 m analysiert. Die in Tab. 2 beschriebenen Parameter werden aus den FWF-ALS-Daten berechnet und dienen als Eingangsgrößen für die Klassifizierung in Nadel- und Laubwald.

Die Klassifizierung in Nadel- und Laubwald erfolgt mit einem Entscheidungsbaum. Die Auswahl der Parameter und deren Schwellwerte erfolgt anhand einer visuellen Analyse der FWF-ALS-Parameter unter Einbeziehung eines NIR-Orthophotos in welchem Laub- und Nadelwald gut differenzierbar sind.

4 Ergebnisse und Diskussion

In Abb. 2 ist eine Auswahl an Raster-basierten FWF-ALS-Parametern dargestellt. Als visuelle Orientierungshilfe für die Unterscheidung von Laub- und Nadelwald steht ein NIR-Orthophoto von 2001 zur Verfügung. In der Reflektivität des Aprilfluges ($R_{\text{veg-top-April}}$) lassen sich Laub- und Nadelwald erwartungsgemäß sehr gut unterscheiden. Dabei weist der Nadelwald deutlich höhere Werte auf als der Laubwald, da dieser im April noch großteils laublos war. Ein gänzlich anderes Bild zeigt sich in der Reflektivität des Augustfluges ($R_{\text{veg-top-August}}$). Hier zeichnet sich der Laubwald aufgrund seines dichten Blätterdaches und der damit einhergehenden hohen Reflektivität im nahen Infrarot aus. Die Differenz der beiden Reflektivitäten ($R_{\text{veg-top-August}} - R_{\text{veg-top-April}}$) eignet sich daher auch sehr gut für die Differenzierung in Laub- und Nadelwald. Da für die Berechnung dieser Reflektivitäten nur Echos aus der obersten Kronenschicht verwendet wurden gibt es keinen Einfluss der Boden- bzw. bodennahen Vegetation. Ein weiteres Kriterium für die Unterscheidung in Laub- und Nadelwald stellt die vertikale Punktverteilung, repräsentiert als 50-% Höhenquantil aller Vegetationsechos, dar. Dabei sind die Änderungen ($HQ_{50\text{-August}} - HQ_{50\text{-April}}$) im Nadelwald sehr gering, wohingegen im Laubwald eine Erhöhung des HQ_{50} um z.T. mehrere Meter auftritt. Schlussendlich wurde noch die Veränderung der vertikalen Punktdichteverteilung in der Vegetationsschicht zwischen belaubten und unbelaubten Zustand ($D_{\text{vegtop-veg-August}} - D_{\text{vegtop-veg-April}}$) für die Klassifizierung genutzt. Die starke Zunahme im Laubwald weist darauf hin, dass im belaubten Zustand die meisten Echos von der obersten Kronenschicht zurückgestreut werden, wohingegen im unbelaubten Zustand die Echos über die gesamte Baumhöhe verteilt sind. Die Amplituden-Ratio Werte weisen ebenso wie die Reflektivitäten eine gute Trennbarkeit zwischen den beiden Klassen auf und stellen eine interessante Alternative zu den kalibrierten Reflektivitäten dar. Dies gilt speziell für diskrete ALS Daten bei denen im Regelfall keine kalibrierten Reflektivitäten vorliegen. Die Parameter basierend auf Echobreiten zeigen für dieses Untersuchungsgebiet keine deutliche Unterscheidung zwischen Laub- und Nadelwald.

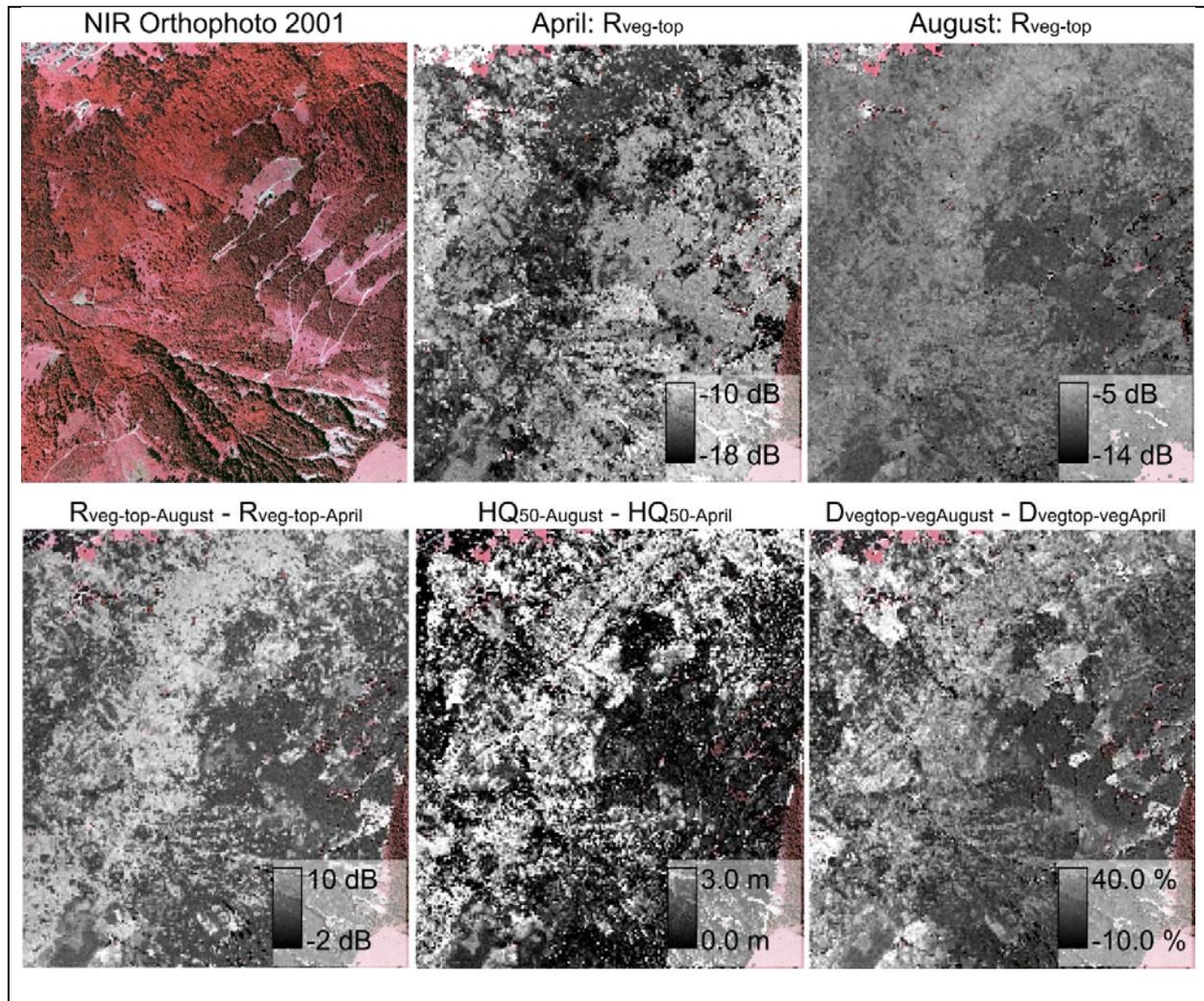


Abb. 2: Rasterbasierte FWF-ALS-Parameter (Zellgröße 9×9 m²). Die Ausdehnung des Ausschnittes beträgt 1600×1900 m².

In Abb. 3 sind ausgewählte Parameter für einzelnen Baumkronen dargestellt. Um den Detailierungsgrad des Einzelbaum-basierten Ansatzes zu verdeutlichen wurde in Abb. 3 nur der südliche Teil von Abb. 2 vergrößert dargestellt. Im Allgemeinen zeigt sich ein ähnliches Bild wie beim Raster-basierten Ansatz. Im Detail zeigt sich jedoch, dass die Reflektivitätswerte eine bessere Trennung als die Änderung der Höhenquantilen (HQ₅₀-August – HQ₅₀-April) oder der Änderung der vertikalen Punktdichteverteilung in der Vegetationsschicht zwischen belaubten und unbelaubten Zustand (D_{vegtop-veg}-August – D_{vegtop-veg}-April) ermöglichen. Dies kann durch die robustere Abschätzung der Quantilenwerte und der Punktdichteverteilungen beim Raster-basierten Ansatz erklärt werden.

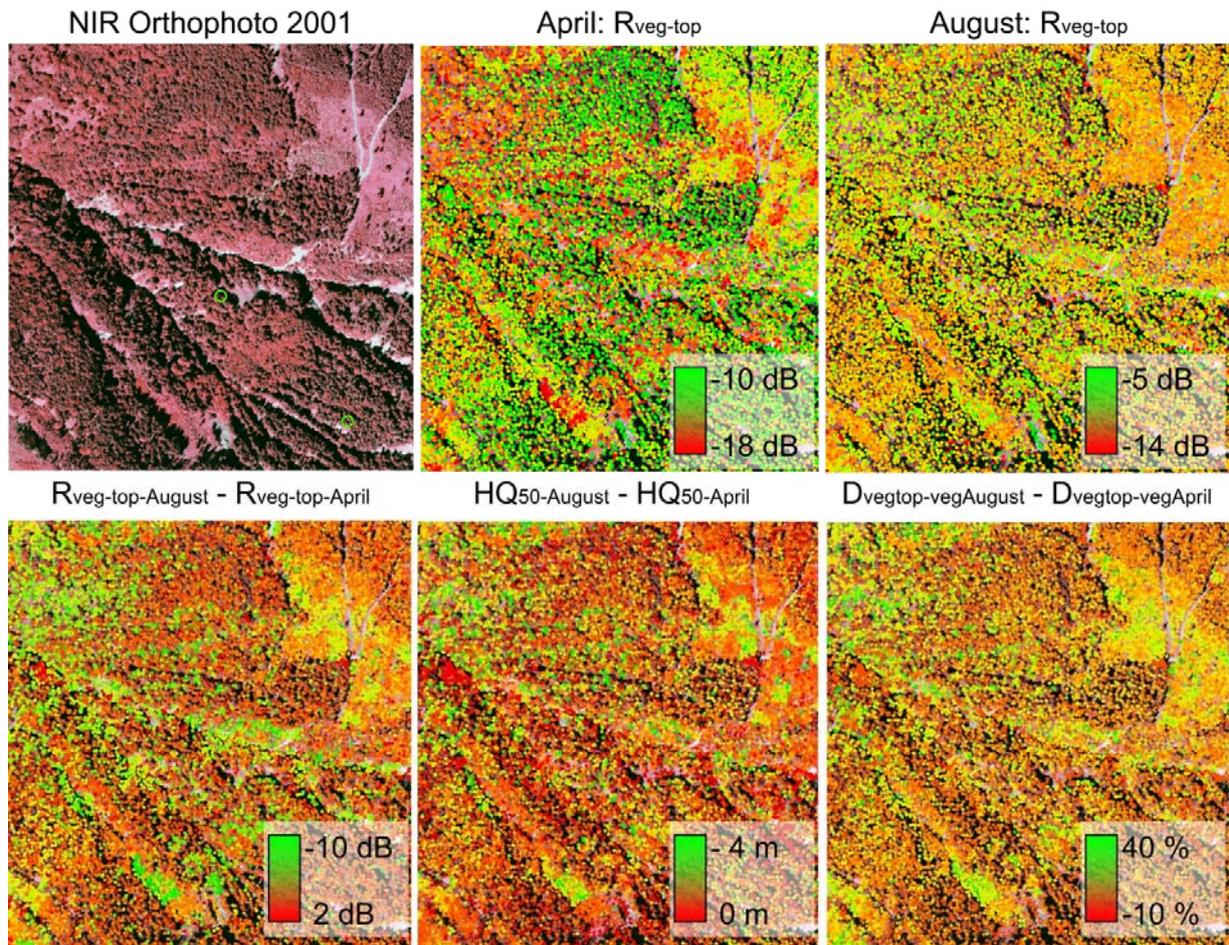


Abb. 3: Einzelbaumbasierte FWF-ALS-Parameter (Kronengröße \varnothing 4 m). Die Ausdehnung des Ausschnittes beträgt $1285 \times 815 \text{ m}^2$.

Die durchgeführte wissenschaftliche Entscheidungsbaumklassifizierung liefert sowohl für den Raster- als auch für den Einzelbaumbasierten Ansatz sehr gute Ergebnisse. Der visuelle Vergleich des Klassifikationsergebnisses mit dem NIR-Orthophoto zeigt eine hohe Übereinstimmung. Zu beachten ist dabei, dass das vorliegende NIR-Orthophoto aus dem Jahr 2001 ist und es speziell im westlichen sowie südwestlichen Bereich des Ausschnitts in Abb. 4 zu Aufforstungen gekommen ist. Von den restlichen Waldflächen kann davon ausgegangen werden, dass es in den Jahren zwischen 2001 und 2015 zu keinen wesentlichen Änderungen der Laub-Nadelwaldverteilung gekommen ist. Die in diesem ersten Testgebiet liegenden vier FI Stichproben (siehe Abb. 4) bestätigen die hohe Genauigkeit der Klassifizierung, lassen jedoch auf der zu geringen Anzahl an FI Stichproben noch keine statistischen Aussagen zu.

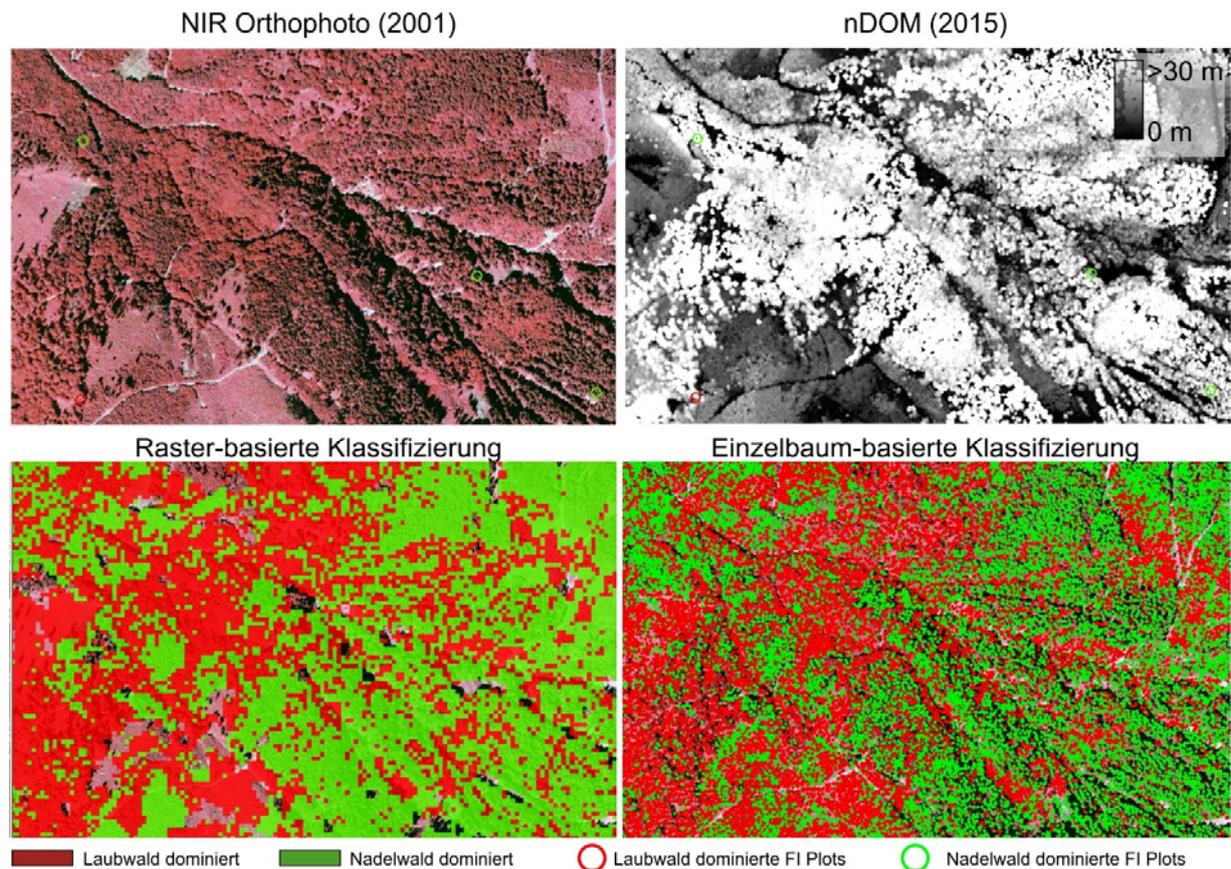


Abb. 4: Raster- und Einzelbaumbasierte Laub- Nadelwaldklassifizierung. Die Ausdehnung des Ausschnittes beträgt $1175 \times 335 \text{ m}^2$.

5 Fazit & Ausblick

In der Untersuchung wurden FWF-Parameter aus ALS über Vegetation untersucht. Speziell wurden auch Parameter vorgestellt, die als Differenzen oder Verhältnisse berechnet werden. Damit sollte die Übertragbarkeit zwischen verschiedenen, auch nicht kalibrierten Missionen erleichtert werden. Die für die Klassifizierung herangezogenen radiometrischen und geometrischen Parameter, sowie deren Veränderungen aufgrund der Änderung der Phänologie, haben sich für die Unterscheidung in Laub- und Nadelwald als sehr nützlich erwiesen. Die Untersuchungen haben ebenso gezeigt, dass mit Hilfe Einzelbaum-basierter Verfahren eine deutlich höhere räumliche Auflösung erzielbar ist. Demgegenüber stehen jedoch deutlich längere Rechenzeiten für die Ermittlung der FWF-ALS-Parameter. Die abgeleitete Baumartenkarte dient als Stratifizierungsgrundlage für die Ableitung weiterer forstlicher Parameter wie z.B. dem Holzvorrat. Die beschriebenen Analysen werden in einem nächsten Schritt auf die Gesamtfläche Liechtensteins ausgeweitet. Damit können anhand der vorhandenen Forstinventur-Stichproben statistische Genauigkeitsanalysen durchgeführt werden.

6 Danksagung

Teile dieser Arbeit wurden im Rahmen des EU FP7/2007-2013 Projekts Advanced_SAR (grant agreement No. 606971) durchgeführt.

7 Literaturverzeichnis

- BRANDTBERG, T., 2007: Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *ISPRS Journal of Photogrammetry and Remote Sensing* **61**, 325-340.
- EYSN, L., HOLLAUS, M., SCHADAUER, K. & RONCAT, A., 2011: Crown coverage calculation based on ALS data. 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems (Silvilaser 2011), **1620**, Hobart, Australia, Oct. 16-20, 10.
- EYSN, L., HOLLAUS, M., SCHADAUER, K. & PFEIFER, N., 2012: Forest Delineation Based on Airborne LIDAR Data. *Remote Sensing* **4** (3), 762-783.
- HOLLAUS, M., DORIGO, W., WAGNER, W., SCHADAUER, K., HÖFLE, B. & MAIER, B., 2009a: Operational wide-area stem volume estimation based on airborne laser scanning and national forest inventory data. *International Journal of Remote Sensing* **30** (19), 5159-5175.
- HOLLAUS, M., MÜCKE, W., HÖFLE, B., DORIGO, W., PFEIFER, N., WAGNER, W., BAUERHANSL, C. & REGNER, B., 2009b: Tree species classification based on full-waveform airborne laser scanning data. *Silvilaser 2009*, October 14-16, College Station, Texas, USA, 54-62.
- HOLLAUS, M., MANDLBURGER, G., PFEIFER, N. & MÜCKE, W., 2010. Land cover dependent derivation of digital surface models from airborne laser scanning data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **38** (3A), 221-226.
- HYYPÄ, J., HYYPÄ, H., LECKIE, D., GOUGEON, F., YU, X. & MALTAMO, M., 2008: Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing* **29** (5), 1339-1366.
- KANGAS, A., EID, T. & GOBAKKEN, T., 2014: Valuation of Airborne Laser Scanning Based Forest Information. *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*, M. Maltamo, E. Næsset and J. Vauhkonen (Editors), Springer Netherlands, 315-331.
- KIM, S., MCGAUGHEY, R.J., ANDERSEN, H.-E. & SCHREUDER, G., 2009: Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sensing of Environment* **113** (8), 1575-1586.
- LEITERER, R., MÜCKE, W., MORSDORF, F., HOLLAUS, M., PFEIFER, N. & SCHAEPMAN, M., 2013: Operational forest structure monitoring using airborne laserscanning. *Photogrammetrie Fernerkundung Geoinformation* **3**, 173-184.
- LIM, K., TREITZ, P., WULDER, M., ST-ONGE, B. & FLOOD, M., 2003: LIDAR remote sensing of forest structure. *Progress in Physical Geography* **27** (1), 88-106.
- NÆSSET, E., 2014: Area-Based Inventory in Norway – From Innovation to an Operational Reality. *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*, M. Maltamo, E. Næsset and J. Vauhkonen (Editors), Springer Netherlands, 215-240.

- NAESSET, E., 2007: Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. *Scandinavian Journal of Forest Research* **22** (5), 433-442.
- NÆSSET, E., GOBAKKEN, T., HOLMGREN, J., HYYPPÄ, H., HYYPPÄ, J., MALTAMO, M., NILSSON, M., OLSSON, H., PERSSON, Å. & SÖDERMAN, U., 2004: Laser scanning of forest resources: the Nordic experience. *Scandinavian Journal of Forest Research* **19** (6), 482-499.
- REITBERGER, J., KRZYSZEK, P. & HEURICH, M., 2006: Full-waveform analysis of small footprint airborne laser scanning data in the Bavarian forest national park for tree species classification. *International Workshop on 3D Remote Sensing in Forestry, Vienna, Austria, 14-15. Feb.*, 218-227.

Physically-based Multi-temporal LiDAR Traits for Species Discrimination in a Temperate Mixed Forest

FELIX MORSDORF¹, MORITZ BRUGGISSER¹, CARLA GUILLEN ESCRIBA¹,
DANIEL KUEKENBRINK¹, REIK LEITERER¹, FABIAN D. SCHNEIDER¹,
HOSSEIN TORABZADEH¹ & MICHAEL E. SCHAEPMAN¹

Zusammenfassung: Basierend auf einem Datensatz von 873 Einzelbäumen untersuchen wir ob physikalisch-basierte ALS Merkmale einen zur Trennung von Arten geeigneten phänologischen Trend aufweisen. Hierzu berechnen wir die Differenz dieser Merkmale zwischen jeweils zwei belaubten und unbelaubten ALS Datensätzen innerhalb der einzelnen Baumkronen. Reine Höhenmerkmale wie die Baumhöhe bleiben in unserem Testgebiet zwischen verschiedenen Sensoren und Jahreszeiten stabil, und Merkmale, die die Dichte der Vegetation repräsentieren, lassen sich in gewissen Grenzen zur Trennung von Arten benutzen. Bei komplexeren Merkmalen zeigt sich jedoch schnell der Einfluss der Aufnahmegeometrie oder des verwendeten Sensors, so dass hier weitere Studien, möglichst auch mit Modellsimulationen, nötig sind.

1 Introduction

The spatial distribution of tree species is an important parameter when judging the quality of the forest habitats and its biodiversity, and is equally relevant for forest management purposes. Due to the limited large-scale applicability of exhaustive in-situ observations, passive remote sensing approaches have been used extensively for the mapping of tree species in different forest ecosystems. However, some species remain indistinguishable using image based features alone. Airborne laser scanning (ALS) is now an established method in remote sensing of forests, with having the unique advantage of providing a direct link to the structural characteristics of the forest canopy. Some early studies hinted at the potential of using structural features for species classification, but either were lacking in the number of species (e.g. Scandinavian studies often only differ between three tree species) or the number of datasets used. With the ever increasing availability of regional and nationwide ALS datasets, multi-temporal aspects and the comparability of datasets acquired with different sensors and survey settings gets more in the focus. In addition, knowledge about the interpretability of features is important when working towards site independent methods. Thus, we use a hypothesis based selection of a smaller feature set, rather than testing a long list of features in a black box approach typical of many empirical studies. The objectives of this study are to (i) test a set of physically-based, multi-temporal and full-waveform features for their tree species classification potential in a temperate mixed forest and to (ii) to differentiate changes in features resulting from objects from those resulting from sensor properties. To do this, we exploit data from four (two leaf-on, two leaf-off) different

¹ University of Zürich, Remote Sensing Laboratories, Department of Geography, Winterthurerstr. 190, CH-8057 Zürich, E-Mail: felix.morsdorf@geo.uzh.ch

ALS campaigns and use crown outlines together with tree species information of more than 800 trees.

2 Study area and data

2.1 Study area

The Laegeren site is located at N 47° 28', 49" and E 8° 21', 05" at 682 m a.s.l. on the south slope of the Laegeren mountain, approximately 15 km northwest of Zürich, Switzerland. The south slope of the Laegeren marks the boundary of the Swiss Plateau, which is bordered by the Jura and the Alps. The mean annual temperature is 8°C. The mean annual precipitation is 1200 mm and the vegetation period is 170-190 days. The natural vegetation cover around the tower is a mixed beech forest. The western part is dominated by broad-leaved trees, mainly beech (*Fagus sylvatica* L.) and ash (*Fraxinus excelsior* L.). In the eastern part beech and Norway spruce (*Picea abies* (L.) Karst.) are dominant. The forest stand has a relatively high diversity of species, ages, and diameters. The ground cover consists of bare soil, boulders, and litter while existing understory is characterized by a dense herb and shrub coverage. Average ALS based canopy height is 24.9 m, with a maximum of 49 m, and the field based stem density is 290 stems per ha.

2.2 Multi-temporal airborne laser scanning data

In 2010, two airborne laser scanning campaigns were carried out, using RIEGLs LMS-Q680i scanner (under leaf-on conditions) and using RIEGLs LMS-Q560 scanner under leaf-off conditions in a mainly semi-natural, deciduous dominated forest stand in Laegeren (Swiss Jura), yielding in two independent datasets. The used flight strips have an overlap of approximately 50%. Full-waveform features, namely echo width and intensity, were extracted from the full-waveform data using the software RiAnalyse and were assigned to each individual return contained in the multiple-echo point cloud. However, we found systematic differences in the intensity calibration between the two datasets and thus will not use the intensity, but only the echo width feature in this study. Another leaf-on and off dataset was acquired in 2014, this time with an LMS-Q680i in both acquisitions, albeit with slightly larger maximum scan angle (± 22 deg) and lower average point density than the 2010 data. The specifications of all four datasets used in this study are displayed in Table 1. For the 2010 data, the DTM was interpolated using our own algorithm (for details see LEITERER et al. 2013), while for the 2014 data we used the DTM provided by the data supplier, which was created using the TerraScan software suite. For each echo, the height above ground was then calculated as a subtraction of the interpolated DTM value at the x,y coordinates of each respective echo.

Tab. 1: Specification of the four ALS data sets used in this study.

Campaign	Laegern (2010)		Aargau (2014)	
ALS Instrument (Riegli)	LMS-Q560	LMS-Q680i	LMS-Q680i	LMS-Q680i
∅ Altitude (ABG) [m]	500		600	700
Scanning method	rotating multi-facet mirror			
Echo detection method	full-waveform processing			
Pulse length [ns]	~ 4			
Max. scan angle [deg]	±15		±22	
∅ echo density	20	40	15	30
Acquisition date	10. April	1. August	March/April	June/July

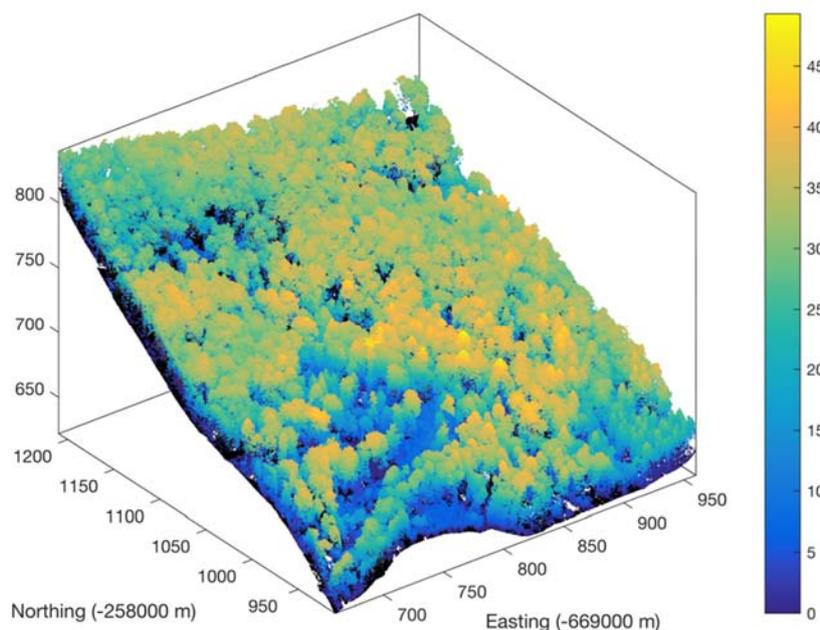


Fig. 1: Leaf-on and leaf-off point cloud of the study area from 2010. Leaf-on data is colored with height above ground, while leaf-off data is using gray-scaled colormap to display height above ground. Note the flux-tower towards the center of the image.

2.3 Validation data

The area of exhaustive single-tree inventory is 330 m × 260 m in size and its elevation ranges from 620 to 810 m above sea level (a.s.l.). The high diversity in respect to species type, age (up to 185 years) and diameter distribution results in a very heterogeneous forest canopy, both horizontally and vertically. Despite the small size of the area, it contains all the tree species

present in the larger Laegeren area and includes very dense deciduous-dominated canopy in the north, and less dense conifer-dominated stands in the south (Fig. 1). Therefore, this area can be considered as a good representative for the forested area around. Collecting tree species data at the single-tree level was part of a comprehensive in-situ forest inventory in April 2013. Each tree with a diameter at breast height (DBH) greater than 0.2 m was accurately located (<1m) using a total station, with a linkage to global coordinates through a set of DGPS points in forest gaps. In addition to tree species, other tree variables were measured as well (e.g. DBH, social status, horizontal crown displacement, number of canopy layers, vitality, approximated crown size, and crown projection area). To get accurate crown segments for the aggregation of features, we used an UAV to acquire images in fall, where leaf senescence aids the delineation of single tree crowns. The platform was an ultra-light drone, called eBee (SenseFly), which was flown on October 21, 2013, at 270 m above ground. Following the geometric rectification of 174 images, the resulting 0.08 m × 0.08 m spatial resolution ortho-mosaic was geometrically co-registered to the ALS dataset. Using a semi-automatic segmentation approach, possible ITCs were detected from combined UAV-derived ortho-mosaic and ALS-derived CHM. The extracted tree crown shapes were visually checked to avoid any conflict or mixed crowns. Crown polygons that appeared to contain more than a single, top-layer tree crown were consequently removed from the final datasets. The in-situ observed trees were linked to the UAV-delineated ITCs using geometrical (coordinates and crown size) and attributive (vertical layer) constraints. As a result, 873 well-matched trees were used as the reference data. Table 2 contains more details about the linked trees and their abundance.

Tab. 2: Abundance of tree species in the reference data set and species identifier used in the box plots.

Species Name	Number of trees	Species identifier
Silver fir (<i>abies alba</i>)	79	11
Norway spruce (<i>picea abies</i>)	40	14
Norway maple (<i>acer platanoides</i>)	23	22
Sycamore maple (<i>acer pseudoplatanus</i>)	118	23
European beech (<i>fagus sylvatica</i>)	376	29
European ash (<i>fraxinus excelsior</i>)	147	31
Large-leaved Lime (<i>tilia platyphyllos</i>)	56	56
Which elm (<i>ulmus glabra</i>)	24	59

3 Methods

We used the crown polygons derived from on the UAV data to subset the raw ALS data based on their x and y coordinates. Only echoes with their x,y-location inside the crown polygon were retained for the subsequent analysis. For the computation of the features, we used an additional height threshold of 3 m above ground to differentiate crown from non-crown returns. The total

number of echoes within the polygon is only used when computing the fraction of crown returns in comparison to all returns. For each of the crown segments, we computed a set of ten traits or features, which are listed in Table 3, along with a description of how they were computed and what was the hypothesis behind choosing this particular feature. Please note that we only had the FW width feature in the 2010 data, while the intensity based features were only useable in the 2014 data. The latter had to do with a non-recoverable semi-empirical intensity correction done by the data provider, which skewed the distribution of intensity values and rendered a multi-temporal comparison of these features impossible. The crown segments used in this study are displayed in Figure 2. After subsetting the point cloud using the crown segments, we computed the features for each crown segment, subtracted the leaf-on feature from the leaf-off one for the same year and displayed these phenological differences as boxplots with mean and stand and deviation for each of the eight species in the results section.

Tab. 3: List of features that were computed from the point clouds contained by the crown segments.

Feature/Trait	Computation	Hypothesis/Remarks
max. height	maximum echo height above ground	should be stable for conifers, impacted by leaf state for deciduous trees
median height	median of all echo heights > 3 m above ground	should be influenced by leaf state in deciduous tree
std. height	standard deviation of all height values > 3m	should be influenced by crown shape
median int.	median of all intensity values within crown > 3 m	should indicate small scale structure and/or reflectance changes (2014 data only)
std. int.	standard deviation of all intensity values > 3 m	should indicate small scale structure and/or reflectance changes (2014 data only)
fraction of single echoes	number of single echoes divided by total of echoes within crown	should change with phenology and species
fraction of ground echoes	number of single echoes divided by total of echoes within crown	should change with phenology and species
median width	median width of echoes (FW feature)	possible linkage with fine-scale vertical structure and incidence angle (2010 data only)
std. width	standard deviation of echoes (FW feature)	possible linkage with fine-scale vertical structure and incidence angle (2010 data only)

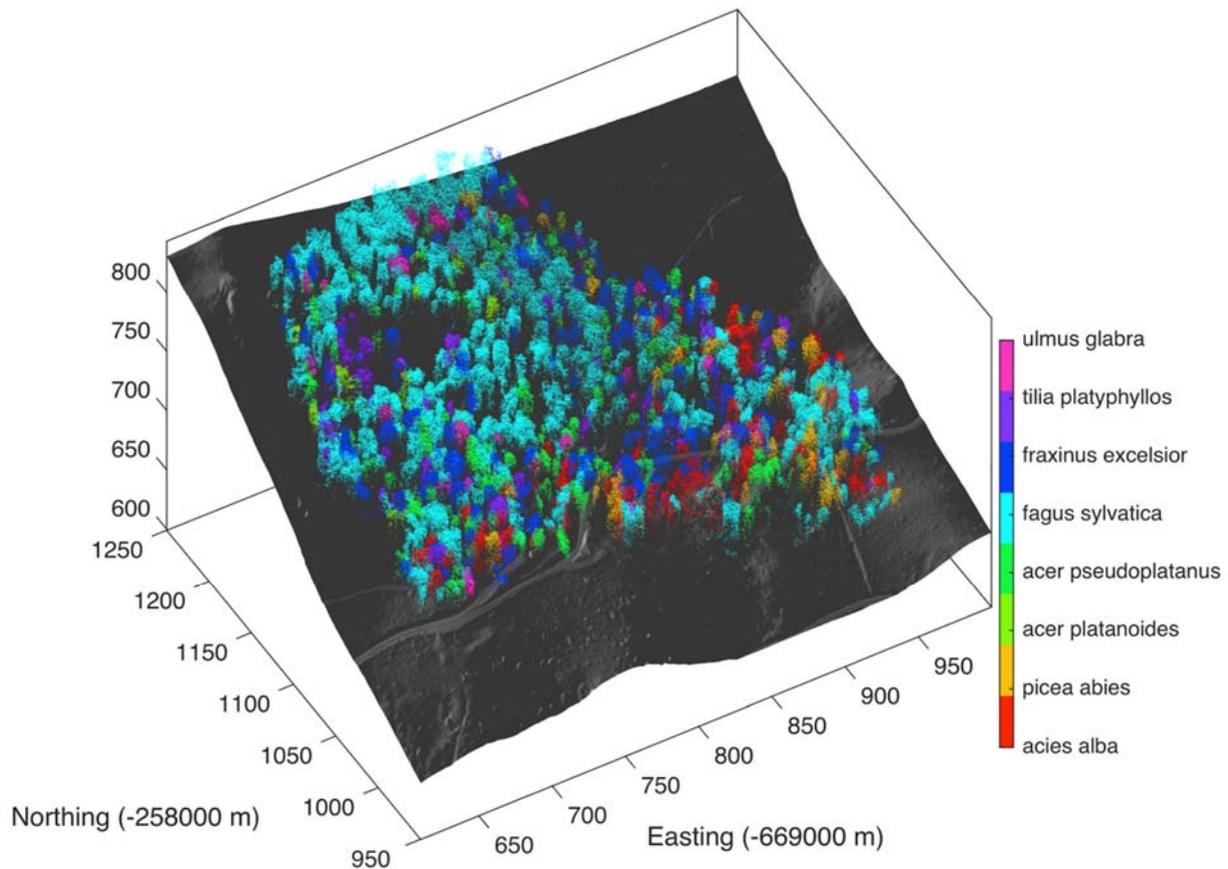


Fig. 2: Point cloud of studied crown segments colored according to their respective field measured species. The ALS data shown here is the 2014 leaf-on data.

4 Results and Discussion

Figure 3 presents the results of both 2010 and 2014 phenologic differences of the features in boxplots, indicating median, 75% and 25% and min/max values of the data. Studying the max and median height features, the conifer tree species (11 and 14) show less seasonal change than the deciduous ones (22-59), more (~ 2 m difference in median) so for median height (especially in 2010) than for maximum height (< 1 m difference in median). Standard deviation is much smaller for maximum height and with the rather small seasonal differences it appears as if leaf-on and leaf-off data would perform comparable when estimating tree height by using the maximum ALS return within a crown segment. There is a much larger difference in the change of median height in 2010 between conifers and deciduous trees (~ 0 m for conifers, ~ -2 m, for deciduous), whereas in 2014 all the median height differences for all trees are negative and in the same order of magnitude, irrespective of the species. A similar behavior can be seen in the fraction of single echoes, which should correspond to the density of canopy material, and is much higher for the conifers in 2010, but again more or less the same in 2014. This unexpected outcome might be partly explained by the two different sensors used in 2010. Echo detection methods and energy output might severely impact both the height distribution of echoes as well

as the fraction of single echoes within a canopy. For the FW data from 2010, the median echo width shows a separation between conifers and deciduous trees, but this again could be an artefact of the two different sensors used. On the other hand, in the 2014 data the median intensity feature shows some differences between conifers and deciduous, potentially linking to the different geometric-optical properties of foliated and non-foliated branches, as the seasonal change for the conifers remains close to zero, while the deciduous trees show higher intensity values in summer, which is likely due to the increased amount of scattering material in summer. Interestingly, the penetration ratio linked feature (fraction of ground echoes) does not show any significant species related difference, neither for the 2010, nor for the 2014 data.

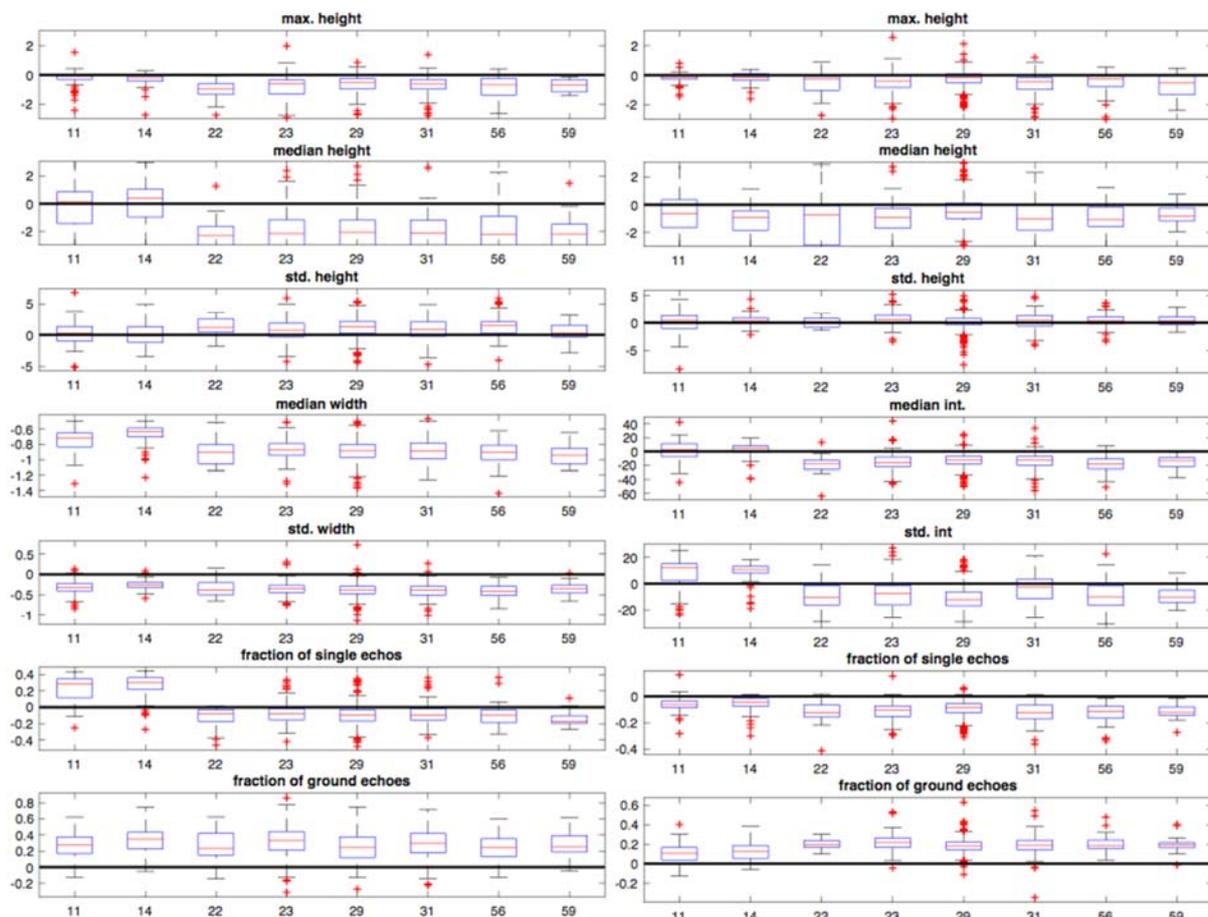


Fig. 3: Boxplots of the **difference** of the selected variables between leaf-off - leaf-on state for 2010 (right panel) and 2014 (left panel). Values are those of the respective leaf-off campaign subtracted by those from the leaf-on campaign. This corresponds well with the negative values for median height seen for the deciduous trees (id > 22). Tree identifiers are: 11 - silver fir, 14 - norway spruce, 22 - norway maple, 23 - sycamore maple, 29 - european beech, 31 - european ash, 56 - large-leaved lime, 59 - which elm. The red line denotes the median, the top and bottom of the blue box the 75% and 25% percentiles, the black bars the min and max values and the red crosses outliers.

5 Conclusions and Outlook

Using a large dataset of 873 field-labelled crown polygons, we derived a set of physically-based features and investigated if their phenological differences could be linked to the eight different tree species present in our dataset. Maximum height was stable across acquisitions and seasons, showing how robust tree height measurements based on ALS data are. Intensity was very useful to show the difference in leaf state. Unfortunately, some of the more complex features showed different behavior and in some cases it was not clear if the changes were due to differences in instruments and survey configurations or actual changes within the vegetation canopy. To get the most species information out of the ALS data, more studies of this type and very likely accompanying model simulations using ray-tracing and virtual tree models are needed.

6 References

LEITERER, R., MÜCKE, W., MORSDORF, F., HOLLAUS, M. PFEIFER, N. & SCHAEPMAN, M.E., 2013: Operational Forest Structure Monitoring Using Airborne Laser Scanning. *Photogrammetrie Fernerkundung Geoinformation* **3**, 173-184.

Analysen zur Datumsfestlegung in photogrammetrischen Projekten großer Volumina

JAN REZNICEK¹, HEIDI HASTEDT¹, TANJA EKKEK¹, THOMAS LUHMANN¹ & CHRISTIAN JEPPIING¹

Zusammenfassung: Photogrammetrische Anwendungen erfordern insbesondere bei industriellen Aufgaben hohe Genauigkeitsniveaus und zwangsfreie Datumsfestlegungen. In der Regel werden bei einer freien Netzausgleichung photogrammetrische Maßstäbe in das Messvolumen eingebracht. Bei der statistischen Analyse der Bündelausgleichungsergebnisse eines großvolumigen Testszenarios mit einer großen Anzahl Maßstabsstrecken sowie einer langen interferometrisch bestimmten Strecke verweisen unterschiedliche Maßstabsfestlegungen auf besondere Effekte im Objektraum. Einerseits kann festgestellt werden, dass erhebliche Korrelationen zwischen den Prüflängen vorliegen, die in Abhängigkeit zur Konfiguration zu betrachten sind. Andererseits kann nachgewiesen werden, dass besondere Effekte in den Standardabweichungen der Objektpunkte durch die Aufnahme Konfiguration zu verzeichnen sind. Die Analyse der Konfidenzellipsoide verweist auf Systematiken, die im Hinblick auf die Datumsfestlegung näher zu betrachten sind. Im folgenden Beitrag sollen die aufgedeckten Effekte auf Basis eines großvolumigen Testszenarios spezifiziert und erläutert werden.

1 Motivation

Sollen Objekte großer Volumina hochgenau und dreidimensional erfasst werden, so sind häufig verschiedene Herausforderungen zu bewältigen. Aufgabenstellung und Genauigkeitsniveau bilden dabei den einschränkenden Faktor hinsichtlich der Wahl des Messverfahrens. Soll berührungslos und taktil mit einer Genauigkeit im Submillimeter-Bereich gemessen werden, so minimiert sich die Auswahl des Messverfahrens im Allgemeinen auf Lasertracking und Photogrammetrie. Bei der photogrammetrischen Erfassung ist dabei zu berücksichtigen, dass die Wahl einer geeigneten rückführbaren Maßverkörperung für eine zwangsfreie Ausgleichung schwierig ist. Handelsübliche Maßverkörperungen liegen in der Regel in Längen zwischen 1 m bis 2.5 m (ggf. 5 m) vor. Bei einem größeren Messvolumen müssen diese also während der Messung versetzt oder ein Verlust in Genauigkeit und Zuverlässigkeit der Skalierung akzeptiert werden. KAHMEN (2015) und KAHMEN & LUHMANN (2016) zeigen einen Ansatz zur Konstruktion einer großen Maßverkörperung auf Basis eines Invardrahtes und verweist darauf, dass die äußere Genauigkeit einer photogrammetrischen Bestimmung unter Berücksichtigung einer langen, weniger genauen Maßverkörperung im Vergleich zur Verwendung kurzer, hochgenauer Maßstäbe steigt. Alternativ zur langen Maßverkörperung können stabile Passpunkte im Messvolumen definiert werden, die beispielsweise mit einem Lasertracker vermessen werden. Die Messgenauigkeit verfügbarer Lasertracker ist jedoch derzeit nicht signifikant höher als die erzielbare Genauigkeit der photogrammetrischen Lösung. Ergänzend ist zu berücksichtigen, dass die Systemlagerung unter Zwang auf nicht signifikant genauer bestimmten Passpunkten als nicht unkritisch einzustufen ist.

¹ Jade Hochschule Wilhelmshaven Oldenburg Elsfleth, Institut für Angewandte Photogrammetrie und Geoinformatik, Ofener Straße 16/19, D-26121 Oldenburg, E-Mail: [Jan.Reznicek, Heidi.Hastedt, Tanja.Ekkel, Thomas.Luhmann, Christian.Jepping]@jade-hs.de

Aus wirtschaftlichen Aspekten ist es zudem wünschenswert, den Einsatz eines Lasertrackers in Kombination mit einer photogrammetrischen Erfassung zu vermeiden.

Im Rahmen eines umfangreichen Praxistests sollen die Ergebnisse aus KAHMEN (2015) verifiziert und vertieft analysiert werden. Ein großvolumiges Testszenario konnte dazu aufgebaut werden, um zu ermitteln, inwieweit die Wahl und Anordnung von Systemmaßstäben in photogrammetrischen Projekten großer Volumina Einfluss auf die äußere Genauigkeit haben. In REZNICEK et al. (2016) wurden die Ergebnisse der Genauigkeitsevaluation verschiedener Konfigurationen zur Maßstabs- und Systemfestlegung beschrieben und analysiert. Geprüft wurde die äußere Genauigkeit durch die Bestimmung von Längenmessabweichungen (LME, nach VDI/VDE 2002). Es wurden Ist-Soll-Vergleiche in Bezug auf Referenzlängen herangezogen, die mit einer interferometrischen Lasertrackermessung bestimmt wurden. Die Ergebnisse in REZNICEK et al. (2016) verweisen darauf, dass die alleinige Betrachtung der Längenmessabweichungen nicht ausreichend ist. Verschiedene Auffälligkeiten konnten im Hinblick auf die Standardabweichungen im Objektraum, Korrelationen zwischen den Prüflängen sowie die Datumsfestlegung ermittelt werden. Diese Auffälligkeiten werden in weiterführenden Analysen näher betrachtet, die Ergebnisse und Erkenntnisse werden im Folgenden erläutert.

2 Voruntersuchungen

2.1 Testszenario

Den nachfolgenden Analysen liegt ein Testszenario zu Grunde, welches schematisiert in Abb. 1 dargestellt wird. Im Messvolumen von (20m x 20m x 4m) werden 116 kalibrierte, photogrammetrische Maßstabslängen, die in verschiedenen Konfigurationen für die Maßstabsfestlegung herangezogen wurden, platziert.

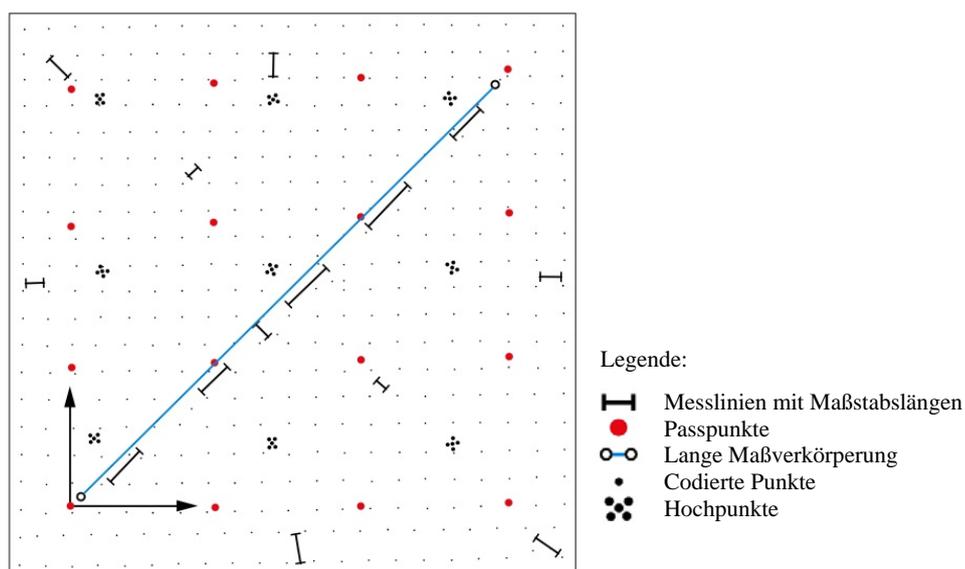


Abb. 1: Grundriss der Testumgebung und Verteilung der Maßverkörperungen und Passpunkte

Zudem wurde eine lange Maßverkörperung in die Diagonale des Messvolumens platziert, deren Länge durch interferometrische Messung mit einem API T3 Lasertracker bestimmt wurde. Für die Untersuchungen wurde eine ALPA 12 WA metric mit einem Leaf Aptus 75 Sensorrückteil eingesetzt (vgl. RIEKE-ZAPP et al. 2005).

2.2 Auswertemethoden

Die Konfiguration der Bilddatenerfassung ist vergleichbar zu üblichen luftgestützten Anordnungen. Die Bündelblockausgleichung mit Simultankalibrierung als freie Netzausgleichung mit Gesamtpurminimierung erfolgte mit dem AICON 3D Studio sowie AXIOS Ax.Ori. In Bezug auf die Verzeichnung wurden Parameter der radial-symmetrischen und der radial-asymmetrischen und tangentialen Verzeichnung sowie Affinität und Scherung mitgeschätzt. Der komplette Bilddatensatz umfasst ca. 2200 Bilder und wird nachfolgend als „erweiterter“ Datensatz bezeichnet. Ein reduzierter Datensatz mit etwa 380 Bildern stellt den Standarddatensatz dar.

Als Maß zur Beurteilung der erzielten photogrammetrischen Ergebnisse dienen Längenmessabweichungen (LME, nach VDI/VDE 2002). Diese wurden anhand von Soll-Distanzen ermittelt, die mittels interferometrischer Lasertrackermessung zu 16 im Messvolumen verteilten Objektpunkten (vgl. Abb. 1, Passpunkte) mit einer Genauigkeit von 10 µm bestimmt wurden. Die verschiedenen Untersuchungsszenarien sind in Tab. 1 zusammengefasst. Für die nachfolgenden, weiterführenden Analysen werden u.a. die Konfigurationen E bis I näher untersucht. Es hat sich gezeigt, dass die Einbringung eines kurzen Maßstabs zu höheren Abweichungen und systematischen Maßstabseffekten im vorliegenden Testszenario führen.

Tab. 1 : Untersuchungsszenarien

	Symbol	Anzahl Maßstäbe	Länge [m]	Std.abw. Maßstäbe [µm]		Symbol	Anzahl Maßstäbe	Länge [m]	Std.abw. Maßstäbe [µm]
A		116	0.15–2.2	10 ⁶	F		1	1.4	10
B		116	0.15–2.2	10	G		1	2.0	10
C		1	21.0	20	H		1	2.2	10
D		5	1.4–2.2	10	I		1	1.5	10
E		1	1.6	10	J		0	–	20

Aufgrund der weiterführenden Analyse der im System verbleibenden Korrelationen wurde ein internes Bündelausgleichungsprogramm erweitert und eingesetzt. Für die Bündelausgleichung werden als Eingangsgrößen die Ergebnisse der Bündelausgleichung mit dem AICON 3D Studio bzw. AXIOS Ax.Ori verwendet. Die Ausgleichungsverfahren basieren alle auf der Ausgleichung nach vermittelnden Beobachtungen. Neben geringfügigen Abweichungen in den Ausgleichungsergebnissen, die unterschiedlichen Verfahren der Ausreißerelimination zugeschrieben werden können, sind die Ergebnisse direkt vergleichbar und übertragbar.

2.3 Ergebnisse

In Tab. 2 werden die erzielten Ergebnisse der Voruntersuchungen (REZNICEK et al. 2016) für die in Tab. 1 aufgeführten Konfigurationen zusammengefasst. Betrachtet werden dabei die maximale Längenmessabweichung (Max LME) sowie der RMS LME aller Längenmessabweichungen einer Konfiguration. Die geringsten Abweichungen sind in Blau markiert, die höchsten in Rot.

Tab. 2: Ergebnisübersicht der verschiedenen Konfigurationen

[μm]	A 	B 	C 	D 	E 
Max LME (Standardsatz)	-389	-236	-217	293	551
Max LME (erweiterter Satz)	-251	-164	-155	204	289
RMS LME (Standardsatz)	125	69	62	111	227
RMS LME (erweiterter Satz)	99	66	62	93	133
mittlere Anzahl Strahlen (Standardsatz / erweiterter Satz)	35/166	35/166	25/129	44/208	53/244
[μm]	F 	G 	H 	I 	J 
Max LME (Standardsatz)	201	318	-197	679	119
Max LME (erweiterter Satz)	226	193	164	223	114
RMS LME (Standardsatz)	70	121	57	287	47
RMS LME (erweiterter Satz)	105	87	70	104	44
mittlere Anzahl Strahlen (Standardsatz / erweiterter Satz)	48/230	43/195	42/210	32/162	-

3 Analysen

Die Ergebnisse der Voruntersuchungen (REZNICEK et al. 2016) zeigen das erzielte Genauigkeitsniveau des vorliegenden Testszenarios. Die unterschiedlichen Konfigurationen verweisen jedoch auch auf Auffälligkeiten und Restsystematiken in den Ausgleichungsergebnissen, die sich allein in der Betrachtung der Längenmessabweichungen nicht wiederfinden. Die Standardabweichungen der Objektpunkte resultieren in ihren Z-Komponenten so, dass minimale Standardabweichungen bei ca. 2/3 des maximalen radialen Abstands zum Schwerpunkt auftreten. Weiterhin ist zu untersuchen, inwiefern die Wahl der Datumsfestlegung Einfluss auf das Bündelausgleichungsergebnis hat. Je nach gewählter Aufnahmekonfiguration ergeben sich erhebliche Korrelationen zwischen den Prüflängen. Es ist zu untersuchen, inwiefern sich die benannten Effekte begründen lassen.

3.1 Analyse der Standardabweichungen der Objektpunkte

Die Standardabweichungen der Objektpunkte einer Bündelausgleichung hängen maßgeblich von der a priori Bildmessgenauigkeit, der Maßstabsdefinition und dem funktionalen Zusammenhang

der zu bestimmenden Parameter ab. Die theoretisch zu erwartende Kurve für die Standardabweichungen im Objektraum, basierend auf Erfahrungswerten, zeigt Abb. 2 in Rot. Die Standardabweichungen sollten aufgrund der geringer werdenden Dichte der Beobachtungen zum Testfeldrand steigen, als erwartete Höhe der Abweichungen wird das mittlere Minimum der Abweichungen angesetzt. Abb. 2 zeigt ergänzend einen typischen Verlauf der Standardabweichungen (Blau) für das vorliegende Testszenario.

Zunächst werden für die Analyse die Standardabweichungen der Objektpunkte in der Z-Komponente in Abhängigkeit zum Schwerpunkt des quadratischen Testfeldes (Mitte) dargestellt. Da insbesondere Auffälligkeiten in der Z-Komponente zu erkennen sind, wird die nähere Analyse in X- und Y-Richtung in diesem Zusammenhang vernachlässigt.

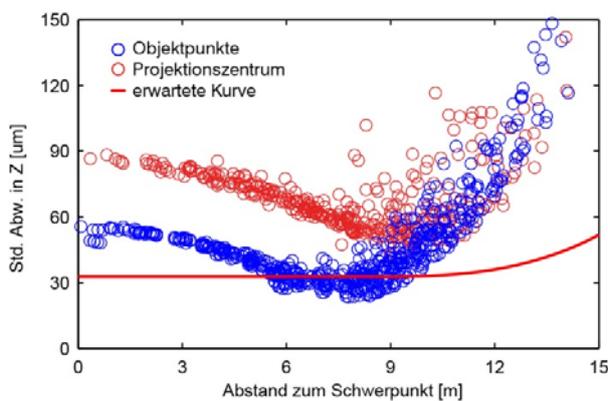


Abb. 2: Abhängigkeiten der Standardabweichungen der Z-Komponenten der Objektpunkte sowie der Projektionszentren vom Schwerpunkt des Messvolumens

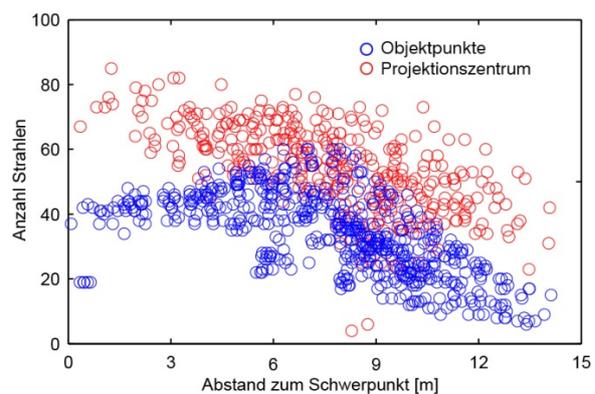


Abb. 3: Anzahl der Strahlen pro Objektpunkt und Projektionszentrum in Abhängigkeit vom Abstand zum Schwerpunkt

In der Z-Komponente stellt sich dabei in Abb. 2 eine ungewöhnliche Systematik dar. Wider Erwarten sind die Standardabweichungen der Objektpunkte (blaue Kreise) in der Mitte des Testfeldes höher als bei einer Entfernung von etwa 8 Meter zum Schwerpunkt. Ein ähnlicher Effekt kann bei den Standardabweichungen in den Z-Komponenten der Projektionszentren (rote Kreise) beobachtet werden. Diese Charakteristiken, welche in allen getesteten Konfigurationen, außer Konfiguration J (Einführung von Passpunkten, vgl. Tab. 1), zu sehen sind, lassen auf eine hohe Korrelation der Parameter schließen.

Zur Analyse dieses Effektes wird Konfiguration C (Standarddatensatz) gewählt. Hier liegt die a priori Genauigkeit des langen Maßstabs bei $0,1 \mu\text{m}$, um den Einfluss einer Maßstabsungenauigkeit auszuschließen. Einer der größten Anteile an der Größe einer Standardabweichung ist die Anzahl der Bildstrahlen pro Punkt. Um diesen Einfluss näher zu untersuchen, wurde in Abb. 3 die Anzahl der Strahlen pro Objektpunkt und Projektionszentrum in Abhängigkeit vom Abstand zum Schwerpunkt aufgetragen. Signifikante Korrelationen können zwischen der Strahlendichte und den Standardabweichungen der Objektpunkte (vgl. Abb. 2) beobachtet werden, wo die höchste Dichte nahe 8m liegt. Dies ist jedoch nicht der Fall zwischen der Strahlendichte und den Projektionszentren. Dies lässt darauf schließen, dass die Strahlendichte nicht die beobachteten

Systematiken in den Standardabweichungen verursacht, da für Objektpunkte und Projektionszentren die gleichen Auffälligkeiten (vgl. Abb. 2) vorherrschen.

Durch weitere Untersuchungen des beobachteten Effekts wurde der Einfluss der Verzeichnungsparameter auf die Standardabweichung in der Z-Komponente analysiert. Durch eine Deaktivierung der radial-symmetrischen Verzeichnungsparameter (A1, A2, A3) während der Simultankalibrierung wird ein großer Teil des beobachteten Effekts eliminiert (Abb. 4).

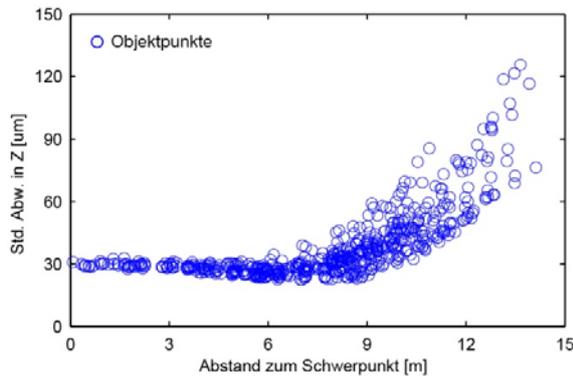


Abb. 4: Abhängigkeit der Standardabweichungen der Objektpunkte in Z vom Abstand zum Schwerpunkt unter Berücksichtigung der Deaktivierung der radial-symmetrischen Verzeichnungsparameter (A1 bis A3) im Ausgleichsprozess

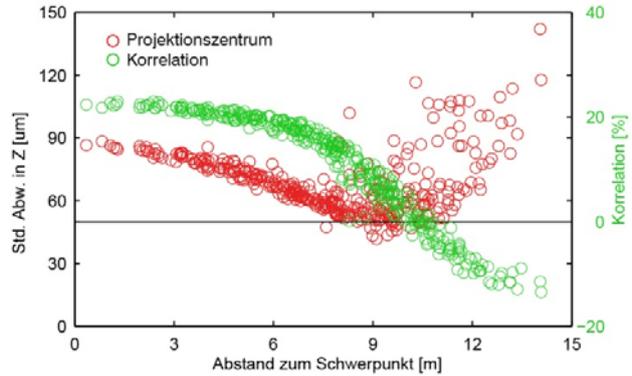


Abb. 5: Standardabweichungen der Projektionszentren in Z (rot) und deren Korrelation mit dem Verzeichnungsparameter A1 (grün)

Darüber hinaus ist festzustellen, dass der Parameter A1 eine erhöhte Korrelation zur Z-Komponente der Projektionszentren aufweist und vom Abstand zum Schwerpunkt abhängig ist. Der Verlauf der Kurven (Abb. 5) könnte den Effekt in der Z-Komponente der Standardabweichungen in den Objektpunkten erklären, ein höherer Korrelationskoeffizient führt zu einer höheren Standardabweichung.

Um die experimentell ermittelten Ergebnisse zu verifizieren wurde die gleiche Analyse an einem simulierten Datensatz durchgeführt, um unbekannte Einflüsse und Fehler auszuschließen. Die Konfiguration dieser simulierten Messung entspricht dem Aufbau der realen Messung. Die Objektpunkte liegen rasterförmig in einem Abstand von 1m in der XY-Ebene. Die Anzahl der Objektpunkte beträgt 360 (20 × 18). Die Anzahl und die Anordnung der Kamerastandpunkte entsprechen denen der Objektpunkte, jedoch mit einer Z-Koordinate von 7 m. Je Kamerastandpunkt werden vier um 90° gekantete Aufnahmen simuliert, was eine Gesamtanzahl von 1440 simulierten Bildern ergibt. Die Objektkoordinaten sowie die Kamerastandpunkte wurden zufällig verrauscht (± 50 mm für die Koordinaten und 0,01 rad für die Drehwinkel). Zur Berechnung der Bildkoordinaten sind die Parameter der inneren Orientierung aus dem praktischen Testszenario zum Einsatz gekommen. Die Bildkoordinaten unterliegen einem normalverteilten Rauschen von 0.5 μm .

Im Gegensatz zur realen Messung basieren die simulierten Daten auf einem planaren Testfeld ohne Hochpunkte, welche jedoch für eine gute Bestimmung der Kamerakonstante notwendig wären. Diese Konfiguration ist gewählt worden, damit vor allem die bisher beobachteten Systematiken hervorgehoben werden und mögliche Effekte der Hochpunkte auszuschließen sind. Es ist zu beachten, dass im praktischen Testszenario die Hochpunkte genau wie die Objektpunkte mit minimalen Standardabweichungen um etwa 8 m vom Schwerpunkt des Testfelds entfernt positioniert sind. Aufgrund des gewählten Simulationsszenarios wird deshalb zunächst auf eine simultane Schätzung der Kamerakonstante in der Prozessierung verzichtet. Ihre Berücksichtigung ergab keinen erkennbaren Einfluss. Alle weiteren Prozessierungen wurden dem des praktischen Testszenarios angepasst. Die Ergebnisse sind in Abb. 6 dargestellt.

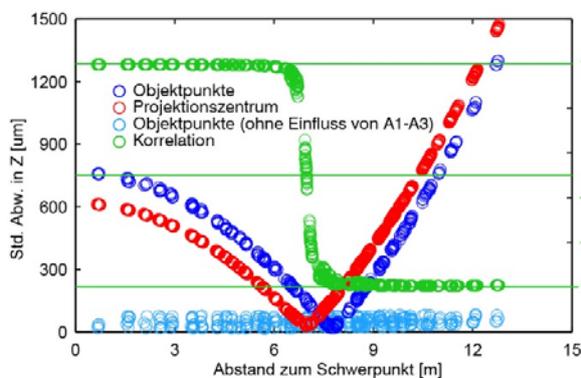


Abb. 6: Abhängigkeit der Standardabweichungen der Objektpunkte (dunkelblau, hellblau) sowie der Projektionszentren (rot) vom Abstand zum Schwerpunkt, Korrelation der Projektionszentren zu A1 (grün)

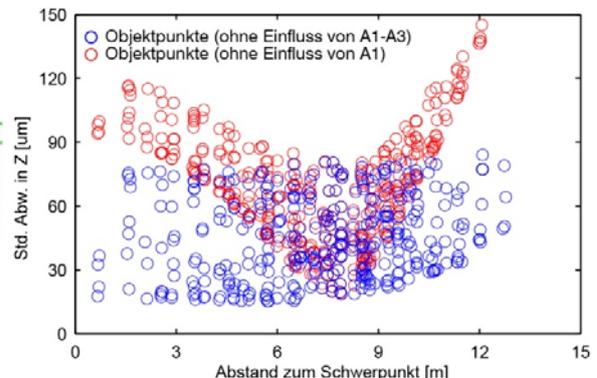


Abb. 7: Abhängigkeit der Standardabweichungen der Objektpunkte in Z vom Abstand zum Schwerpunkt (blau: ohne Einfluss von A1-A3, rot: ohne Einfluss von A1)

Die Graphik in Abb. 6 zeigt ähnliche Effekte wie die Ergebnisse der realen Messung. Die Kurve der Standardabweichungen der Objektpunkte (dunkelblau) nimmt bis zu einem bestimmten Abstand zum Schwerpunkt (etwa 8 m) ab und steigt danach wieder an. Ein annähernd gleicher Verlauf kann für die Standardabweichungen der Projektionszentren beobachtet werden (rot). Die grüne Kurve kennzeichnet die Korrelation zwischen den Standardabweichungen der Projektionszentren und des Verzeichnungsparameters A1. Der Kurvenverlauf zeigt eindeutig den Zusammenhang, dass die geringste Korrelation (0%) und die niedrigste Standardabweichung der Z-Komponenten im Objektraum im gleichen radialen Abstand zum Schwerpunkt vorliegen. Die in hellblau abgetragenen Daten zeigen die Standardabweichungen der Objektpunkte, wenn die Verzeichnungsparameter A1 bis A3 bei der Simultankalibrierung deaktiviert und dementsprechend nicht bestimmt werden. Wie in Abb. 6 zu sehen, ist der untersuchte systematische Effekt eliminiert und erhöht gleichzeitig die Präzision in der Z-Komponente der Objektpunkte. Gegenübergestellt werden in Abb. 7 die Standardabweichungen der Objektpunkte aus Auswertungen ohne Einfluss von A1-A3 sowie ohne Einfluss von A1 in einer kleineren Achsskalierung. Verbleibende Systematiken zeigen, dass alle radial-symmetrischen

Verzeichnungsparameter einen Einfluss auf die Standardabweichung in der Z-Komponente der Objektkoordinaten haben. Basierend auf diesen Analysen kann darauf geschlossen werden, dass der beobachtete Effekt durch die simultane Schätzung der radial-symmetrischen Verzeichnisparameter hervorgerufen wird.

3.2 Einfluss der Datumsfestlegung

In diesem Kapitel wird der Einfluss verschiedener Datumsfestlegungen auf die Genauigkeit der resultierenden Objektpunktkoordinaten untersucht. Für diese Untersuchungen wurde die Konfiguration C (Standarddatensatz) gewählt. Die a priori Genauigkeit der langen Maßverkörperung liegt hier bei 100 μm , um einen deutlichen Effekt durch eine ungenaue Skalierung hervorzurufen. Insgesamt werden drei verschiedene Datumsdefinitionen untersucht, bei denen die Lagerung auf

- I. allen Objektpunkten,
- II. drei Objektpunkten in einer Ecke des Testfelds und
- III. den vier äußeren Eckpunkten

definiert ist (Abb. 8). Datum II. und III. wurden so gewählt, dass Unterschiede möglichst gut aufgedeckt und sichtbar gemacht werden können. Die Prozessierung erfolgt mit der bereits erwähnten eigenen Implementierung einer Bündelausgleichung. Im Folgenden werden die Ergebnisse der einzelnen Auswertungen vorgestellt.

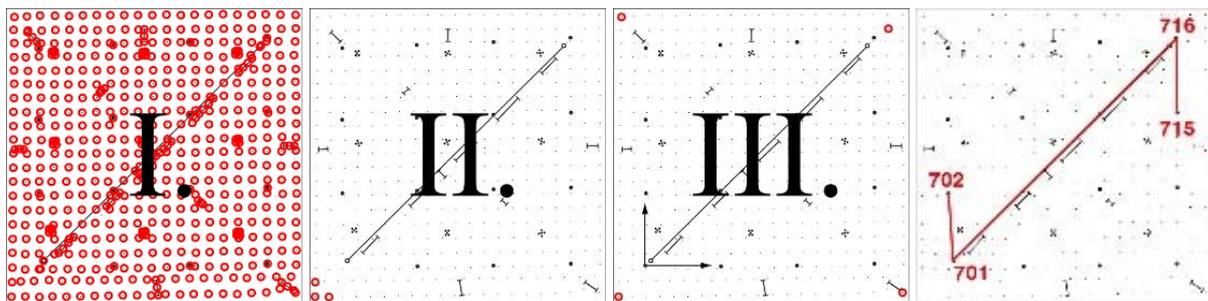


Abb. 8: von links: Schematische Darstellung der drei verschiedenen Datumsdefinitionen sowie die Lage der drei Prüfrecken. Die roten Kreise repräsentieren die Objektpunkte, welche als Datumspunkte in die Ausgleichung eingeführt werden.

Die Lage der Objektpunktkoordinaten der Definitionen II. und III. werden zunächst mit der Lage der Objektpunktkoordinaten von Definition I. durch Differenzenbildung verglichen. Die Datensätze II. und III. unterliegen dabei einer Best-fit-Transformation in das Koordinatensystem des ersten Datensatzes. Die Analyse der Koordinatendifferenzen zeigt einen vernachlässigbaren Einfluss der Datumsdefinition auf die absolute Lage der Objektpunktkoordinaten (vgl. Tab. 3).

Tab. 3: Statistische Analyse des Koordinatenvergleichs der Objektpunkte zwischen den Datensätzen (I.) - (II.) und (I.) - (III.).

[μm]	RMS		ABS MAX		Mittel	
	I.-II.	I.-III.	I.-II.	I.-III.	I.-II.	I.-III.
d_x	0.09	0.09	0.18	0.11	-0.08	-0.09
d_y	0.05	0.05	0.14	0.15	0.03	0.04
d_z	0.04	0.04	0.11	0.10	0.01	0.01
d_{xyz}	0.11	0.11	0.22	0.19	-	-

Tab. 4: Objektpunktgenauigkeit pro Datumsdefinition; Standardabweichungen der Objektpunkte s_x, s_y, s_z und Eigenwert der großen Halbachse s_e

[mm]	RMS			ABS MAX		
	I.	II.	III.	I.	II.	III.
s_x	0.03	0.29	0.04	0.07	0.56	0.07
s_y	0.03	0.55	0.04	0.07	0.99	0.07
s_z	0.05	2.22	0.11	0.14	4.03	0.14
s_e	0.05	2.28	0.11	0.14	4.12	0.14

Der Einfluss auf die erreichte a posteriori Genauigkeit kann durch Konfidenzellipsoide dargestellt werden, die Ausrichtung wird durch den Eigenwert der großen Halbachse des Ellipsoids bestimmt. Die maximalen statistischen Werte für die Standardabweichungen sowie die maximalen Eigenwerte pro Datensatz sind in Tab. 4 zusammengefasst. Generell sind hohe Genauigkeitsunterschiede feststellbar, die höchsten Abweichungen sind bei der zweiten Datumsdefinition zu sehen (vgl. Abb. 9). Sie werden durch hohe Korrelationen zwischen den Objektkoordinaten hervorgerufen. Der Einfluss dieser hohen Korrelationen soll anhand einer Streckenanalyse näher spezifiziert werden.

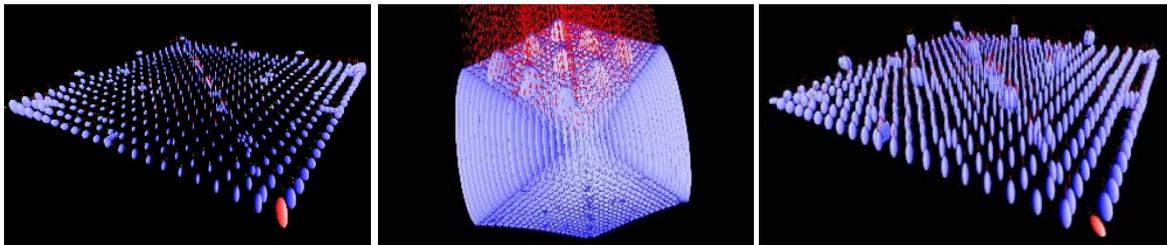


Abb. 9: Darstellung der Konfidenzellipsoide für die Datumsdefinitionen I. (links), II. (Mitte) und III. (rechts). Alle Graphiken weisen den gleichen Skalierungsfaktor für die Ellipsen auf. Die mittlere Graphik zeigt neben den extrem hohen Eigenwerten die Systematiken, die aufgrund von hohen Korrelationen zwischen den Objektkoordinaten hervorgerufen werden.

Für die Untersuchung des Einflusses der Datumsfestlegung auf die Prüflängen wurden drei Längen zwischen vier Objektpunkten ausgewählt, anhand derer der Einfluss der hohen Korrelationen zwischen den einzelnen Objektpunktkoordinaten gezeigt werden soll. Für die berechneten Längen wurde die komplette Kovarianzmatrix hinzugezogen. Die resultierenden Längen und deren Standardabweichungen sind jedoch bis auf einen numerischen Fehler identisch (vgl. Tab. 5).

Tab. 5: Standardabweichungen sowie Längen der ausgewählten Prüflängen

[μm]	Standardabweichungen der Prüflängen			Länge [mm]
	I.	II.	III.	
701–702	0.87	0.87	0.87	4111.11
701–716	8.99	8.99	8.99	21886.80
715–716	1.25	1.25	1.25	4431.60

Die Prozessierung der Datensätze wurde über eine Simultankalibrierung durchgeführt. Die ermittelten Parameter der inneren Orientierung einschließlich der Korrelationen zwischen diesen sind für alle Datensätze identisch. Für die Parameter der äußeren Orientierung trifft dies nicht zu, die Korrelationen zwischen diesen und den Parametern der inneren Orientierung sinken für den Datensatz II. signifikant.

Zusammenfassend ergibt sich, dass verschiedene Datumsdefinitionen nicht die relative Lage der Objektkoordinaten zueinander ändern, aber zu einer unterschiedlichen Genauigkeit a posteriori führen. Die Auswertung der vollständigen Kovarianzmatrix verweist darauf, dass die a posteriori Genauigkeit von Prüflängen für alle Datumsdefinitionen identisch ist. Insgesamt führt eine hohe Anzahl an Datumpunkten in einer guten Testfeldkonfiguration zu niedrigen Korrelationen. Höhere Korrelationen werden durch Unsicherheiten in der Datumsfestlegung verursacht.

3.3 Analyse von Korrelationen zwischen Referenzlängen

Im Rahmen der Ergebnisanalyse der verschiedenen Konfigurationen und insbesondere der damit einhergehenden unterschiedlichen Systemfestlegungen werden Korrelationen zwischen den zur Bestimmung der Längenmessabweichungen definierten Referenzlängen bestimmt und analysiert. Insgesamt liegen 60 Referenzlängen vor, die zwischen 16 rasterförmig angeordneten Objektpunkten aufgespannt sind. Zur Korrelationsbestimmung wird die Kovarianzmatrix der Strecken, bestimmt durch eine Varianz-Kovarianz-Fortpflanzung, der Bündelausgleichung hinzugezogen.

In der Betrachtung wird angenommen, dass eine höhere Korrelation zwischen Strecken vorliegen kann, wenn diese über mindestens einen gleichen Objektpunkt aufgespannt werden. Aus diesem Grund werden zur Analyse der in den verschiedenen Konfigurationen vorliegenden Korrelationen lediglich diejenigen näher betrachtet, bei denen die korrelierten Strecken jeweils zwischen vier unterschiedlichen Objektpunkten aufgespannt werden. Mit dieser Selektion ergeben sich 1408 zu betrachtende Streckenkombinationen. In Tab. 6 sind die Ergebnisse der Korrelationen der Streckenkombinationen aus den Referenzlängen zusammenfassend aufgeführt. Betrachtet wird jeweils die Anzahl der Streckenkombinationen, die eine Korrelation größer 75% bzw. >90% aufweisen.

Die Korrelationsergebnisse verdeutlichen, dass eine Konfiguration mit Passpunkten (J) neben der besten äußeren Genauigkeit, gemessen an den Längenmessabweichungen an den Referenzlängen, auch die geringsten Streckenkorrelationen mit maximal 25% aufweisen. Die geringen Korrelationen bei der Wahl einer langen, hochgenauen Maßverkörperung (C) festigen die Ergebnisse der Längenmessabweichungen und unterstützen die Schlussfolgerung, dass diese Konfiguration alternativ zu sehr guten und zuverlässigen Ergebnissen führt. Es ist jedoch zu bedenken, dass die Korrelationen in Abhängigkeit zur a priori Standardabweichung der Maßverkörperung zu betrachten sind. Werden die weiteren untersuchten Konfigurationen näher betrachtet, so kann festgestellt werden, dass die Wahl nur einer kurzen Maßverkörperung (~10% des Messvolumens) am Rande des Messvolumens dazu führt, dass alle vermeintlich unabhängigen Referenzlängen zu über 75% miteinander korreliert sind. Davon sind 64% der Streckenkombinationen zu >90% miteinander korreliert. Beim erweiterten Datensatz, d.h. unter Berücksichtigung zusätzlicher gekanteter und schräger Aufnahmen, steigt die Korrelation der

Streckenkombinationen. 98% aller möglichen Streckenkombinationen, die über verschiedene Objektpunkte aufgespannt werden, sind in diesem Fall zu >90% korreliert.

Tab. 6: Anzahl korrelierter Streckenkombinationen zwischen Referenzlängen (aufgespannt über vier unterschiedliche Objektpunkte)

Konfiguration		Korrelation		Konfiguration		Korrelation	
		>75%	>90%			>75%	>90%
B standard		181	0	E standard		1408	904
C standard (a priori 20µm)		2	0	E erweitert		1408	1376
C standard (a priori 100µm)		342	11	G standard		1386	590
C erweitert (a priori 20µm)		3	0	I standard		1408	1324
D standard		423	27	J standard		0	0
D erweitert		895	176				

Die graphische Auswertung der korrelierten Streckenkombinationen zeigt, dass kein eindeutiges Korrelationsmuster zu erkennen ist. Es kann festgestellt werden, dass sich die korrelierten Streckenkombinationen häufig kreuzen und vermehrt Streckenkombinationen hoch korreliert sind, die entweder in der Mitte des Messvolumens angeordnet sind oder dieses abdecken. Es kann bei Konfiguration C eine leicht homogenere Anordnung der korrelierten Streckenkombinationen im Messvolumen beobachtet werden. Bei der Betrachtung der erweiterten Datensätze mit einem kurzen Maßstab lässt sich erkennen, dass die innere Genauigkeit des Systems bei gleichbleibender Genauigkeit des Maßstabs steigt. Es resultieren jedoch deutlich höhere Korrelationen zwischen den Prüflängen. Die Anzahl der Beobachtungen für die Maßstabspunkte übt diesbezüglich keinen besonderen Einfluss auf die Ergebnisse aus. Wird die Anzahl der Beobachtungen im erweiterten Datensatz soweit reduziert, dass sie der des Standarddatensatzes entspricht, so erhöht sich der Korrelationsanteil weiter.

Insgesamt führen die Analysen zu dem Ergebnis, dass ein kurzer Maßstab einen hohen Einfluss auf das Bildbündel ausübt und dieser das Messvolumen nicht ausreichend repräsentiert. Die Genauigkeit des Maßstabs ist in diesen Fällen (Konfiguration E-I) nicht ausreichend.

4 Fazit & Ausblick

Die statistische Analyse der Bündelausgleichungsergebnisse eines großvolumigen Testszenarios ergab bei unterschiedlichen Maßstabfestlegungen besondere Effekte im Objektraum. Eine eingehende Untersuchung der Standardabweichungen der Objektpunkte deckt eine von simultan geschätzten radial-symmetrischen Verzeichnungsparametern beeinflusste Bestimmung der Z-Komponenten auf. Minimale Standardabweichungen resultieren bei Simultankalibrierung der radial-symmetrischen Verzeichnung sowohl für die Objektpunkte als auch die Projektionszentren

in Z bei ca. $2/3$ des maximalen radialen Abstands zum Systemschwerpunkt. Ein Zusammenhang zur Anzahl der Bildstrahlen kann ausgeschlossen werden. Die Untersuchungen zum Einfluss der Datumsdefinition konnten keine relativen Lageänderungen der Objektpunktkoordinaten verschiedener Lagerungen aufdecken, jedoch wird eine schwankende a posteriori Genauigkeit der Objektpunktkoordinaten ermittelt. Insgesamt führt eine hohe Anzahl Datumpunkte bei einer guten Testfeldkonfiguration zu niedrigen Korrelationswerten. Im vorliegenden Testszenario konnten des Weiteren hohe Korrelationen in Abhängigkeit von der Maßstabsfestlegung aufgedeckt werden. Ein kurzer Maßstab im Objektraum verbunden mit einem hochredundanten Bündel führt zwar zu besseren Genauigkeiten im Objektraum, repräsentiert jedoch nicht das komplette Messvolumen, da erhebliche Korrelationen zwischen den Prüflängen vorliegen.

Zusammenfassend lässt sich sagen, dass die Wahl der Maßverkörperung, deren Länge und Lage einen erheblichen Einfluss auf die Ergebnisse in großvolumigen Körpern ausübt. Die Ergebnisse der vorliegenden Untersuchungen zeigen, dass die radial-symmetrische Verzeichnung im vorliegenden Experiment die innere Genauigkeit des Systems maßgeblich bestimmt. Die Untersuchungsergebnisse führen zu der weiterführenden Fragestellung, inwieweit der Einfluss der radial-symmetrischen Verzeichnung und die Bestimmbarkeit der Kamerakonstante näher spezifiziert werden können.

5 Danksagung

Wir danken Ralph Rosenbauer und der ALPA Ltd. für die Bereitstellung der ALPA metric Kamera. Die vorliegenden Arbeiten wurden unter anderem durch die Forschungsprofessur des Landes Niedersachsen unterstützt.

6 Literaturverzeichnis

- KAHMEN, O., 2015: Entwicklung einer großen Invardraht-Maßverkörperung zur Anwendung in der Industriephotogrammetrie. Bachelorarbeit Jade Hochschule Oldenburg, unveröffentlicht.
- KAHMEN, O. & LUHMANN, T., 2016: Entwicklung einer großen Invardraht-Maßverkörperung zur Anwendung in der Industriephotogrammetrie. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 25, T. Kersten (Hrsg.), Dreiländertagung der DGPF, der OVG und der SGPF in Bern, 7.-9. Juni 2016, 538-544.
- REZNICEK, J., EKKEL, T., HASTEDT, H., LUHMANN, T., KAHMEN, O. & JEPPIING, C., 2016: Zum Einfluss von Maßstäben in photogrammetrischen Projekten großer Volumina. Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage, T. Luhmann & C. Schumacher (Hrsg.), VDE Verlag GmbH, Berlin/Offenbach, 286-295.
- RIEKE-ZAPP, D., OLDANI, A. & PEIPE, J., 2005: Eine neue, hochauflösende Mittelformatkamera für die digitale Nahbereichsphotogrammetrie. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 14, E. Seyfert (Hrsg.), 263-270.
- VDI/VDE, 2002: Optische 3-D-Messsysteme - Bildgebende Systeme mit flächenhafter Antastung. VDI/VDE Richtlinie 2634, Blatt 1, Beuth Verlag, Berlin.

Genauigkeitsuntersuchungen handgeführter Scannersysteme

THOMAS P. KERSTEN¹, HEINZ-JÜRGEN PRZYBILLA², MAREN LINDSTAEDT¹,
FELIX TSCHIRSCHWITZ¹ & MARTIN MISGAISKI-HASS³

Zusammenfassung: Handgeführte Scannersysteme verschiedener Hersteller sind in zunehmendem Maße am Markt verfügbar, jedoch ist über deren geometrische Leistungsfähigkeit bei vielen Anwendern wenig bekannt. Daher hat das Labor für Photogrammetrie & Laserscanning der HafenCity Universität Hamburg in Zusammenarbeit mit der Hochschule Bochum (Labor für Photogrammetrie) sowie der Humboldt-Universität zu Berlin (Institut für Informatik) geometrische Genauigkeitsuntersuchungen mit folgenden Systemen durchgeführt: DOTProduct DPI-7, Artec Spider, Mantis Vision F5 SR, Kinect v1 + v2, Structure Sensor und Google's Project Tango. Im Rahmen dieser vergleichenden Untersuchungen wurden geometrisch stabile Referenzkörper eingesetzt. Die zugehörigen Referenzdaten wurden durch Messung mit zwei Streifenprojektionssystemen (smartSCAN und GOM ATOS I 2M) erfasst. Die umfassenden Untersuchungsergebnisse der verschiedenen Testszenerien werden in diesem Beitrag vorgestellt und kritisch diskutiert.

1 Einleitung

In den letzten Jahren wurde der Markt zunehmend mit sogenannten handgeführten 3D-Scannern im unteren bis mittleren Preissegment (EUR 500 bis 20.000) erweitert. Die typischen Anwendungsbereiche dieser 3D-Scanner beschränken sich zumeist auf den Nahbereich, d. h. für Messaufgaben mit Distanzen unter einem Meter bis zu einigen wenigen Metern. Aufgrund der aktuellen technologischen Vielfalt im Bereich der 3D-Scanner ist es jedoch eine Herausforderung für eine konkrete Anwendung ein geeignetes 3D-Scanningsystem zu wählen. TALLIG et al. (2015) stellen eine Konzeption und Verifikation für Auswahlverfahren von 3D-Scantechnologien mit einem vielseitigen Kriterienkatalog vor.

Durch den günstigen Preis und die teilweise einfache Handhabung stellen diese handgeführten Scanner ggf. eine signifikante Konkurrenz zu den teuren und präzisen Streifenlichtprojektionssystemen dar. Es stellt sich daher die Frage, wie genau diese 3D-Scanner im Vergleich zu klassischen Streifenlichtprojektionssystemen sind (z. B. der Hersteller GOM, Steinbichler, AICON u. a.) und welche Qualität der Anwender für die aufgenommenen 3D-Daten erwarten kann. Dazu gibt es in der Literatur bereits erste veröffentlichte Ergebnisse, z. B. von Systemen aus dem Spielbereich sowie von sogenannten low-cost Systemen (Streifenprojektionssystem David SLS-1 und Kinect v1/ReconstructMe) zur 3D-Rekonstruktion kleiner Objekte (HIERONYMUS et al. 2011 und OMELANOWSKY et al. 2013, BÖHM 2014). Hier hat sich wie erwartet gezeigt, dass weder Stabilität noch die Qualität dieser Systeme mit den High-End Systemen derzeit mithalten können.

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

² Hochschule Bochum, Fachbereich Geodäsie, Labor für Photogrammetrie, Lennershofstraße 140, D-44801 Bochum, E-Mail: heinz-juergen.przybilla@hs-bochum.de

³ Humboldt-Universität zu Berlin, Institut für Informatik, Rudower Chaussee 25, D-12489 Berlin-Adlershof, E-Mail: martin.misgaiski@informatik.hu-berlin.de

Im nachfolgenden Beitrag werden geometrische Genauigkeitsuntersuchungen verschiedener handgeführter 3D-Scanner vorgestellt. Hierzu wurden verschiedene Referenzkörper eingesetzt, deren Datensätze aus Messungen mit High-End Streifenprojektionssystemen (AICON 3D Systems smartSCAN und GOM ATOS I 2M) stammen.

2 Referenzkörper

Für den Vergleichstest wurden folgende Objekte genutzt (Abb. 1): eine Büste von Einstein aus Gips (Höhe 160 mm), eine Radnabe aus Gusseisen mit den Dimensionen $232 \times 120 \times 232 \text{ mm}^3$ und vier sogenannte „Testys“ (Höhe 380 mm) des Instituts für Informatik der Humboldt-Universität zu Berlin (REULKE & MISGAISKI 2012). Weitere Prüfungen erfolgten mit den folgenden geometrisch stabilen Referenzkörpern der Hochschule Bochum: ein kreuzförmiger Körper mit Kugeln aus Stahl (max. Abstand 450 mm von fünf Kugeln mit einem Durchmesser von 65 mm) und eine ebene Granitplatte (Größe $300 \times 300 \text{ mm}^2$). Als Referenzobjekt für größere Flächenaufnahmen diente die Wand des 3D-Testfeldes der HCU Hamburg, die mit einem terrestrischen Laserscanner IMAGER 5010 der Firma Zoller+Fröhlich mit der Auflösung high (6 mm @ 10 m) von drei Standpunkten erfasst wurde.



Abb. 1: Referenzkörper für die Untersuchungen der handgeführten Scannersysteme – v.l.n.r. Einstein-Büste, Radnabe, Testy, Testkörper HSBO mit Kugeln und Granitplatte.

3 Untersuchte Scannersysteme

Folgende handgeführte Scannersysteme standen zur Verfügung (Abb. 2), von denen ausgewählte technische Daten in Tabelle 1 zusammengefasst sind: drei DotProduct DPI-7 (Landeskriminalamt Hamburg (LKA), Hochschule Bochum (HSBO) und Dr. Hesse und Partner Ingenieure (dhp:i)), Artec Spider (LKA), Mantis Vision F5 Short Range (MexConsult), Kinect v1 und v2 (HCU Hamburg), Structure Sensor (Humboldt-Universität zu Berlin) und Google's Project Tango (dhp:i). Außerdem wurden alle Objekte mit einer digitalen Spiegelreflexkamera Nikon D800 und einem Smartphone Galaxy Note 1 aufgenommen.



Abb. 2: Untersuchte handgeführte Scannersysteme (v.l.n.r.): DPI-7, Artec Spider, Mantis Vision F5, Kinect v2 (oben), Kinect v1, Structure Sensor und Google's Project Tango (unten).

Tab. 1: Ausgewählte technische Daten der untersuchten 3D-Scanner (Herstellerangaben)

System	Messverfahren	Messbereich [m]	Genauigkeit [mm]	Gewicht [kg]	Preis [EUR]
DPI-7 SR (dhp:i)	Musterprojektion	0,60 – 1,20	2 (auf 1m)	< 1,00	5.000
DPI-7 LR (LKA/HSBO)	Musterprojektion	0,60 – 3,30	2 (auf 1m)	< 1,00	5.000
Artec Spider	Musterprojektion	0,17 – 0,35	0,05	< 1,00	15.700
Mantis F5 SR	Musterprojektion	0,30 – 0,80	0,05 (auf 50cm)	1,70	15.000
Structure Sensor	Musterprojektion	0,40 – 3,50	0,5 (auf 40cm)	0,10	450
Kinect v1	Musterprojektion	0,80 – 4,00	5 (auf 1,5m)	1,36	100
Kinect v2	Time-of-flight	0,40 – 4,50	keine Info	2,04	200
Google's Project Tango	Time-of-flight	0,50 – 4,00	keine Info	0,37	450
ATOS I 2M	Streifenlicht	0,16 – 1,28	0,02	3,50	50.000
smartSCAN	Streifenlicht	0,03 – 1,50	0,009 (Fläche)	4,00	80.000

3.1 DotProduct DPI-7

Wesentlicher Bestandteil der Hardware des DPI-7 Scanners (DotProduct, USA) ist ein PrimeSense Sensor, wie er auch in der Kinect v1 verbaut ist (NIR-Projektor sowie NIR- und RGB-Kamera). Nach einem Kaltstart benötigt das System ca. 20 Minuten Aufwärmzeit. Die Steuerung des Sensors erfolgt über ein angeschlossenes Android-Tablet mit der Software Phi.3D. Zur Ausrichtung der Punktwolke werden die internen Beschleunigungsmesser und Gyroskope des Tablets genutzt. Bei ausreichender Überdeckung (Prüfung über visuelle Farbinformation am Tablet) erfolgt eine Vorregistrierung durch einen ICP-Algorithmus (BESL & MCKAY 1992). Nach Abschluss des Scanvorganges schließt sich eine Optimierung der Registrierung an, bei der auch fehlerhafte Punkte (z. B. Mixed Pixel) eliminiert werden. Eine Untersuchung zum DPI-7 findet sich in JAHRAUS et al. (2015), eine Anwendung in AHERN & SPRING (2015).

3.2 Artec Spider

Artec Spider (Artec 3D, Luxemburg) ist ein handgeführter 3D-Scanner, der speziell für CAD-Anwender entwickelt wurde, um kleine Gegenstände mit komplexer Oberflächenstruktur, scharfen Kanten und dünnen Rippen mit 7,5 Bildern oder mit 1,0 Mio. Punkten pro Sekunde zu scannen. Das System benötigt eine Aufwärmphase von ca. 30 Minuten und arbeitet mit einem linearen

Sichtfeld zwischen $90 \times 70 \text{ mm}^2$ bis $180 \times 140 \text{ mm}^2$. Der Aufnahmebereich liegt zwischen 0,17 - 0,35 m. Der Artec Spider nutzt eine Musterprojektion mit blauen LEDs als Lichtquelle und eine Farbkamera mit 1,3 Megapixel (24-bit Radiometrie) für die Texturierung der Objekte. In Verbindung mit der Artec Studio Software ist er (laut Hersteller) „ein Desktoptool für Designer, Ingenieure und Erfinder jeder Art“. Anwendungsbeispiele für das System finden sich bei ADAMS et al. (2015), FRIEDMAN et al. (2015) sowie INZERILLO et al. (2015).

3.3 Mantis Vision F5 Short Range

Der Mantis Vision F5 (Israel) ist ein auf Musterprojektion basierender handgeführter Scanner mit einem Messbereich von 0,5 m - 4,5 m (MVC F5) bzw. 0,3 m - 0,8 m (MV F5 Short Range). Die Hardware des Sensors besteht aus zwei Modulen: einer Videokamera und einem Projektor, welche in den Handgriff eingelassen sind. Der Projektor emittiert infrarotes Licht (firmeneigenes Muster) auf das Objekt, welches als kodiertes Licht von der Videokamera erfasst wird. Der Triangulationsalgorithmus berechnet daraus eine Punktwolke (500.000 Punkte/s). Die Punktdichte in XY beträgt für ein einzelnes Bild 1,6 mm bei einer Entfernung von 0,5 m. Durch die geringe Sensitivität gegenüber dem Umgebungslicht ist das System sowohl bei Dunkelheit als auch bei Tageslicht einsetzbar. WRONA (2014) und ZHANG et al. (2015) beschreiben verschiedene Anwendungen für den Scanner.

3.4 Structure Sensor

Der Structure Sensor wurde von der Firma Occipital (USA) entwickelt und wird heute auch als iSense™ 3D-Scanner für iPad vermarktet. Der Sensor wird an einem iPad oder iPhone 6 (Structure App) adaptiert, welches mit einem PC (Mac OS X oder Windows) mit Nvidia Grafikkarte mit CUDA Funktionalität über WLAN kommuniziert. Die Datenprozessierung erfolgt auf dem PC mit der Software Skanect Pro. Skanect Pro bietet dabei Funktionen, wie z. B. Punktwolkenreduktion oder automatische Eliminierung von Artefakten. Der Sensor selbst besteht aus einem NIR-Laser-Projektor und einer NIR-Kamera (Auflösung in Pixel von 640×480 bei 30 Bildern pro Sekunde (fps) bzw. 320×240 mit 60 fps). Zusätzlich werden die RGB-Kamera und Lagesensorik des iPads verwendet. Letztere ermöglicht eine Positionungsverfolgung während des Scanvorgangs. Somit ist es möglich, texturierte und orientierte 3D-Scans zu erstellen. Der Aufnahmebereich beträgt 0,4 m bis 3,5 m, wobei die Tiefengenauigkeit mit steigender Entfernung abnimmt, bei 40 cm 0,15% (0,5 mm), bei 100 cm 0,3% und bei 350 cm 1,1% (35 mm). Sämtliche Daten werden in Echtzeit an Skanect Pro übertragen, dort verarbeitet und gespeichert.

3.5 Kinect v1 & Kinect v2

Die Kinect v1 wurde von Microsoft zusammen mit der Firma PrimeSense als Hardware zur Steuerung der Videospielekonsole Xbox 360 entwickelt und ist somit quasi einer der ersten am Markt verfügbaren low-cost 3D-Sensoren. Das Funktionsprinzip der Kinect v1 ist detailliert in WUJANZ et al. (2011) beschrieben. Genauigkeitsuntersuchungen der Kinect v1 finden sich in WUJANZ et al. (2011), KHOSHELHAM (2011), CHOW et al. (2012), SMISEK et al. (2013) sowie GONZALEZ-JORGE et al. (2013).

Die Kinect v2 besteht aus einer HD-Weitwinkelkamera mit einer Auflösung von 1920×1080 Pixel bei 30 fps und einem IR-Sensor mit 512×424 Pixel. Drei IR-Projektoren beleuchten aktiv die zu

beobachtende Szene. Die RGB-Kamera erfasst die Farbinformationen, während die IR-Kamera simultan sowohl eine Tiefenkarte als auch IR-Bilddaten in Echtzeit erfasst. LACHAT et al. (2015) stellen geometrische Genauigkeitsuntersuchungen der Kinect v2 vor.

3.6 Google's Project Tango

Google's Project Tango verbindet 3D Motion Tracking mit Tiefenmessungen, um dem mobilen Gerät die eigene Position und die Bewegung durch den Raum mitzuteilen. Es handelt sich also um ein Tablet, das als Multi-Sensor-System zur Navigation und Orientierung im Raum in Echtzeit agiert. Das Tablet besitzt eine RGB-D Kamera mit Fisheye-Objektiv (170° Sichtfeld) und Tiefensensor (Infineon 3D-Bildsensorchip IRS1645C mit 160 x 120 Pixeln) zur Erfassung von dichten Punktwolken. Als weitere Sensoren sind im System eine RGB-IR Kamera mit 4 MPixel, eine Frontkamera mit 1 MPixel und fixiertem Fokus, ein Beschleunigungsmesser, ein Barometer, ein Kompass, GPS und ein Gyroskop enthalten. Dazu kommen noch ein IR-Projektor (von Mantis Vision) und ein Blitzlicht. Mit dieser Hardware und entsprechend entwickelter Software wird die 3D-Bewegung durch die Motion-Tracking-Kamera verfolgt, laufend die Position und Orientierung bestimmt und eine Umgebungskarte aus 0,25 Million 3D-Messungen pro Sekunde erstellt. LOIANNO et al. (2015) zeigen, dass das Project Tango in mobilen Anwendungen (als Sensor auf einem UAV) eingesetzt wird. Den Einsatz für Aufnahmen von Innenräumen stellen WINTERHALTER et al. (2015) und KLINGENSMITH et al. (2015) vor.

3.7 Kameras – Nikon D800 und Smartphone Galaxy Note 1

Parallel zu den handgeführten 3D-Scannern wurden Aufnahmen mit einer Nikon D800 sowie dem Smartphone Galaxy Note 1 erstellt. Die Auswertung der Bildverbände erfolgte mit der SfM-Software Agisoft PhotoScan (Version 1.2.3). Die Nikon D800 ist eine DSLR-Kamera mit einer Sensorgröße (CMOS) von 35,9×24,0 mm² und 36,3 Mio. Pixeln. Die Kamera des Smartphones verfügt über einen deutlich kleineren Sensor mit vermutlich 4,5×3,4 mm² (8 MPixel) und einem Objektiv mit Festbrennweite von 3,97 mm. Beispiele für die Auswertung von Bildverbänden ähnlicher Objekte wurden bereits publiziert (KERSTEN & LINDSTAEDT 2012, OMELANOWSKY et al. 2013).

3.8 Referenzsysteme – GOM ATOS I 2M und smartSCAN 3D

Der ATOS (Advanced Topometric Sensor) I 2M der Gesellschaft für optische Messtechnik GmbH (GOM) aus Braunschweig ist ein Streifenprojektionssystem, das aus zwei CCD-Kameras mit der Auflösung 1624×1236 Pixel und einem Projektor besteht. Je nach verwendetem Objektiv können Messfelder mit 500×400 mm² oder 250×200 mm² eingesetzt werden. Der ATOS I 2M wurde in verschiedenen Anwendungen als Erfassungs- und Referenzsystem eingesetzt (KELLER et al. 2011, RAU & YEH 2012, KERSTEN & LINDSTAEDT 2012, OMELANOWSKY et al. 2013).

Der smartSCAN 3D der Firma AICON 3D Systems GmbH ist ein Streifenprojektionssystem (Weißlichtscanner), der nach dem kombinierten Gray-Code/Phase-Shift-Verfahren arbeitet. Die Systemkameras (hier mit einer Kameraauflösung von 5 Megapixeln) nehmen das projizierte Streifenmuster (Lichtquelle: LED weiß, alternativ grün, blau oder rot) unter einem vordefinierten Betrachtungswinkel in einer Messsequenz von etwa einer Sekunde auf. Der Scanner besitzt ein Messbereich, abhängig von der Basis zwischen den Kameras, von 30 mm bis 1500 mm.

Anwendungsbeispiele des smartSCAN präsentieren SLIZEWSKI et al. (2010) und BATHOW & BREUCKMANN (2011).

4 Aufnahme

Die Testmessungen mit den handgeführten Scannern fanden am 5. und 6. Januar 2016 an der HafenCity Universität Hamburg im Labor für Geomatik statt. An verschiedenen Arbeitsplätzen wurden die Referenzkörper durch die zu untersuchenden 3D-Scanner, die Kameras sowie die beiden Referenzsysteme erfasst. Vor der Aufnahme wurden die Kugeln des Testkörpers HSBO sowie die Radnabe mit Kalkspray eingesprüht, um aus den glänzenden Metallkörpern matte und damit diffus streuende Oberflächen zu erzeugen. Tabelle 2 zeigt die Aufnahmen der verschiedenen Objekte, sortiert nach eingesetzten Systemen.

Tab. 2: Zusammenstellung der durchgeführten Aufnahmen nach Systemen und Objekten

System	T1	T2	T3	T4	Einstein	Radnabe	Platte	RK HSBO	Testfeld
DPI-7 (dhp:i)	-	✓	-	-	✓	✓	✓	✓	-
DPI-7 (LKA)	✓	✓	-	-	✓	✓	✓	✓	-
DPI-7 (HSBO)	✓	✓	✓	✓	✓	✓	✓	✓	✓
Artec Spider	✓	-	-	-	✓	✓	✓	-	-
Mantis F5 SR	-	-	✓	-	✓	✓	-	✓	-
Structure Sensor	-	-	✓	-	✓	✓	✓	✓	-
Kinect v1	-	✓	✓	✓	✓	✓	✓	✓	-
Kinect v2	-	✓	✓	✓	✓	✓	✓	✓	✓
Google's Project Tango	-	✓	-	-	✓	✓	✓	✓	✓
Nikon D800	✓	✓	-	✓	✓	✓	✓	✓	✓
Galaxy Note 1	-	✓	-	✓	✓	✓	✓	✓	-
ATOS I 2M	✓	✓	✓	-	✓	✓	✓	✓	-
smartSCAN	✓	✓	✓	✓	✓	✓	✓	✓	-
T Testy, RK Referenzkörper									

Die handgeführten 3D-Scanner wurden während der Objektaufnahme in einer langsamen, gleichförmigen Bewegung um und über das Aufnahmeobjekt geführt. Dabei wurden die erfassten Daten bei fast allen Systemen in Echtzeit auf dem angeschlossenen Computer (Kinect, Artec Spider) oder auf dem Tablet (Structure Sensor, Google's Project Tango) angezeigt.

Bei den **DPI-7** sind zu schnelle oder abrupte Bewegungen bei der Datenerfassung zu vermeiden, da sie unweigerlich zu einem Signalabriss führen. In diesem Fall muss der Operateur das System über einem bereits hinreichend erfassten Bereich des Objektes neu orientieren. Die Erfahrung des Bedieners übt somit einen nicht unwesentlichen Einfluss auf Geschwindigkeit und Qualität der Datenerfassung aus. Der Status der Datenerfassung (z. B. Vollständigkeit) wird während der Messung (durch entsprechende Farbcodierung) direkt auf dem Bildschirm des angeschlossenen Tablets angezeigt.

Die Aufnahme erfolgt beim **Artec Spider** nach der Aufwärmzeit wie beim DPI-7, d. h. durch langsame, gleichförmige Bewegung um das Objekt wird ein Signalabriss verhindert. Die Datenerfassung dauerte mit diesem System zwischen 2 und 10 Minuten pro Objekt.

Beim **Mantis F5 Scanner** werden mögliche Signalabriss erst im Postprocessing erkannt, sie können jedoch durch die Registrierung behoben werden. Die Datenfrequenz liegt bei 10 fps, eine

typische Aufnahme von einer Minute führt also zu ca. 15 Mio. Punkten. Das System arbeitet ohne Passpunkte, doch es sollten im Aufnahmebereich immer genügend Objektelemente in alle drei Koordinatenrichtungen vorhanden sein.

Alle räumlichen Testobjekte konnten mit dem **Structure Sensor** durch langsame und gleichmäßige Handführung in etwa 1 Minute aufgenommen werden. Die Aufnahme der Granitplatte erfolgte dagegen in nur 15 Sekunden. Die Nachbearbeitung der erfassten Daten dauerte je nach Objekt auf dem verwendeten Macbook zwischen 10 und 15 Minuten. Darin enthalten waren die Segmentierung des Objektes, die Dreiecksvermaschung sowie die Texturierung.

Auch das **Project Tango** wurde während der Aufnahme langsam und gleichmäßig über das Objekt geführt. Dabei zeigte sich schnell, dass dieser Sensor wegen der großen Pixelgröße am Boden von ca. 3 cm für die Aufnahme der kleinen Referenzkörper nicht geeignet ist. Daher wurde mit diesem Sensor nur das 3D-Testfeld aufgenommen.

Die mit der **Nikon D800** und dem **Smartphone Galaxy Note 1** aufgenommenen Objekte wurden durch einen Rundum-Bildverband aus verschiedenen Höhen erfasst. Die Objektaufnahme mit der Nikon D800 (alle Objekte mit Brennweite 85 mm, Testfeld mit 24 mm, Sensorempfindlichkeit ISO 200, Blende 22) erfolgte wegen der langen Belichtungszeiten von einem Stativ, während das Note 1 aus der Hand frei bedient wurde. Bei beiden Kameras war der Fokus auf „automatisch“ eingestellt. Mit der Nikon wurden Bildverbände mit 40 bis 80 Fotos aufgenommen, die je nach Objekt zwischen 20 und 45 Minuten dauerten. Die Freihandaufnahmen mit dem Galaxy Note 1 (Bildverbände mit 28 bis 68 Fotos) erfolgten mit drei bis sechs Minuten deutlich schneller.

Zur Messung der Objekte mit den Streifenprojektionssystemen **ATOS I 2M** und **smartSCAN** wurden diese auf Drehtellern positioniert. Beide Systeme wurden zu Beginn der Messtage kalibriert. Alle Objekte wurden in mehreren Scans aufgenommen, die in der jeweiligen Systemsoftware automatisch registriert wurden und daher mit jeder weiteren Aufnahme das Modell des Objektes vervollständigten. Beim ATOS wurden für die jeweilige Registrierung des Folgescans Zielmarken verwendet, die auf dem Drehteller und teilweise auf dem Objekt aufgeklebt waren. Die Registrierengenauigkeit lag zwischen 0,023 mm und 0,041 mm. Die Anzahl der Einzelscans variierte von 24 (Granitplatte) bis 120 Aufnahmen (Radnabe). Die Verknüpfung der Scans des smartSCAN erfolgte unter Nutzung von „Matchinghilfen“ - kleinen geometrischen Körpern mit signifikanten Oberflächenstrukturen, insbesondere Kanten, die im Objektraum verteilt angeordnet wurden.

5 Auswertung und Ergebnisse

Für die Auswertung lagen die Daten der einzelnen Referenzkörper der verschiedenen Messsysteme in unterschiedlichen Formaten vor. Einige Systeme lieferten Punktwolken (DPI-7), andere generierten bereits 3D-Modelle durch Dreiecksvermaschung on the Fly (Project Tango, Kinect v1 / v2) oder nachträglich in der System-Software (Structure Sensor, Mantis F5, Artec Spider, ATOS, smartSCAN). Die Kameras lieferten Fotos als Bildverbände, aus denen später mittels Dense Image Matching Punktwolken und nachfolgend Dreiecksvermaschungen generiert wurden (Agisoft PhotoScan). Zur Skalierung der Punktwolken wurden in den Fotos drei kalibrierte Maßstäbe

(Länge 140 mm, 290 mm und 480 mm) gemessen und in den Bündelausgleichungen mit Genauigkeiten von 0,1 - 0,6 mm ausgeglichen.

Die Referenzkörper (Testy, Radnabe und Einstein-Büste) wurden mit den Streifenprojektionssystemen sehr detailliert und präzise aufgenommen sowie in Geomagic Studio 2012 modelliert. Für Testy 1, 2 und 3 dienten die durch das ATOS-System erfassten Daten als Referenzdatensatz, während bei Testy 4, der Radnabe sowie der Einsteinbüste die Daten des smartSCAN als Referenz fungierten.

Zur Abnahme und Überwachung optischer Messsysteme mit flächenhafter Antastung bildet die Richtlinie VDI/VDE 2634, Blatt 2 und 3, ein anerkanntes Regelwerk (VDI/VDE, 2002 & 2006). Im Rahmen fest definierter Prüfzenarien werden mittels geeigneter Testobjekte (Artefakte) verschiedene Kenngrößen zur Bewertung der Qualität eines Messsystems ermittelt bzw. bezüglich der Herstellerangaben überprüft. Ziel ist es dabei auch, über die Betriebsdauer eines Systems Aussagen zu dessen Messqualität machen zu können.

In Anlehnung an die Richtlinie erfolgten die Prüfungen mit dem Kugeltestkörper der HSBO sowie mit der Granitplatte. Die im Bewertungsprozess ableitbaren Kenngrößen sind:

- *Antastabweichung PS (Maß)* – die Differenz aus gemessenem Durchmesser und kalibriertem Durchmesser der Kugel.
- *Antastabweichung PF (Form)* – die Spanne der radialen Abstände der Messpunkte von einer berechneten Ausgleichskugel. Die Ausgleichskugel wird nach der Methode der kleinsten Fehlerquadratsumme bei freiem Radius bestimmt.
- *Kugelabstandsabweichung SD* – Differenz zwischen dem gemessenen Wert und dem kalibrierten Wert eines Abstands zweier Kugelmittelpunkte. Anders als in der VDI/VDE 2634 Blatt 2 wird der gemessene Abstand aus den Messwerten mehrerer, flächenhafter Antastungen abgeleitet. Der Grenzwert SD der zulässigen Kugelabstandsabweichung ist die Kenngröße Kugelabstandsabweichung. Sie wird als längenunabhängige Größe spezifiziert und muss im gesamten spezifizierten Messvolumen eingehalten werden.
- *Ebenheitsmessabweichung RE* – die Spanne der vorzeichenbehafteten Abstände der Messpunkte von der Ausgleichsebene, die nach der Methode der kleinsten Fehlerquadrate berechnet wird.

Die Auswertung der Datensätze und Berechnung der Kenngrößen erfolgte in Geomagic Studio.

5.1 Referenzkörper Hochschule Bochum

Abb. 3 stellt die ermittelten Kenngrößen *Antastabweichung* dar. Es erfolgt vorab keine Datenfilterung. Die Kennlinien des Referenzsystems smartSCAN beziehen sich hierbei auf eine Vergleichsmessung des Kugelkörpers mit einem Lasertracker API T3, während alle anderen Graphen sich auf das smartSCAN-System referenzieren.

Bezogen auf die *Antastabweichung PS* (Abb. 3 links) ist auffällig, dass einige Sensoren (DPI7-HSBO, Mantis, Galaxy Note 1) systematische Abweichungen aufzeigen, d. h. zu kleine oder zu große Messwerte liefern, die insgesamt betrachtet auf ein Maßstabsproblem hinweisen. Gegenüber dem Referenzsystem ist die *Antastabweichung PS* um den Faktor 5 - 35 größer.

Die *Antastabweichung PF* (Abb. 3 rechts) zeigt das Rauschverhalten der Sensoren. Die Spanne der radialen Abweichung von der Ausgleichskugel ist hierbei insbesondere bei den Kamerasystemen auffällig groß. Ein Grund hierfür dürfte in der homogen grauen Oberfläche der

Kugeln liegen, die für die Messungen mit den Scannern mit Kalkspray behandelt wurden, jedoch durch diese Oberflächenausgestaltung weniger geeignet für das Dense Image Matching sind. Die Ergebnisse für die DPI-7-Handscanner sind für alle Systeme homogen und pendeln sich bei einer *Antastabweichung PF* von ca. 10 mm ein. Beim Mantis-Scanner liegt der Wert im Mittel unter 2 mm. Der Structure Sensor zeigt hier mit unter 4 mm ein ebenfalls erstaunlich gutes Rauschverhalten.

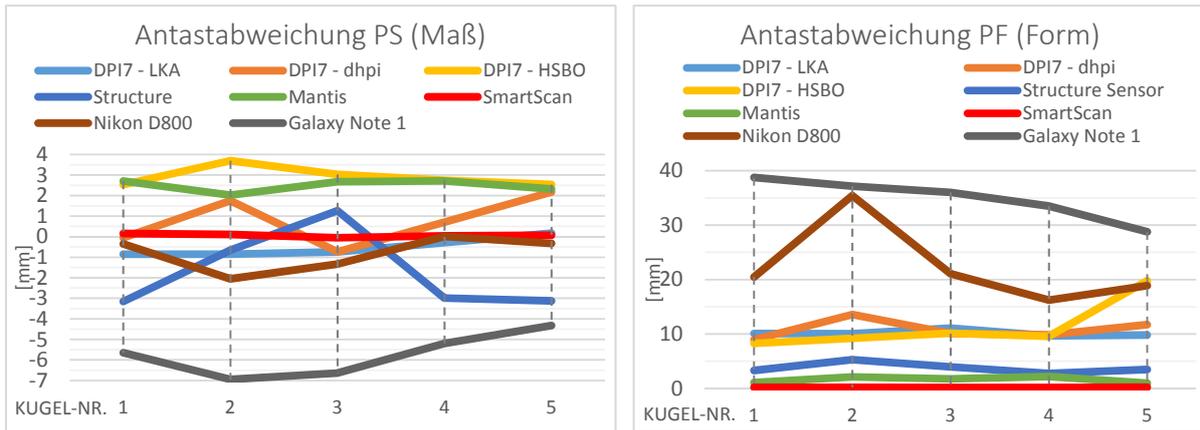


Abb. 3: Kenngröße Antastabweichung (PS / PF) entsprechend VDI/VDE 2634, Blatt 2.

Die in Abbildung 4 dargestellten *Kugelabstandsabweichungen SD* zeigen für nahezu alle Systeme systematisch positive oder negative Abweichungen. Besonders ausgeprägt sind diese Effekte beim Mantis F5 sowie dem Structure Sensor (im Mittel ca. 1% der Streckenlänge). Das bereits erwähnte Maßstabsproblem der Scanner ist hier als Hauptursache zu vermuten.

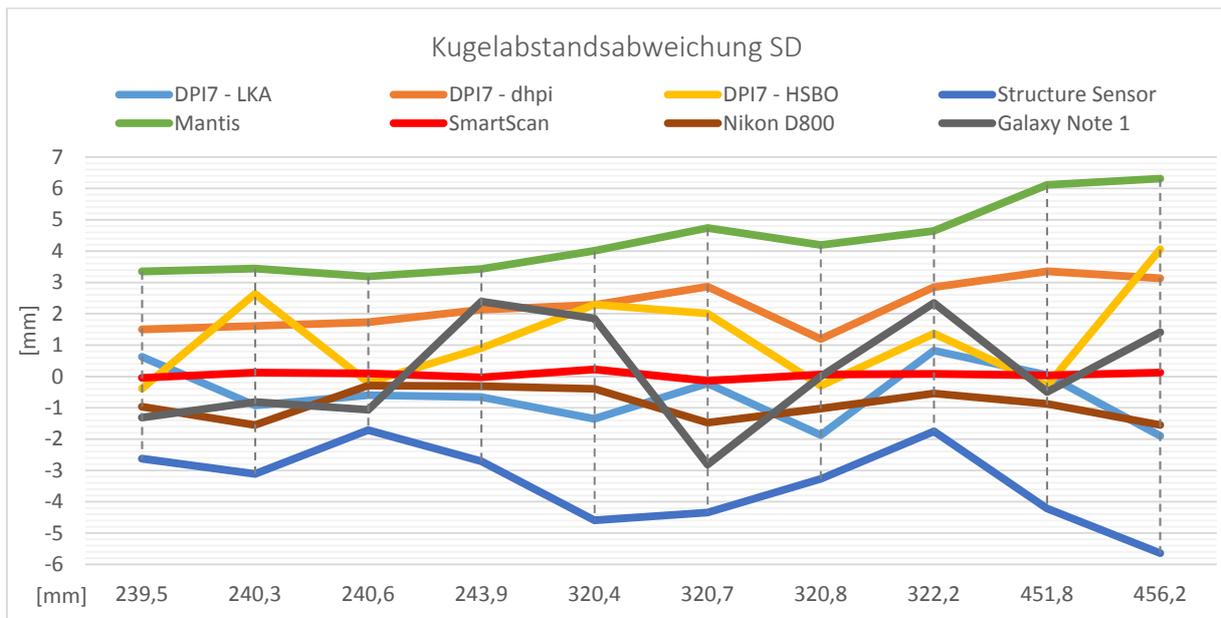


Abb. 4: Kenngröße Kugelabstandsabweichung (SD) entsprechend VDI/VDE 2634, Blatt 3.

5.2 Referenzkörper Granitplatte

Die Diagramme der Abb. 5 geben die Ergebnisse zur *Ebenheitsmessabweichung* R_E wieder. Auch hier ist eine Abhängigkeit zwischen der Ausgestaltung der Objektoberfläche (Granitplatte mit zufälligen Mustern, bestehend aus hellen und dunklen Bereichen) und dem Messprinzip des jeweiligen Sensors erkennbar. Wie aus den Daten der Streifenprojektoren smartSCAN und ATOS erkennbar ist, sind beide Systeme in der Lage, die Oberfläche der Platte mit nahezu gleicher Qualität zu messen, wenngleich die Punktzahl signifikant unterschiedlich ist. Ein Grund ist in den verschiedenartigen Projektor-Einheiten der Scanner zu vermuten (eine aktuelle LED-Beleuchtung beim smartSCAN bzw. Halogenlicht beim ATOS), ein anderer in der unterschiedlichen Auflösung der Kameras.

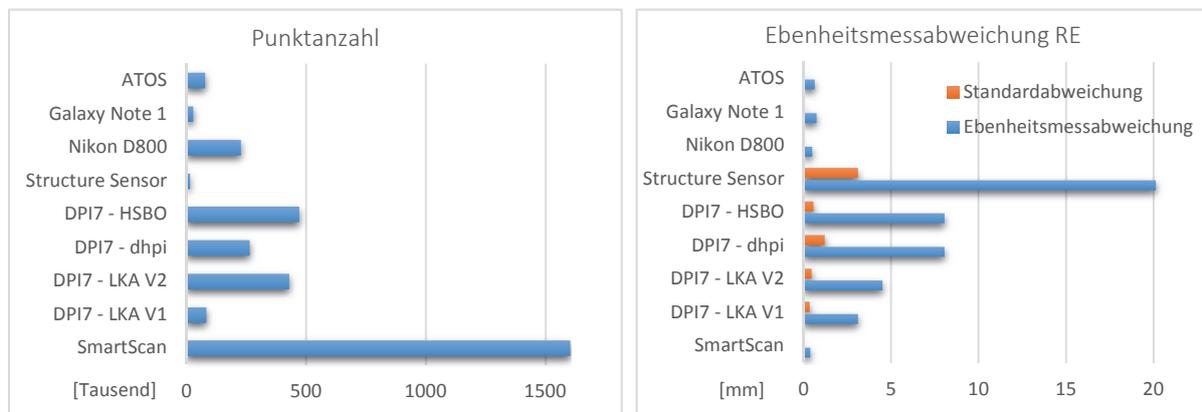


Abb. 5: Kenngröße Ebenheitsmessabweichung (R_E) entsprechend VDI/VDE 2634, Blatt 2.

Die durch Bildtriangulation und Dense Image Matching erfassten Granitplatten weisen Ebenheitsmessabweichungen auf, die in einer zu den Streifenprojektionssystemen vergleichbaren Genauigkeitsklasse liegen. Der Messmethodik kommt hier sicherlich die natürliche Oberflächenstruktur der Granitplatte entgegen, die für das Dense Image Matching von Vorteil ist. Letzteres erweist sich für die Systeme mit aktiver Projektion dagegen eher als Nachteil. Insbesondere der Structure Sensor liefert hier vergleichsweise wenige Messdaten (ca. 15.000 Punkte), die zudem extrem streuen (Abb. 5).

Interessant ist auch noch der bei den DPI-7 Scannern zu erkennende Effekt, dass offensichtlich kein unmittelbarer Zusammenhang zwischen der erfassten Punktmenge und der Qualität der Oberflächenmessung existiert.

5.3 Referenzkörper Testy

Die Ergebnisse der 3D-Vergleiche mit den vier Testys sind in Tabelle 3 zusammengestellt sowie in Abb. 6 farblich illustriert. Die durchschnittliche Abweichung zeigt, dass die zwei Streifenprojektionssysteme sehr geringe Abweichungen zueinander von 4 μm bzw. 12 μm aufweisen und dass sogar die Spanne, die sich aus der Differenz der durchschnittlichen negativen und positiven Abweichungen errechnet, mit 30 μm bis 50 μm sehr gering ist. Damit ist die gute Qualität der Streifenprojektionssysteme als Referenzsystem mit einer übergeordneten Genauigkeit bestätigt. Die beste Qualität konnte mit den Modellen, die aus den Bildverbänden der D800

abgeleitet wurden, erreicht werden, welche durch die geringe Abweichung von bis zu 70 µm und die visuelle Kontrolle (Abb. 6) nachgewiesen wurden. Alle anderen 3D-Scanner konnten diese Genauigkeiten nicht erreichen. Kaum ein System konnte den Testy aufgrund der komplexen Geometrie komplett erfassen und einige Systeme zeigen offensichtlich systematische Maßstabsunterschiede.

Tab. 3: Durchschnittliche Abweichungen an den Testys – 3D-Vergleich mit Geomagic Studio zwischen den untersuchten Systemen und dem Referenzsystem ATOS (Dreiecke ca. 250.000). Einheit [mm]

System	T	#Dreiecke	Ø	SD	Sp
DPI-7 dhp:i	2	588.221	2,16	2,55	3,86
DPI-7 LKA	1	466.114	1,17	0,76	1,81
DPI-7 LKA	2	448.993	1,03	1,09	1,92
DPI-7 HSBO	1	466.846	0,73	1,21	1,71
DPI-7 HSBO	2	554.247	2,01	2,16	3,24
DPI-7 HSBO	3	173.908	1,57	1,19	2,84
DPI-7 HSBO	4	639.945	0,70	1,18	2,07
Artec Spider	1	2.621.776	-1,05	0,85	1,60
Mantis F5	3	4.371.360	0,72	0,67	1,33
Kinect v1	2	76.703	0,61	3,91	6,60
Kinect v2	2	80.267	-2,68	4,28	8,52
StructureSensor	3	99.340	-0,24	1,01	1,63
Nikon D800	1	2.059.835	0,09	0,09	0,18
Nikon D800	2	2.609.969	0,22	0,21	0,47
Nikon D800	4	2.145.685	0,07	0,11	0,18
Galaxy Note 1	2	322.148	0,64	1,19	2,05
smartSCAN	2	1.156.735	-0,012	0,02	0,03
smartSCAN	3	2.626.807	-0,004	0,03	0,05

T... Testy, SD... Standardabweichung, Sp... Spanne

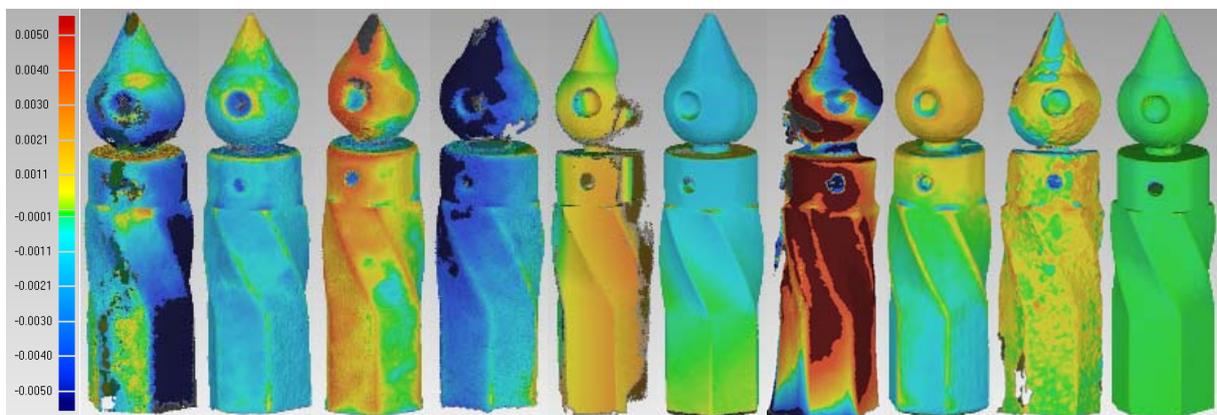
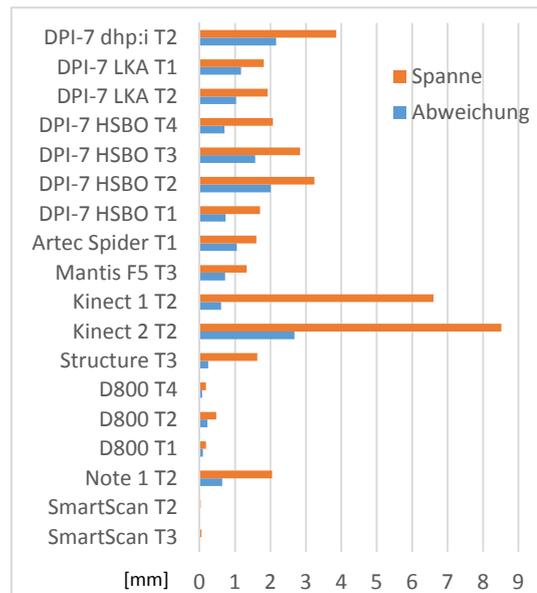


Abb. 6: Abweichungen [m] an verschiedenen Testys nach 3D-Vergleich zur ATOS-Referenz (Testdaten vs. Referenz) (v.l.n.r) – DPI-7 dhp:i T2, DPI-7 LKA T1, DPI-7 HSBO T1, DPI-7 HSBO T2, Artec Spider T1, Mantis F5 T3, Kinect v2 T2, Structure Sensor T3, Galaxy Note 1 T2, Nikon D800 T1.

5.4 Referenzkörper Einstein-Büste

Die Ergebnisse der 3D-Vergleiche mit der Einstein-Büste sind in Tabelle 4 zusammengestellt und in Abb. 7 farblich präsentiert. Das beste numerische Ergebnis erreichte wiederum die Nikon D800, obwohl das 3D-Modell an einigen Stellen wegen der homogenen weißen Oberfläche stark rauscht. Es ist jedoch interessant, dass die Ergebnisse der beiden Dreiecksvermaschungen der Nikon-D800

mit stark unterschiedlicher Auflösung (Faktor 3,7) gleich sind. Aus den Bilddaten des Galaxy Note 1 konnte kein richtiges Modell generiert werden, obwohl die Zahlen in Tabelle 4 ein optimistisches Ergebnis wiedergeben. Aus den Daten der 3D-Scanner DPI-7 und Artec Spider konnten nur Modelle mit einigen großen Lücken erstellt werden. Die hohen Abweichungen bei den DPI-7 zeigen, dass sie mit den homogenen weißen Oberflächenstrukturen der Einstein-Büste nicht zurechtgekommen sind. Auch die beiden höher eingeschätzten Systeme Mantis F5 und Artec Spider haben bei diesem Objekt die vom Hersteller spezifizierten Genauigkeiten nicht erreicht.

Tab. 4: Abweichungen an der Einstein-Büste – 3D-Vergleich mit Geomagic zwischen den untersuchten Systemen und dem Referenzsystem smartSCAN (Dreiecke 1.110.302). Einheit [mm]

System	#Dreiecke	Ø Abw.	SD	Sp
DPI-7 dhp:i	175.696	0,86	1,81	2,57
DPI-7 LKA	167.556	2,20	2,36	3,53
DPI-7 HSBO	77.896	0,88	1,28	2,10
Artec Spider	1.299.298	0,04	0,26	0,45
Mantis F5	8.089.764	0,23	0,25	0,42
Kinect v1	Keine verwertbaren Daten verfügbar			
Kinect v2	Keine verwertbaren Daten verfügbar			
Structure Sensor	1.114.088	0,19	2,68	2,62
D800 V1	3.699.360	0,04	0,11	0,19
D800 V2	99.998	0,03	0,13	0,20
Galaxy Note 1	1.984.979	-0,19	0,94	1,33
ATOS	769.262	0,01	0,02	0,03
SD... Standardabweichung, Sp... Spanne				

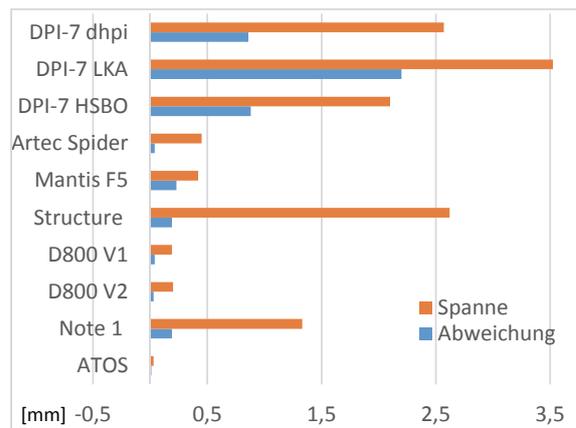


Abb. 7: Abweichungen [m] an der Einstein-Büste nach 3D-Vergleich (Testdaten vs. Referenzdaten) (v.l.n.r) - DPI-7 dhp:i, DPI-7 LKA, DPI-7 HSBO, Artec Spider, Mantis F5, Structure Sensor, Nikon D800, Galaxy Note 1. Farbskala: Grün < 0,0001 m, rot < 0,005 m und blau > -0,005 m.

5.5 Referenzkörper Radnabe

Den komplexesten und schwierigsten Referenzkörper stellte die Radnabe dar. Aufgrund ihrer Symmetrie lässt sich diese nur an einigen wenigen kleinen Objektteilen eindeutig ausrichten, einem gerillten Profil auf der Rückseite und einer erhabenen Schrift im Inneren. Wenn diese in den Scans aufgrund von niedriger Auflösung nicht erkennbar waren, ließ sich das Objekt nicht zum Referenzobjekt ausrichten und somit nicht vergleichen. Außer mit dem smartSCAN, der die Referenzdaten lieferte, wurde mit keinem der untersuchten Systeme ein kompletter Datensatz generiert, da die Radnabe nur aus einer Position aufgenommen und nicht gedreht wurde. Die generierten 3D-Modelle der untersuchten Scanner bzw. Systeme und die Abweichungen zwischen Testdaten und Referenzmodell sind in Abb. 8 farblich dargestellt. Darin zeigt sich, dass neben den beiden Streifenprojektionssystemen der Mantis F5 und die Nikon D800 die optisch besten

Ergebnisse lieferten. Der visuelle Eindruck wird auch durch die Ergebnisse in Tabelle 5 bestätigt, d. h. die geringsten durchschnittlichen Abweichungen weisen die beiden genannten Systeme auf. Dagegen zeigen einige Systeme signifikante systematische Effekte. Der Structure Sensor hat bei diesem Referenzkörper die größten Abweichungen von allen untersuchten Systemen. Er konnte, wie auch der Mantis F5 und der Artec Spider, bei dieser Untersuchung die vom Hersteller angegebene Genauigkeitsspezifikation nicht einhalten. Die anderen Systeme erreichen die spezifizierten Genauigkeiten, allerdings sind sie aufgrund der optischen Qualität des generierten Modells für diese industrielle Anwendung eher ungeeignet.

Tab. 5: Radnabe im 3D-Vergleich mit Geomagic Studio zwischen den untersuchten Systemen und dem Referenzsystem smartSCAN (Dreiecke 6.352.367). Einheit [mm]

System	#Dreiecke	Ø Abw.	SD	Sp
DPI-7 dhp:i	454.234	0,66	1,42	2,19
DPI-7 LKA	Keine Regis. der Scans möglich			
DPI-7 HSBO	Keine Regis. der Scans möglich			
Artec Spider	1.193.774	0,54	2,22	3,00
Mantis F5	5.494.803	0,29	0,60	1,06
Kinect v1	96.477	-0,15	2,20	2,96
Kinect v2	140.701	-0,74	1,94	3,08
Structure Sensor	1.114.088	1,43	1,61	2,67
Nikon D800	5.225.836	0,04	0,38	0,47
Galaxy Note 1	661.749	-0,26	0,77	1,06
ATOS	642.677	-0,18	0,27	0,33
SD...Standardabweichung, Sp.... Spanne				

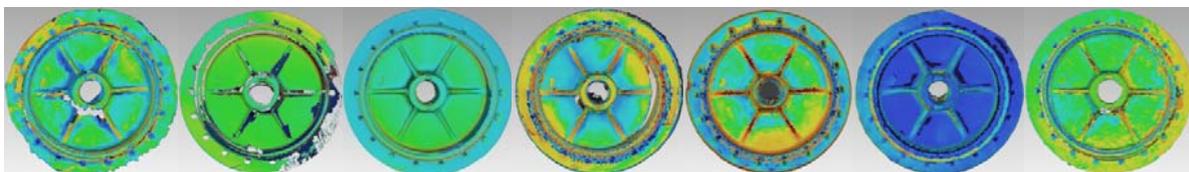
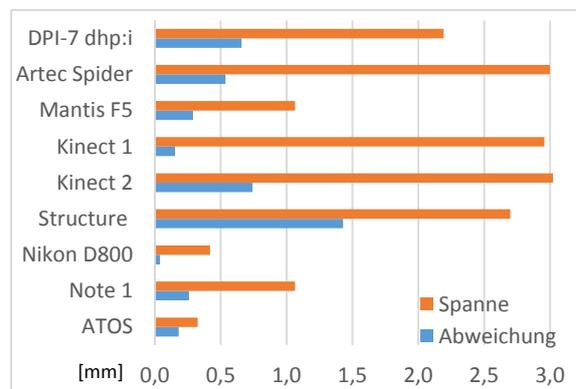


Abb. 8: Abweichungen [m] an der Radnabe - 3D-Vergleich mit Geomagic (Testdaten vs. Referenzdaten) (v.l.n.r.) – DPI-7 dhp:i, Artec Spider, Mantis F5, Kinect v1, Kinect v2, Structure Sensor und Galaxy Note 1. Farbskala: Grün < 0,0001 m, rot < 0,005 m und blau > -0,005 m.

5.6 3D-Testfeldwand

Das temporäre 3D-Testfeld der HCU umfasst ca. 40 schwarz-weiße Zielzeichen aus Papier (DIN A4) auf zwei Wandebenen sowie vorgelagerten Säulen. Die Dimension des Testfeldes von ca. 3×3×2 m³ erlaubt eine Untersuchung unter Testbedingungen, die dem Messvolumen einzelner Sensoren eher entsprechen als die Aufnahme kleiner Referenzkörper. Der 3D-Vergleich erfolgte mit den Ebenen der Wand, des Bodens sowie der Säulen. Als Referenz wurde das Testfeld mit einem IMAGER 5010 von drei Standpunkten gescannt. Die Registrierung der drei Scans erfolgte mit einer Standardabweichung von besser als 2 mm. Verglichen wurden hier die Sensoren, die aufgrund ihrer Eigenschaften auch für größere Objekte gute Ergebnisse versprechen, namentlich die Kinect v2, DPI-7, Google's Project Tango sowie die Nikon D800.

Eine Registrierung der Testdatensätze zur Referenz erfolgte über den ICP-Algorithmus, da außer mit dem Datensatz der Nikon D800 keine Registrierung über die im Testfeld vorhandenen Zielmarken (Passpunktkoordinaten) durchgeführt wurde. Die Ergebnisse der 3D-Vergleiche sind in Abb. 9 dargestellt. Die geringste Abweichung zeigt der Datensatz der Nikon D800. Größere

Abweichungen zeigten sich nur im Nahbereich an einer Säule sowie in Form von groben Ausreißern am Übergang von Boden zur Wandfläche, die manuell gefiltert wurden. Für dieses Objekt zeigt die Kinect v2 ein relativ gutes Ergebnis. Durch die Glättung bei der Erfassung wird die Wandfläche gleichmäßig sehr gut dargestellt. Größere Unterschiede zeigen sich entsprechend des flachen Auftreffwinkels auf den Seiten der Sockel sowie auf der Bodenfläche. Bei den Daten des DPI-7 zeigt sich ein Problem in der Registrierung der Einzelscans als verkettete Registrierung bei gestreckten Objekten. Die Wandfläche erscheint dabei als Ausschnitt einer Zylinderoberfläche, sodass die Randflächen deutlich von der Sollfläche abweichen. Es ist allerdings auch nicht auszuschließen, dass die Datenerfassung in den Randbereichen nicht hinreichend erfolgte. Für den Vergleich wurde mit dem Project Tango nur ein kleiner Ausschnitt des Testfelds erfasst. Dieser relativ dünn besetzte Datensatz zeigt große Ausreißer bei Punkten entlang der Bodenebene sowie zwischen den Oberflächen der Säulen.

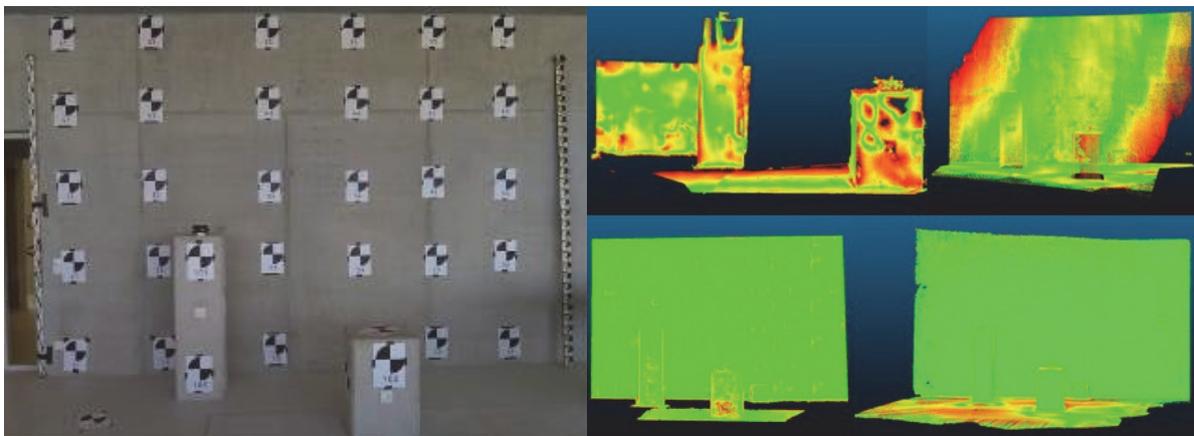


Abb. 9: 3D-Testfeld der HCU Hamburg (links), Abweichungen im HCU 3D-Testfeld – 3D-Vergleich mit CloudCompare (Referenzdaten vs. Testdaten) (v.l.n.r.) – Google's Project Tango, DPI-7 HSBO (oben), Nikon D800 und Kinect v2 (unten), Farbskala: Grün 0 mm, rot 25 mm Abweichung.

6 Fazit und Ausblick

Im Beitrag wurden die Ergebnisse vergleichender geometrischer Genauigkeitsuntersuchungen verschiedener handgeführter 3D-Scanner vorgestellt. Es konnte gezeigt werden, dass die evaluierten Systeme die Genauigkeiten und die Qualität der Referenzdaten aus Messungen von High-End Streifenprojektionssystemen aktuell nicht erreichen. Dabei ist zu berücksichtigen, dass nicht alle verwendeten Referenzkörper dem (jeweils typischen) Anwendungsspektrum der getesteten 3D-Scanner optimal entsprechen.

Die Handhabung der Systeme ist grundsätzlich einfach, jedoch erfordert die Aufnahme durch langsame, gleichförmige Bewegungen – um und über dem Aufnahmeobjekt – eine entsprechende Erfahrung seitens der Nutzer, damit es nicht zum Signalabbruch kommt. Die Aufnahmegeschwindigkeit ist bei allen eingesetzten Systemen mit wenigen Minuten je Objekt recht hoch.

Die in Anlehnung an die Richtlinie VDI/VDE 2634, Blatt 2 und 3, ermittelten Kenngrößen (Antast- und Kugelabstandsabweichungen) weisen nach, dass bei einigen Handscannern der Geräte-Maßstab nicht exakt bestimmt wurde, bzw. auf Grund eines mechanisch zu wenig soliden

Aufbaus des Sensors möglicherweise nicht stabil ist. Prozeduren zur Feldprüfung bzw. einer einfachen, durch den Anwender durchführbaren Selbstkalibrierung, sind hier sinnvoll und notwendig. Die Ergebnisse zur Ebenheitsmessabweichung dokumentieren, dass die bildbasierten Aufnahmeverfahren mit den Kameras sehr geringe Abweichungen gegenüber den Streifenprojektionssystemen haben, während die Systeme mit aktiver Projektion um den Faktor 8 - 50 höhere Abweichungen aufweisen als das Referenzsystem. Das Signal-Rauschverhältnis ist hier verbesserungsbedürftig.

Die zwei im höheren Preissegment angesiedelten Systeme Mantis F5 und Artec Spider konnten die von den Herstellern angegebenen Genauigkeiten bei den Untersuchungen mit den Referenzkörpern Testy, Radnabe und Einstein-Büste nicht einhalten. Allerdings müssen auch der visuelle Zustand und die Vollständigkeit der erfassten Messobjekte neben den reinen Genauigkeitswerten (durchschnittliche Abweichung und Spanne) als Kriterien zur Beurteilung der gesamten Qualität eines untersuchten Systems herangezogen werden. Die visuelle Qualität der Modelle war beim Mantis F5 besser als bei den anderen handgeführten Scannern. Mit den Daten der DPI-7 und dem Artec Spider konnten wegen der vielen Löcher in den Datensätzen keine zufriedenstellenden Modelle der Referenzkörper generiert werden.

Die Ergebnisse der Nikon D800 waren von allen Systemen die besten, d. h. sie zeigen, dass diese Kamera in Kombination mit entsprechender Auswertesoftware gegenüber den handgeführten Scannern eine genaue und preiswerte Alternative darstellt. Die Qualität, der aus den Daten des Galaxy Note 1 erstellten Modelle, war wegen der tlw. großen Löcher nicht brauchbar. Allerdings bleibt abzuwarten, wie die neue Generation der Smartphone-Kameras bei zukünftigen Tests abschneiden wird. Es hat sich auch gezeigt, dass die Qualität eines mit einem bestimmten System generierten Modells sehr stark von der Erfahrung des Operateurs abhängt. Weitere verfügbare, aber noch nicht ausgewertete Daten des FARO Freestyle stehen für ergänzende Analysen bereit. Zukünftige Untersuchungen sollen im Rahmen alternativer Testszenarien, z. B. mit größeren Referenzkörpern durchgeführt werden.

7 Dank

Die Autoren bedanken sich beim Landeskriminalamt Hamburg, bei Dr. Hesse und Partner Ingenieure (dhp:i) und bei MexConsult für die Bereitstellung der Aufnahmesysteme. Ein besonderer Dank gebührt den Masterstudierenden des Studienganges Geomatik der HCU Hamburg, den Studierenden der Hochschule Bochum und den Mitarbeitern des Landeskriminalamtes Hamburg für die tatkräftige Unterstützung bei der Datenerfassung.

8 Literaturverzeichnis

- ADAMS, J.W., OLAH, A., MCCURRY, M.R. & POTZE, S., 2015: Surface Model and Tomographic Archive of Fossil Primate and Other Mammal Holotype and Paratype Specimens of the Ditsong National Museum of Natural History, Pretoria, South Africa. *PloS one*, **10** (10), e0139800.
- AHERN, C., & SPRING, R., 2015: Handheld 3D Capture. *GeoInformatics* **18** (2), 18-19.

- BATHOW, C. & BREUCKMANN, B., 2011: High-definition 3D acquisition of archaeological objects: An overview of various challenging projects all over the world. 23rd CIPA Symposium, 12-16.
- BESL, P.J., & MCKAY, N.D., 1992: Method for registration of 3-D shapes. Robotics-DL tentative, International Society for Optics and Photonics, 586-606.
- BÖHM, J., 2014 Accuracy Investigation for Structured-light Based Consumer 3D Sensors. *Photogrammetrie – Fernerkundung – Geoinformation*, (2), 117-127.
- CHOW, J., ANG, K., LICHTI, D., & TESKEY, W., 2012: Performance analysis of a low-cost triangulation-based 3D camera: Microsoft Kinect system. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39**, 175-180.
- FRIEDMAN, C., JOEL, B.W., SCHULT, A.R. & LEFTWICH, M.C., 2015: Noninvasive 3D Geometry Extraction of a Sea Lion Foreflipper. *Journal of Aero Aqua Bio-Mechanisms* **4** (1), 25-31.
- GONZALEZ-JORGE, H., RIVEIRO, B., VAZQUEZ-FERNANDEZ, E., MARTÍNEZ-SÁNCHEZ, J., & ARIAS, P., 2013: Metrological evaluation of microsoft kinect and asus xtion sensors. *Measurement* **46** (6), 1800-1806.
- HIERONYMUS, J., MISGAISKI, M. & REULKE, R., 2011: Genauigkeitsvergleich von 3D-Sensoren aus dem Freizeit- und Spielmarkt. *Photogrammetrie Laserscanning Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2011*, T. Luhmann/C. Müller (Hrsg.), Wichmann Verlag, Heidelberg, 232-241.
- INZERILLO, L., DI MINO, G., DI PAOLA, F. & NOTO, S., 2015: The Diagnostics of Road Surface Distresses Through Image-Based Modeling Techniques. *Experimental Survey on Laboratory-Rutted Samples. Life Safety and Security* **3** (8), 31-35.
- JAHRAUS, A., LICHTI, D. & DAWSON, P., 2015: Self-Calibration of a Structured Light Based Scanner for Use in Archeological Applications. *Proceedings SPIE 9528, Videometrics, Range Imaging, and Applications XIII*, 95280E.
- KELLER, F., SÄNGER, J., KERSTEN, T., SCHIEWE, J., 2011: Historisches 4D-Stadtmodell der Freien und Hansestadt Hamburg - Automatisierte Generierung und Darstellung innerhalb der Google Earth Engine. *Photogrammetrie – Fernerkundung – Geoinformation*, (3), 155-169.
- KERSTEN, T. & LINDSTAEDT, M., 2012: Image-Based Low Cost Systems for Automatic 3D Recording and Modelling of Archaeological Finds and Objects. *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)*, Volume **7616**, Springer-Verlag Berlin Heidelberg, 1-10.
- KHOSHELHAM, K., 2011: Accuracy Analysis of Kinect Depth Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **38** (5/W12), 133-138.
- KLINGENSMITH, M., DRYANOVSKI, I., SRINIVASA, S.S. & XIAO, J., 2015: CHISEL: Real Time Large Scale 3D Reconstruction Onboard a Mobile Device using Spatially-Hashed Signed Distance Fields. *Proceedings of Robotics: Science and Systems 2015*.
- LACHAT, E., MACHER, H., LANDES, T. & GRUSSENMEYER, P., 2015: Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modeling. *Remote Sensing* **7** (10), 13070-13097.

- LOIANNO, G., CROSS, G., QU, C., MULGAONKAR, Y., HESCH, J.A., & KUMAR, V., 2015: Flying smartphones: Automated flight enabled by consumer electronics. *Robotics & Automation Magazine* **22** (2), 24-32.
- OMELANOWSKY, D., KERSTEN, T. & LINDSTAEDT, M., 2013: Untersuchungen von Low-Cost Systemen zur 3D-Rekonstruktion kleiner Objekte. *Photogrammetrie Laserscanning Optische 3D-Mess-technik, Beiträge der Oldenburger 3D-Tage 2013*, T. Luhmann & C. Müller (Hrsg.), Wichmann Verlag, Heidelberg, 217-228.
- RAU, J. Y. & YEH, P. C., 2012: A semi-automatic image-based close range 3D modeling pipeline using a multi-camera configuration. *Sensors* **12** (8), 11271-11293.
- REULKE, R. & MISGAISKI, M., 2012: Test body “Testy” for Laser Scanning and Optical Systems. *PFG – Photogrammetrie Fernerkundung Geoinformation*, **6**, zum Titelbild.
- SLIZEWSKI, A., FRIESS, M. & SEMAL, P., 2010: Surface scanning of anthropological specimens: nominal-actual comparison with low cost laser scanner and high end fringe light projection surface scanning systems. *Quartär* **57**, 179-187.
- SMISEK, J., JANCOSEK, M. & PAJDLA, T., 2013: 3D with Kinect. *Consumer Depth Cameras for Computer Vision*, Springer London, 3-25.
- TALLIG, G., ZENDER, R. & LUCKE, U., 2015: Konzeption und Verifikation eines Auswahlverfahrens für 3D-Scantechnologien. *Proceedings der Pre-Conference Workshops der 13. E-Learning Fachtagung Informatik*, S. Rathmayer & H. Pongratz (Hrsg.), München, 19-27.
- VDI/VDE, 2002: *Optische 3-D-Messsysteme - Bildgebende Systeme mit flächenhafter Antastung*. VDI/VDE Richtlinie 2634, Blatt 2, Beuth Verlag, Berlin.
- VDI/VDE, 2006: *Optische 3-D-Messsysteme - Bildgebende Systeme mit flächenhafter Antastung in mehreren Einzelansichten*. VDI/VDE Richtlinie 2634, Blatt 3, Beuth Verlag, Berlin.
- WINTERHALTER, W., FLECKENSTEIN, F., STEDER, B., SPINELLO, L., & BURGARD, W., 2015: Accurate indoor localization for RGB-D smartphones and tablets given 2D floor plans. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 3138-3143.
- WRONA, M., 2014: Using Optical NIR Handheld Scanner for Close Range 3D Mapping. *The 9th International Conference on Environmental Engineering 2014*, *Procedia Engineering*, Vilnius.
- WUJANZ, D., WEISBRICH, S. & NEITZEL, F., 2011: 3D-Mapping mit dem Microsoft® Kinect Sensor – erste Untersuchungsergebnisse. *Photogrammetrie Laserscanning Optische 3D-Mess-technik, Beiträge der Oldenburger 3D-Tage 2011*, T. Luhmann & C. Müller (Hrsg.), Wichmann Verlag, Heidelberg, 274-283.
- ZHANG, W., WANG, C., & XI, X., 2015: 3D Scan of Ornamental Column (huabiao) Using Terrestrial LiDAR and Hand-held Imager. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (5/W7), 491-494.

The Colour and Stereo Surface Imaging System (CaSSIS) for ESA's ExoMars Trace Gas Orbiter

N. THOMAS¹, G. CREMONESE², R. ZIETHE¹, M. GERBER¹, M. BRÄNDLI¹, M. ERISMANN¹, L. GAMBICORTI¹, T. GERBER³, K. GHOSE¹, M. GRUBER¹, P. GUBLER¹, H. MISCHLER¹, J. JOST¹, D. PIAZZA¹, A. POMMEROL¹, M. RIEDER¹, V. ROLOFF¹, A. SERVONET¹, W. TROTTMANN¹, T. UTHAICHAROENPONG¹, C. ZIMMERMANN¹, D. VERNANI⁴, M. JOHNSON⁴, E. PELO⁴, T. WEIGEL⁴, J. VIERTL⁴, N. DE ROUX⁴, P. LOCHMATTER⁴, G. SUTTER⁴, A. CASCIELLO⁴, T. HAUSNER⁴, I. FICAI VELTRONI⁵, V. DA DEPPO⁶, P. ORLEANSKI⁷, W. NOWOSIELSKI⁷, T. ZAWISTOWSKI⁷, S. SZALAI⁸, B. SODOR⁸, G. TROZNAI⁸, M. BANASKIEWICZ⁷, J.T. BRIDGES⁹, S. BYRNE¹⁰, S. DEBEI¹¹, M.R. EL-MAARRY¹, E. HAUBER¹², C.J. HANSEN¹³, R. HENSON⁹, A. IVANOV¹⁴, L. KESTAY¹⁵, R. KIRK¹⁵, R. KUZMIN¹⁶, N. MANGOLD¹⁷, L. MARINANGELI¹⁸, W.J. MARKIEWICZ¹⁹, M. MASSIRONI²⁰, A.S. MCEWEN¹⁰, C. OKUBO¹⁵, L.L. TORNABENE²¹, P. WAJER⁷ & J.J. WRAY²²

Zusammenfassung: Die CaSSIS genannte Kamera (für Colour and Stereo Surface Imaging System) wurde kürzlich an Board des ExoMars Satelliten 'Trace Gas Orbiter' (TGO) von der Europäischen Raumfahrtagentur Richtung Mars gestartet. Dort wird TGO im Oktober 2016 ankommen. Die eigentliche wissenschaftliche Mission beginnt im Juni 2017. CaSSIS kann nicht nur hochauflösende Aufnahmen von der Marsoberfläche in vier verschiedenen Farben machen. Der spezielle Rotationsmechanismus erlaubt es ausserdem Stereo-Bilder anzufertigen, wobei die entsprechende Region dafür nur einmal überflogen werden muss.

¹ University of Bern, Physikalisches Institut, Sidlerstr. 5, CH-3012 Bern, Switzerland
E-Mail: nicolas.thomas @space.unibe.ch

² INAF-Osservatorio Astronomicodi Padova, Vicolo Osservatorio 5, 35122 Padova, Italy

³ Wavelab GmbH, Gewerbestrasse 11, CH-3053 Laetti, Switzerland

⁴ RUAG Space, Schaffhauserstrasse 580, CH-8052 Zuerich, Switzerland

⁵ Selex ES, Via A. Einstein 35, 50013 Campi Bisenzio (FI), Italy

⁶ CNR-IFN UOS Padova LUXOR, via Trasea, 7, 35131 Padova, Italy

⁷ PAS Space Research Center, Bartycka 18A, 00-716 Warsaw, Poland

⁸ SGF Technology Associates Co. Ltd., Pipiske u. 1-5/20 1121 Budapest, Hungary

⁹ Space Research Centre, Dept. of Physics & Astronomy, University of Leicester, LE1 7RH Leicester, UK

¹⁰ University of Arizona, Lunar and Planetary Laboratory, Tucson, Arizona 85721-0092, USA

¹¹ CISAS, University of Padova, Via Venezia 15, 35131, Padova, Italy

¹² DLR, Institute of Planetary Research, Rutherfordstr. 4, Berlin-Adlershof, Germany

¹³ Planetary Science Institute, 1700 E. Fort Lowell, Suite 106, Tucson, AZ 85719, USA

¹⁴ eSpace, École Polytechnique Fédérale de Lausanne (EPFL), Station 11, CH-1015 Lausanne, Switzerland.

¹⁵ Astrogeology Science Center, USGS, Flagstaff, Arizona 86001-1637, USA

¹⁶ Vernadsky Institute of Geochemistry and Analytical Chemistry, Russian Academy of Sciences, Kosygin Street 19, Moscow 177975, Russia

¹⁷ LPGN/CNRS, Université Nantes, 44322 Nantes, France

¹⁸ IRSPS, Università d'Annunzio, Pescara, Italy

¹⁹ Max-Planck-Institut für Sonnensystemforschung, Göttingen, Germany

²⁰ Dipartimento di Geoscienze, Università di Padova, via Giovanni Gradeno 6, 35131 Padova, Italy

²¹ Centre for Planetary Science and Exploration/Department of Earth Sciences, University of Western Ontario, London, Ontario N6A 5B7, Canada

²² School of Earth and Atmospheric Sciences, Georgia Institute of Technology, 311 Ferst Drive, Atlanta, Georgia 30332-0340, USA

1 Introduction

The Colour and Stereo Surface Imaging System (CaSSIS) is a moderately high resolution imaging system specifically designed to fly on the European Space Agency's ExoMars Trace Gas Orbiter (TGO) (VAGO et al. 2015; THOMAS et al. 2016a). The spacecraft was launched successfully from Baikonur on 14 March 2016 at 09:31 UT on a PROTON provided by the Russian space agency, ROSCOSMOS. At the time of writing, TGO is in the inter-planetary cruise and will encounter Mars on 19 October 2016. CaSSIS is the only imager onboard and has been designed to both complement the other payload elements and to provide new observations of the surface of Mars once TGO reaches its final orbit in June 2017. The paper presents a brief overview of the instrument and serves as an introduction to a more detailed paper that is currently in preparation for submission to *Planetary and Space Sciences* (THOMAS et al. 2016b). In view of the nature of the conference, we will particularly draw attention to the stereoscopic properties of the instrument.

2 Aims and Objectives

In the 1990s, the European Space Agency (ESA) set up a series of study teams to establish the scientific aims and objectives for a mission focussed on the search for past or present extra-terrestrial life. Outputs from these studies included a detailed report (ESA 1999) and associated papers (e.g. BRACK et al. 1999). This work eventually led to the setting up of the ExoMars mission as part of ESA's optional programme which was aimed at performing an in situ search for extinct or extant life on Mars. During the initial development, it became clear that a two-step programme was needed to guarantee adequate data return. The first element, an orbiter, was to be followed by a rover, 26 months later. The orbiter element was to provide communications while the rover was specifically targeted at investigating the sub-surface where primitive lifeforms might have survived having been protected from the harmful UV radiation present at the surface. Although the orbiter was primarily foreseen as communication infrastructure for the rover element, the possibility to place payload on the orbiter was attractive and the detection of methane in the atmosphere of Mars (FORMISANO et al. 2004; MUMMA et al. 2009) led to the concept of providing the orbiter with the capability to detect trace gases. The instrument definition team report (ZUREK et al. 2009) suggested provision of very high spatial resolution imaging or mapping instruments (e.g., cameras and multi-beam active lasers) to provide geological context and location of small-area sources of trace gases should they exist (e.g., a volcanic vent, rift or crater). The definition team focused on a high-resolution colour stereo camera concept expanding upon the successful HRSC experiment on Mars Express (NEUKUM et al. 2004) that would provide geologic characterization.

Following various re-configurations of the ExoMars programme, the ExoMars Trace Gas Orbiter (TGO) was finally approved with a payload of 4 instruments and an entry, descent, and landing demonstrator (subsequently called Schiaparelli). Schiaparelli was also to be equipped with sensors for measuring atmospheric parameters during its descent. The final orbiter payload comprised two spectrometers (NOMAD and ACS) designed to detect trace gases, a neutron

spectrometer (FRIEND) to investigate water in the first metre of the surface layer, and an imaging system (CaSSIS) to provide geological context.

The main scientific objectives of CaSSIS were defined as

- to characterize sites which have been identified as potential sources of trace gases
- to investigate dynamic surface processes (e.g. sublimation, erosional processes, volcanism) which may contribute to the atmospheric gas inventory and
- to certify potential future landing sites by characterizing local slopes, rocks, and other potential hazards.

Given the limited mass and volume available, the technical solution selected emphasized colour and stereo capability with a moderately high resolution and not attempting to reach the remarkable high resolution imagery (0.25 m/px) of the HiRISE instrument on NASA’s Mars Reconnaissance Orbiter (MCEWEN et al. 2007). The top level technical requirements are shown in Table 1 and were required to be obtained from a circular orbit 400 km above the Martian surface. The prime mission is to last 1 Mars year. Additional science is expected to come from the fact that TGO will not be placed in a Sun-synchronous orbit (unlike most of its predecessors) and hence diurnal variations (e.g. in surface frost coverage) also form major scientific objectives.

Tab. 1: Basic technical requirements for CaSSIS

Spatial resolution	<5 m/px
Number of colour channels	4
Swath width	>8km in full colour
Swath length	>30 km
Stereo capability	Yes, on one pass
Digital resolution	14 bit

A consortium was put together to build the instrument comprising the University of Bern as instrument development lead with support from the Astronomical Observatory in Padova and the Space Research Centre in Warsaw. Local industries were used for specific hardware elements and additional support from SGF in Hungary and the University of Arizona was also provided.

3 Instrument Concept

3.1 Functional Units

CaSSIS comprises two units (Fig. 1) – the Camera Rotation Unit (CRU) and the Electronics Unit (ELU). The CRU includes the optics (telescope) and the detector but also includes a rotation mechanism which will be used to generate the stereo pair. The ELU contains electronics boards in the form of modules for controlling the instrument activities.

3.2 Detector sub-system

The detector sub-system was based around the hybrid silicon CMOS device called Osprey from Raytheon. This 2k x 2k 10 micron pixel pitch device allows rapid read-out from 6 arbitrarily defined areas of the detector. The CaSSIS detector is a flight spare of the detector assembly to be flown in the SIMBIO-SYS experiment slated to fly on ESA’s BepiColombo mission in 2018

(FLAMINI et al. 2010). The approach for CaSSIS was to use this device to produce 2048×256 images of the surface with sufficient rapidity so that they overlap on the ground-track. This approach is sometimes referred to as “push-frame”. The advantage of this approach is that jitter can be determined easily by co-registration of the frames. However, the read-out speed constrains the best resolution one can obtain. Detailed signal to noise calculations in combination with the known properties of the detector led to a target spatial scale of around 4.6 m/px from the nominal 400 km circular orbit intended for TGO.

The timing and the internal buffering of the instrument is sufficient to allow exposure sequences of 40 “framelets” in one pass over a target. Considering the overlap needed to ensure accurate co-registration of the framelets, this results in a maximum swath length of about 40 km. The swath width is around 9.5 km.

The digital resolution of the detector is 14 bit and on ground calibration currently indicates that the device is stable and linear at the framing rates and exposure times required for CaSSIS. A radiator has been mounted to the detector system to cool the device. A heater has also been implemented. This allows us to stabilize the temperature at a value of our choice. The best option seems to be around 0°C when dark current is almost negligible for the millisecond exposure times to be used.

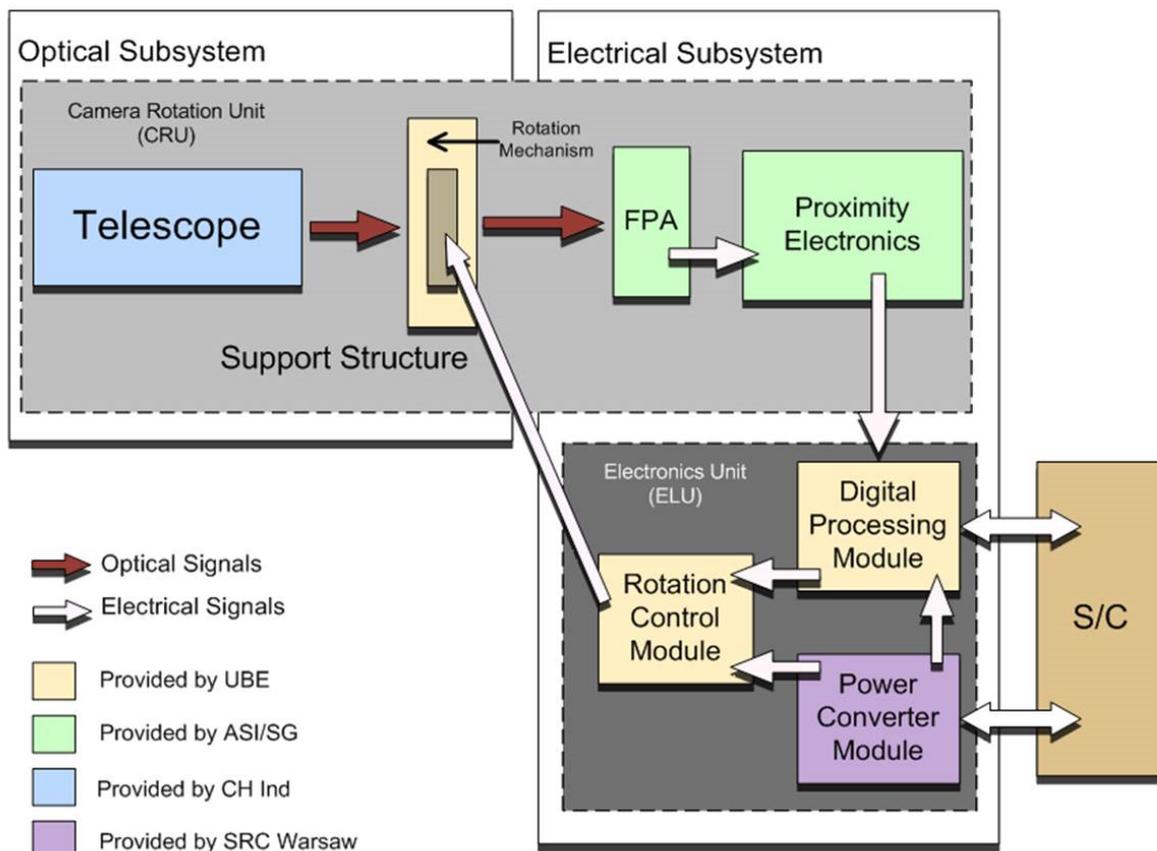


Fig. 1: Functional block diagram of the CaSSIS system.

3.3 Filter sub-system

Individual filters can be placed directly above the detector surface so that different sections of the detector provide different colour response. Custom-designed filters were deposited onto a fused silica substrate to provide 4 different bandwidths. The substrate was then mounted into the detector module. Gaps between the filters were blackened to reduce stray light and cross-talk between the different filter images. The nominal bandwidths are given in Table 2. The NIR is a long-pass filter and the bandwidth limiting is effectively provided by the detector cut-off at 1100 nm which is a consequence of the steep drop in the absorption coefficient of silicon at this wavelength. The timing and data transfer rates are such that only 3 of the 4 filter images can be obtained for each swath if full resolution and full swath width are required. Reduction of swath width and binning of each filter individually is possible providing a significant degree of operational flexibility despite this limitation.

Tab. 2: Filters implemented in CaSSIS

Filter name	Filter type	Central wavelength or cut-on wavelength [nm]
BLU	Bandpass	480
PAN	Bandpass	676
RED	Bandpass	839
NIR	Longpass	860

The capability of CaSSIS to distinguish between specific minerals on the surface of Mars has the subject of a preliminary investigation by TORNABENE et al. (2016).

3.4 Telescope sub-system

To reach the necessary image scale, an 880 mm focal length, F/6.5 telescope was designed using a carbon-fibre reinforced plastic (CFRP) structure and a three-mirror anastigmatic concept based around Zerodur™ mirrors. A fourth mirror was included primarily as a fold mirror but was also required to be powered to compensate for other constraints. The telescope design provides a small (~1%) but non-negligible geometric distortion. The optimum focus position was found in the laboratory using interferometric techniques. The mounting of the focal plane was however slightly shifted to pre-compensate the expected shrinkage of the CFRP structure caused by moisture release.

3.5 Stereo approach

The key concept of the CaSSIS system is that it uses a rotation mechanism to produce the stereo pair. The telescope in this case is not nadir-pointing but at an angle of 10° with respect to nadir. Initially, the telescope is rotated so that it points 10° in front of the nadir position on the ground-track. After acquisition of the first image, the telescope is rotated through 180° so that it now points 10° aft. The second image is then acquired. The off-nadir angle leads to an almost negligible decrease in resolution.

The idea of rotating or tilting a camera to obtain a stereo pair from one pass over an object is not new. For example, the LANYARD system from 1963 (part of the first generation of U.S. photo-reconnaissance satellites) was programmed to tilt between fore and aft to cover the same land area twice during a photographic pass and thus to acquire stereo coverage at 2 m resolution from

around 160 km altitude (see <http://www.fas.org/spp/military/program/imint/corona.htm> and <http://www.earsel.org/SIG/3D/Dati/Porto/Jacobsen1.pdf>). One of the most recent implementations of this technique can be seen in the GeoEye-1 programme (see https://dg-cms-uploads-production.s3.amazonaws.com/uploads/document/file/97/DG_GeoEye1.pdf; retrieved 22 March 2016).

In addition to simplifying the optical solution and allowing a higher resolution, the rotation also neatly solves a further issue. TGO was designed such that one side (the $-Y$ panel) is always nadir pointing. However, the spacecraft rotates about this vector in order to maintain sunlight orthogonal to the spacecraft solar panels. The CaSSIS concept follows an original idea described in MCEWEN et al. (2011) and is an elegant solution to the dual problem of the rotation of the platform about the nadir vector plus the requirement for stereoscopic imaging. Here it is the instrument that moves and not the whole spacecraft (as in the case of GeoEye-1) and, by allowing any rotation position, accomplishes not only the stereoscopic objective but also the compensation for the spacecraft attitude.

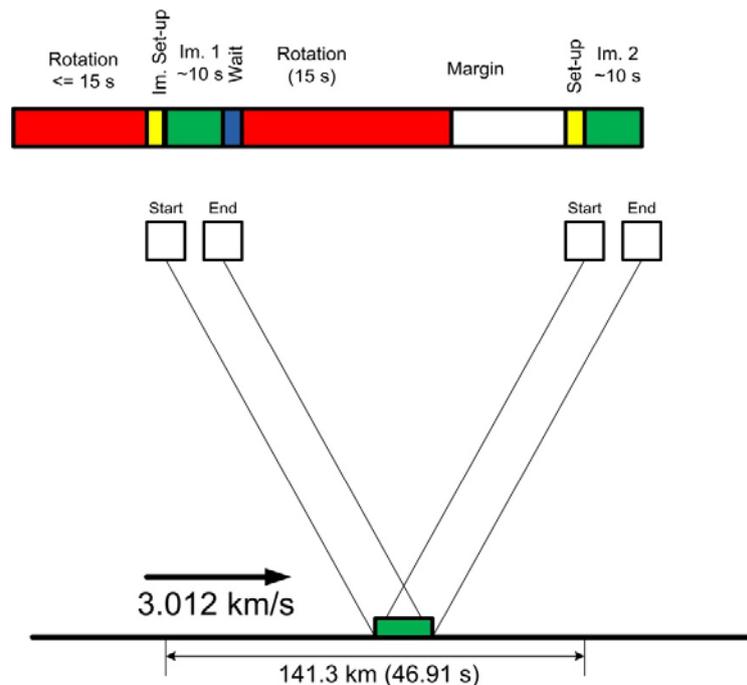


Fig. 2: Timing diagram for the stereo acquisition approach of CaSSIS.

The timing of the stereo acquisition from the 400 km nominal circular orbit is shown in Figure 2. The second image is started 46.91 s after the first which allows more than 30 s for the rotation to be achieved. When combined with the spherical shape of Mars, the resulting stereo convergence angle is nominally 22.4° . On-ground testing suggests that the convergence angle will be slightly smaller although this remains to be confirmed by observations of standard stars in flight.

3.6 Structural concept

A complexity introduced by the rotation is the need to ensure proper routing of the cables needed to drive the electronics. A cable management system (referred to as the “twist capsule”) which

ensures that the wiring for the detector and proximity electronics do not snag was therefore introduced. The instrument structure was then based around a T-like structure with the optics and focal plane assembly on one side of a main support element, the focal plane electronics in the centre, and the cable management system on the other side of the support. The items balance across the support which also contains the rotation mechanism necessary to align the detector and its filters with respect to the ground-track motion of the spacecraft. This neat arrangement is lightweight although care has to be taken to ensure that the structure is sufficiently stiff so that nodding provoked by vibration during launch is controlled. AlBeMet was the material chosen to give the structural support sufficient rigidity.

The CRU was decoupled from the spacecraft bench using a honeycomb panel. This improved mechanical stability as well as allowing the instrument to control its temperature rather precisely using heaters without having to be concerned with heat leaking from/to the spacecraft. The entire structure was wrapped in multi-layer insulation to reduce losses.

3.7 Rotation mechanism

The rotation mechanism consists of a hollow shaft supported by two ceramic bearings and driven by a worm gear, whereby the worm wheel is integral part of the hollow shaft. The reduction ratio is around 200:1.

High-strength titanium alloys are used for the gear component, which are hard coated to provide durability. The housing is made of AlBeMet. A stepper motor (modified Portescap P430) is connected to the worm shaft via a bellow coupling. End switches are used for zeroing; backlash is compensated by software and is calibrated in-flight. The current estimate is that positioning of the mechanism is accurate to about 0.1 deg. The current system rotates the telescope through 180 degrees in <16 s – nearly a factor of 2 faster than required.

3.8 Electronics unit (ELU)

The ELU has three main functions. Firstly, the power converter module (PCM) takes the spacecraft provided 28 V and splits this into the various voltages required by the sub-systems using appropriate DC/DC converters. The digital processing module (DPM) using a radiation-hard LEON-3 processor to control the instrument function. It is also responsible for telecommand and telemetry processing. The rotation control module (RCM) controls the mechanism and its logic.

The flight software (FSW) has been developed with simplicity in mind. On boot, the FSW takes thermal control of the instrument and ensures that all elements (and specifically the detector and telescope) reach their nominal operating temperatures. The FSW also receives and decodes telecommands, drives the detector in a master/slave configuration, and prepares the data received from the detector and proximity electronics for delivery across a SpaceWire interface to the spacecraft. Housekeeping (HK) values are also collected to monitor instrument health. A 1553 interface provides regular HK information to ground via the spacecraft.

CaSSIS can time image acquisition to an accuracy of about 1 ms which corresponds to <1 px on the ground-track implying that the size of the overlap between framelets can be maintained and/or modified accurately. The instrument timing has been verified pre-launch using stroboscopic testing.

The instrument has been allocated a data volume of 2.9 Gb/day during the prime mission. A JPEG2000 compressor is planned to be uploaded to the instrument in flight to provide data compression ratios of 3-10 depending upon the required fidelity. Typically this would result in about 6 stereo pairs per Earth day being returned together with 6-10 individual images (non-stereo) from other targets.

4 Conclusion and future work

4.1 Experimental

The complete instrument (without MLI) is shown in Figure 3. It was successfully mounted on the TGO spacecraft on 11 November 2015 and, following a series of communication tests with the spacecraft, was launched on 14 March 2016. The spacecraft is currently in interplanetary cruise. On ground testing appears to confirm that the instrument build has met its requirements and analysis of the calibration dataset acquired prior to delivery is on-going with publication intended for later this year.

4.2 Geometric analyses

A key part of the in-flight testing will be the verification of the geometric distortion. This will be investigated using standard star fields. The aim is to establish a distortion map based on the high precision astrometry provided by Hipparcos and Gaia. Images of Mars will then be inverted to eliminate the distortion. Once this has been done, co-registration of the overlapping areas in the push-frame sequence will be performed to generate final images around 9.5 km x 40 km in dimension. The stereo pairs can then be matched to provide the 3D topography of specific targets. A key element will be the development of tools for digital terrain model production.

4.3 Target selection

A target database is being constructed for CaSSIS to address the scientific goals of the experiment and the mission. It is intended eventually to open this to the public. The data rate expected and the targeting should result in ~3% of the Martian surface being studied at <5 m/px in colour and stereo every Mars year.

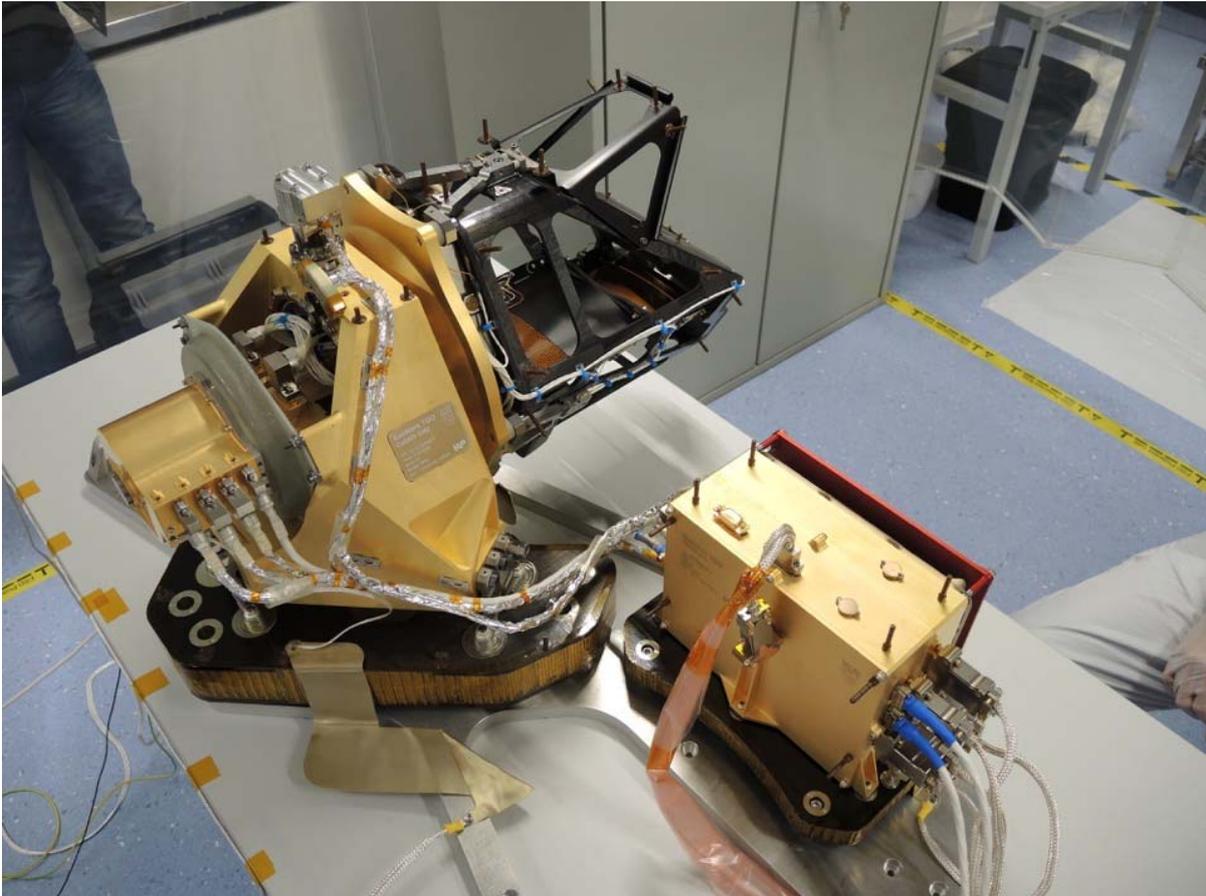


Fig. 3: The CaSSIS instrument on the bench at the University of Bern. The CRU is to the left, the ELU to the right. The black structure on the right of the CRU is the telescope. The multi-layer insulation is not mounted.

5 Acknowledgement

The authors wish to thank the spacecraft and instrument engineering teams for the successful completion of the instrument. CaSSIS is a project of the University of Bern and funded through the Swiss Space Office via ESA's PRODEX programme. The instrument hardware development was also supported by the Italian Space Agency (ASI) (ASI-INAF agreement no.I/018/12/0), INAF/Astronomical Observatory of Padova, and the Space Research Center (CBK) in Warsaw. Support from SGF (Budapest), the University of Arizona (Lunar and Planetary Lab.) and NASA are also gratefully acknowledged.

6 References

- BRACK, A., CLANCY, P., FITTON, B., HOFMANN, B., HORNECK, G., KURAT, G., MAXWELL, J., ORI, G., PILLINGER, C., RAULIN, F., THOMAS, N. & WESTALL, F., 1999: An integrated exobiology package for the search for life on Mars. *Adv. Space Res.* **23**, 301-308.
- ESA, 1999: Exobiology In The Solar System And The Search For Life On Mars Report from the ESA Exobiology Team Study 1997-1998. ESA SP-1231, October 1999.
- FLAMINI, E., CAPACCIONI, F., COLANGELI, L., CREMONESE, G., DORESSOUNDIRAM, A., JOSSET, J.-L., LANGEVIN, Y., DEBEL, S., CAPRIA, M.T., DE SANCTIS, M.C., MARINANGELI, L., MASSIRONI, M., MAZZOTTA EPIFANI, E., NALETTO, G., PALUMBO, P., ENG, P., ROIG, J.F., CAPORALI, A., DA DEPPO, V., ERARD, S., FEDERICO, C., FORNI, O., SGAVETTI, M., FILACCHIONE, G., GIACOMINI, L., MARRA, G., MARTELLATO, E., ZUSI, M., COSI, M., BETTANINI, C., CALAMAI, L., ZACCARIOTTO, M., TOMMASI, L., DAMI, M., FICAI VELTRONI, I., POULET, F., HELLO, Y. & SIMBIO-SYS TEAM, 2010: SIMBIO-SYS: The spectrometer and imagers integrated observatory system for the BepiColombo planetary orbiter. *Planetary and Space Science* **58**, 125-143.
- FORMISANO, V., ATREYA, S., ENCRENAZ, T., IGNATIEV, N. & GIURANNA, M., 2004: Detection of Methane in the Atmosphere of Mars. *Science* **306**, 1758-1761.
- MC EWEN, A.S., ELIASON, E.M., BERGSTROM, J.W., BRIDGES, N.T., HANSEN, C.J., DELAMERE, W.A., GRANT, J.A., GULICK, V.C., HERKENHOFF, K.E., KESZTHELYI, L., KIRK, R.L., MELLON, M.T., SQUYRES, S.W., THOMAS, N. & WEITZ, C.M., 2007: Mars Reconnaissance Orbiter's High Resolution Imaging Science Experiment (HiRISE). *Journal of Geophysical Research (Planets)* **112**, E05S02.
- MUMMA, M.J., VILLANUEVA, G.L., NOVAK, R.E., HEWAGAMA, T., BONEV, B.P., DISANTI, M.A., MANDELL, A.M. & SMITH, M.D., 2009: Strong release of methane on Mars in northern summer 2003. *Science* **323**, 1041-1045.
- NEUKUM, G., JAUMANN, R. & THE HRSC CO-INVESTIGATOR AND EXPERIMENT TEAM, 2004: HRSC: The High Resolution Stereo Camera of Mars Express. ESA Special Publication, SP-1240, 1-19.
- THOMAS, N. & 58 OTHERS. 2016a: The Colour and Stereo Surface Imaging System for ESA's Trace Gas Orbiter. *LPSC 47*, #1306 (extended abstract).
- THOMAS, N. & 58 OTHERS. 2016b: The Colour and Stereo Surface Imaging System for ESA's Trace Gas Orbiter. *Planetary and Space Science*, submitted.
- TORNABENE, L.L., SEELOS, F.P., POMMEROL, A., HANSEN, K.T., SEGAL, N., THOMAS, N., CREMONESE, G., MC EWEN, A.S., SUTTON, S. & CHOJNACKI, M., 2016: Analysis of Colour And Stereo Surface Imaging System (CaSSIS) colour capabilities and simulated images generated from MRO datasets. *LPSC 47*, #2695 (extended abstract).
- VAGO, J., WITASSE, O., SVEDHEM, H., BAGLIONI, P., HALDEMANN, A., GIANFIGLIO, G., BLANCQUAERT, T., MCCOY, D. & DE GROOT, R., 2015: ESA ExoMars program: The next step in exploring Mars. *Solar System Research* **49** (7), 518-528.
- ZUREK, R.W., CHICARRO, A., ALLEN, M.A., BERTAUX, J.L., CLANCY, R.T., DAERDEN, F., FORMISANO, V., GARVIN, J. B., NEUKUM, G. & SMITH, M. A., 2009: Final Report from the 2016 Mars Orbiter Bus Joint Instrument Definition Team (JIDT). November 2009.

Möglichkeiten und Einschränkungen des kombinierten Einsatzes digitaler Senkrecht- und Schrägluftbildsensoren in einer Großformatkamera

JENS KREMER¹ & PHILIPP GRIMM¹

Zusammenfassung: Verschiedene Weiterentwicklungen der Auswertelgorithmen und der Computer- und Visualisierungstechnik erlauben die Erzeugung von neuen luftbildbasierten Produkten wie z.B. automatisch erzeugten texturierten Stadtmodellen oder hochauflösenden Schrägluftbildern mit entsprechenden Nutzungsumgebungen. Die dadurch entstehende größere Vielfalt an unterschiedlichen Auswertungs- und Nutzungstechniken stellt sehr verschiedene Anforderungen an das zugrundeliegende Luftbildmaterial und damit auch an die verwendeten Kamerasysteme und Fluggeräte. Aufgrund der sehr verschiedenen Anforderungen kann es keinen optimalen Sensor für alle denkbaren Anwendungen geben. Nichtsdestotrotz kann mit einem modularen Kamerasystem ein sehr weiter Bereich der tatsächlich auftretenden Anforderungen abgedeckt werden.

Neben der Auswahl der Kameramodule sind die variablen Parameter eines solchen Systems die Brennweite der eingesetzten Objektive, die Anzahl der verwendeten Kameras sowie die Ausrichtung der einzelnen Module.

In der Praxis zeigt es sich, dass bei der Auswahl der eingesetzten Kameramodule die Fähigkeiten einer hohen Bildwiederholrate und der Möglichkeit zur Nutzung kurzer Belichtungszeiten besonders wichtig sind.

Bisher ist die sequentielle Erzeugung von Senkrecht- und Schrägluftbildern mit doppelten Flugstunden verbunden, da die Streifenbreite der Nadir-Komponente von Schrägluftbildkameras nicht an die einer Großformatkamera heranreicht.

Es wird eine neue Generation von Kameramodulen vorgestellt. Mit diesen Modulen lässt sich ein Kamerasystem aufbauen, welches eine hohe Streifenbreite im Nadir-Blick mit der Möglichkeit zur Erzeugung von Schrägluftbildern vereint.

1 Einleitung

Verschiedene Weiterentwicklungen der Auswertelgorithmen und der Computer- und Visualisierungstechnik erlauben die Erzeugung von neuen luftbildbasierten Produkten wie z.B. automatisch erzeugten texturierten Stadtmodellen oder hochauflösenden Schrägluftbildern mit entsprechenden Nutzungsumgebungen. Die dadurch entstehende größere Vielfalt an unterschiedlichen Auswertungs- und Nutzungstechniken stellt durchaus sehr verschiedene Anforderungen an das zugrundeliegende Luftbildmaterial und damit auch an die verwendeten Kamerasysteme und Fluggeräte. Aufgrund der sehr verschiedenen Anforderungen die z.B. an ein Kamerasystem in einem kleinen UAV im lokalen Einsatz und für den Einsatz in einem schnell und hoch fliegenden Jet bei den Befliegung großer Flächen gestellt werden, kann es keinen optimalen Sensor für alle denkbaren Anwendungen geben. Nichtsdestotrotz kann mit einem modularen Kamerasystem, bei dem entsprechend leistungsfähige Kameramodule je nach Anwendung auf verschiedene

¹ IGI mbh, Langenauer Str. 46, D-57223 Kreuztal, E-Mail: [j.kremer, p.grimm]@igi-systems.com

Weise integriert werden können, ein sehr weiter Bereich der tatsächlich auftretenden Anforderungen abgedeckt werden.

Mit der DigiCAM Serie stellt die IGI mbH seit vielen Jahren ein solches modulares System zur Verfügung. Neben dem Einsatz als einzelne Mitteformatkamera bzw. als komplementärer Sensor bei Airborne Lidar Systemen werden die DigiCAM Module in der Quattro DigiCAM Konfiguration und in der Penta DigiCAM Konfiguration als Oblique-System mit einem Nadir Sensor mit reduzierter Streifenbreite eingesetzt. Diese Konfigurationen können von Nutzer zur jeweils anderen Konfiguration umgebaut werden. Allerdings bedeutet dies, dass eine simultane Erfassung von hochauflösenden Nadir Aufnahmen und von Oblique Bildern nicht möglich ist.

Ein einziger Aufbau, der die beiden Konfigurationen zusammenfasst war bisher aus Platzgründen nicht möglich, weil alle Kameramodule in einem Kameraträger Platz finden müssen, der in eine stabilisierte Aufhängung passt. Weiterhin müssen auch alle Kameras durch ein Kameraloch mit üblichen Maßen blicken können.

Die Einführung eines neuen, leistungsfähigeren Kameramoduls hat nun diese Situation geändert. Mit diesen Modulen ist es möglich, ein kombiniertes Großformat-Nadir und Vierfach-Oblique System mit den notwendigen Maßen zu realisieren.

2 Anforderung an Oblique- und Nadir-Luftbildkameras

Bei der Auswahl eines Luftbildkamarasystems für eine bestimmte Anwendung sind offensichtlich die Eigenschaften des optischen Systems und der Anzahl der verfügbaren Bildpunkte die wichtigsten Kenngrößen. Bei der Planung des Bildflugs und bei der praktischen Durchführung des Bildflugs stellen sich aber auch besonders die im Betrieb erreichbare Bildwiederholrate sowie die nutzbaren Belichtungszeiten als entscheidend heraus.

2.1 Bildwiederholrate

Bei der Anwendung von „Dense Matching“ Techniken zur Weiterverarbeitung der Luftbilder sind hohe Bildüberlappungen in Vorwärtsrichtung notwendig. Um eine hohe Bildähnlichkeit zwischen aufeinanderfolgenden Bildern zu erreichen und um eine Redundanz der Bestimmung der 3D-Punkte zu ermöglichen wird heute mit Vorwärtsüberdeckungen von 80% und mehr gearbeitet. Weil gleichzeitig aber auch immer bessere Bodenauflösungen erreicht werden sollen, ist eine hohe Bildwiederholrate zum Erreichen der notwendigen Überdeckungen unbedingt notwendig, wenn keine Einschränkungen beim Einsatz von relativ schnell fliegenden Plattformen hingenommen werden sollen.

Beispiel:

Soll bei einer Fluggeschwindigkeit von 160 kn (82 m/sec) ein Pixelabstand am Boden (GSD) von 5cm, und eine Überdeckung von 80% erreicht werden, so ist bei einer Kamera mit 10.000 Pixeln in Flugrichtung eine Bildfolgezeit von mindestens 1,2 Sekunden notwendig. Zieht man noch praktische Einflüsse, wie Geschwindigkeitsvariationen durch Wind in Betracht, wird man bei diesem Beispiel eher eine Kamera mit einer Bildrate von etwa einem Bild pro Sekunde einsetzen wollen.

Diese Anforderung an eine hohe Bildrate besteht aber nicht nur für senkrecht blickende Kameras, sondern auch für Schrägluftbildsysteme. Wenn auch bisher die Schrägluftbilder vorwiegend für Anwendungen zum Einsatz kamen, bei denen nur eine geringe Überdeckung gefordert war, so erfordert die Auswertung von Schrägluftbildern mit „Dense Matching“ Techniken auch eine ebenso hohe Überdeckung wie bei der Nadir-Anwendung.

2.2 Belichtungszeit

Bewegt sich die Kamera oder das Motiv während der Belichtung des Bildes, so wirkt sich diese Bewegung als eine Verwischung im Bild aus.

Bei Kameras die in Nadir-Richtung blicken, kann dieses Problem der Bildverwischung durch verschiedene FMC („Forward Motion Compensation“) Techniken stark minimiert werden. Bei der Erzeugung von Schrägluftbildern ist eine solche Kompensation der Bildverwischung mit den aktuell verwendeten Techniken aber nicht möglich. Die Stärke der Verwischung hängt in diesem Fall von der Position innerhalb des Bildes ab und kann deshalb durch eine Korrektur, die auf das ganze Bild gleichmäßig wirkt, nicht kompensiert werden.

Wird wie im obigen Beispiel von einer Fluggeschwindigkeit von 160 kn und einer GSD von 5cm ausgegangen, so ist eine Belichtungszeit von etwa 1/1600 Sekunde notwendig um die Bildverwischung auf einen Pixel zu begrenzen.

Ein ähnliches Problem tritt bei der Nutzung der Kameras ohne stabilisierte Aufhängungen bzw. ohne aufwändige Schwingungsreduzierung auf. In diesem Fall kommt zur wohldefinierten Bildverwischung durch die Fluggeschwindigkeit noch eine Bildverwischung durch die Winkelbewegung der Kamera in beliebige Richtungen. Auch in diesem Fall ist die Reduktion der Bildverwischung durch die Wahl einer kurzen Belichtungszeit wünschenswert.

Wenn ein Kameramodul die oben abgeschätzten Belichtungszeiten erreicht, können also nicht nur die Voraussetzungen für verwischungsfreie Oblique-Bilder geschaffen werden und Einzelkameras auf mechanisch einfachen Aufhängungen betrieben werden, sondern es kann auch für kombinierte, großformatige Systeme für die meisten Flugbedingungen auf aufwändige FMC Techniken verzichtet werden.

3 Modulares Kamerasystem IGI DigiCAM

Die IGI mbH stellt seit einigen Jahren das modulare DigiCAM Luftbildkamarasystem bereit. Es basiert auf Mittelformatkamera-Modulen, die je nach Anforderung entweder einzeln oder als Multi-DigiCAM System eingesetzt werden. Die Kameramodule besitzen vom Nutzer austauschbare Objektive und ein Datenerfassungssystem mit einer komfortablen grafischen Benutzeroberfläche. Der Umbau zwischen den Multi-DigiCAM Konfigurationen ist vom Nutzer selbst in kurzer Zeit (wenige Stunden) durchführbar.

3.1 IGI Quattro DigiCAM

Um die große Abdeckung einer großformatigen Kamera zu erreichen, wird die Konfiguration „Quattro DigiCAM“ eingesetzt. Bei dieser Konfiguration erzeugen vier Kameramodule die hier gezeigte Bodenabdeckung.

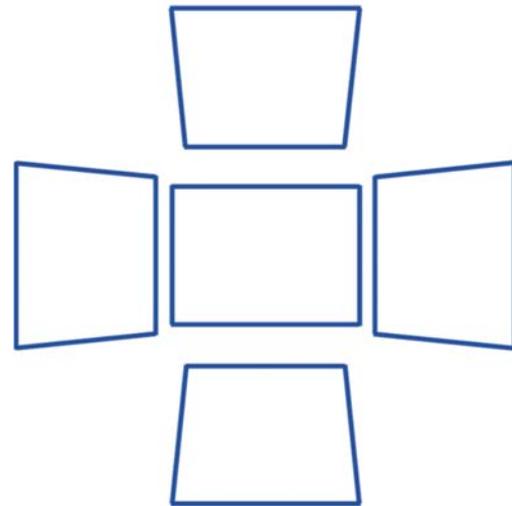
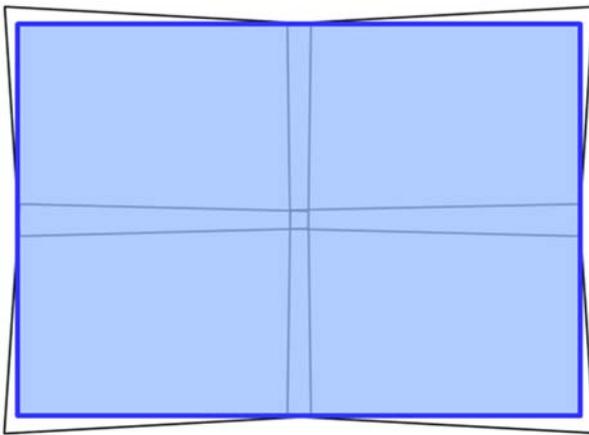


Abb. 1: Bodenabdeckung einer Quattro DigiCAM Abb. 2: Bodenabdeckung einer Penta DigiCAM

Die Kameramodule werden hochgenau synchronisiert ausgelöst. Deshalb kann durch eine geometrische und radiometrische Anpassung der Kameramodule ein homogenes und geometrisch korrektes Gesamtbild erzeugt werden. (In der Abbildung blau markiert.)

3.2 IGI Penta DigiCAM

Eine andere erfolgreich eingesetzte Konfiguration ist die Penta DigiCAM. Bei dieser Konfiguration werden vier Kameras unter 45° montiert und zur Erzeugung von Schrägluftbildern eingesetzt. Die fünfte Kamera blickt nach unten.

4 IGI DigiCAM 100

Während die oben dargestellten DigiCAM Konfigurationen mit CCD basierten DigiCAM Modulen mit bis zu 60 Megapixeln realisiert worden sind, stehen nun auch Kameramodule auf CMOS Basis mit 100 Megapixeln zur Verfügung.

Tab. 1: Eigenschaften der DigiCAM 100

Anzahl Bildpunkte	11.600 × 8.700
Technologie	CMOS, Back Side Illuminated
Pixel Pitch	4,6µm
Min. Belichtungszeit	1/1600 sec
Dynamikbereich	84dB
Min. Bildfolgezeit	0,85 sec

Die Kameramodule werden über modulare Rechereinheiten, sogenannte „Sensor Management Units“ (SMU) gesteuert. In diesen SMU werden auch die Bilder auf im Flug austauschbaren Speichereinheiten abgelegt. Die grafische Nutzerschnittstelle wird entweder auf Touch-Screen Anzeigemodulen oder auf einem speziellen Tablett-PC bereitgestellt.



Abb. 3: DigiCAM 100 mit SMU und User Interface

Diese DigiCAM 100 Module zeigen nicht nur eine größere Anzahl an Pixeln, sondern durch die verwendete CMOS Technologie und den speziellen Aufbau der Pixel eine stark verbesserte Empfindlichkeit und Dynamik, die den Einsatz selbst bei widrigen Wetterbedingungen ermöglicht.

Abb. 4 zeigt beispielsweise einen Ausschnitt aus einem mit der DigiCAM 100 erzeugten Luftbild. Der Bildflug fand am 22. Dezember 2015 in Kreuztal statt. Wie an dem Fehlen von Schatten zu erkennen ist, herrschte während des Fluges eine komplett geschlossene Wolkendecke. Die Bilder des photogrammetrischen Blocks wurden mit 2cm GSD aufgenommen.



Abb. 4: Luftbild mit 2cm GSD, aufgenommen unter komplett geschlossener Wolkendecke am 22. Dezember 2015 in Kreuztal.

4.1 IGI Penta DigiCAM „Urban Mapper“

Die Eigenschaften des DigiCAM 100 Kameramoduls machen neue Multi-DigiCAM Konfigurationen möglich, welche die Vorteile der Quattro DigiCAM mit denen der Penta DigiCAM verbindet.

Die hohe Wiederholrate in Verbindung mit der größeren Pixelzahl ermöglichen es, eine sehr hohe Überdeckung in Flugrichtung zu realisieren, ohne in dieser Richtung mehrere Bilder zu einem Bild kombinieren zu müssen. So kann eine Kamera aufgebaut werden, die nicht nur eine große Nadir Abdeckung aufweist, sondern auch ein vollwertiges Oblique-Kamera System mit 100 Mpixel Schrägluftbildern darstellt. Anders als bei den oben kurz beschriebenen Quattro DigiCAM und Penta DigiCAM Konfigurationen können die Großformat-Nadir Option und die Vierfach-Oblique Version aber nicht nur nacheinander, sondern zeitgleich in einem Sensoraufbau realisiert werden. Die neue IGI Penta DigiCAM „Urban Mapper“ gehört aus diesem Grund zu einer neuen Generation digitaler Luftbildkamarasysteme, die in Lage sind großformatige Nadir- und Schrägluftaufnahmen innerhalb eines Fluges zu generieren.

Die Konfiguration „Urban Mapper“ ist optimiert für die Erfassung von großen urbanen Gebieten. Sie erlaubt eine gleichzeitige Erzeugung von klassischen Senkrechtluftbildern mit über 80% Vorwärtsüberdeckung zusammen mit Schrägluftbilder hoher Qualität. Sie ermöglicht damit die effiziente Erzeugung von hochwertigen, texturierten Digitalen Stadtmodellen aus dem Ergebnis einer Befliegung mit einem einzigen Sensorsystem.

4.1.1 Aufbau des „Urban Mapper“

Die Kamera besteht aus acht Kameramodulen. Vier Module bilden den Nadir Blick. Drei der Module werden dabei in RGB Konfiguration betrieben und definieren das Format des Nadir-Blicks der Kamera. Sie sind nebeneinander und mit der langen Seite entlang der Flugrichtung ausgerichtet. Das vierte Nadir-Modul erfasst den optionalen NIR Kanal. Die weiteren vier Kameramodule (RGB) blicken nach vorn, rückwärts, nach rechts und nach links. Der Oblique-Winkel ist mit 42° etwas niedriger gewählt als bei der Penta DigiCAM, um in innerstädtischen Häuserschluchten eine komplette Erfassung der Hausfassaden zu gewährleisten. Für alle RGB Kameramodule werden Objektive der Firma Rodenstock mit 90 mm Brennweite eingesetzt.

Die acht Kamera Module sind so angeordnet, dass sie in einer der gebräuchlichen kreiselstabilisierten Kameraaufhängungen betrieben werden können. Das IGI AEROcontrol GNSS/IMU System ist in der Kamera integriert. Zur Flugführung wird das IGI CCNS5 eingesetzt.

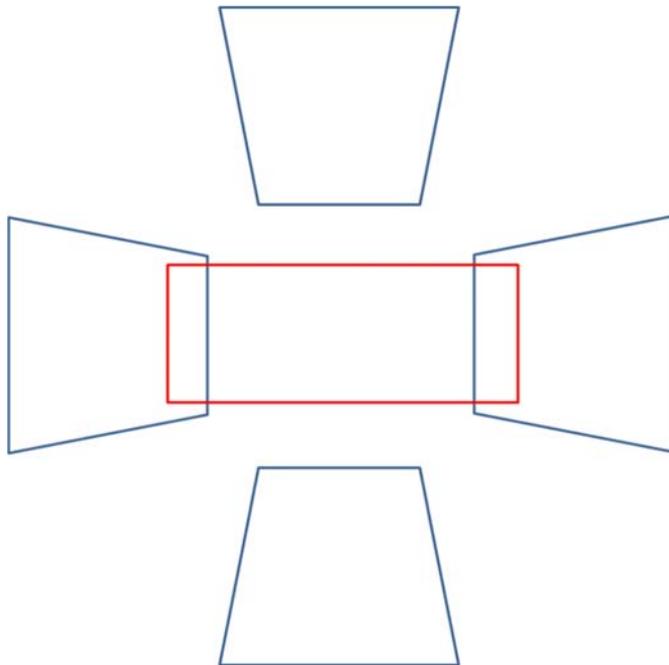


Abb. 5: Footprint der DigiCAM „Urban Mapper“



Abb. 6: Penta DigiCAM „Urban Mapper“

4.1.2 Kameraparameter / Flugplanung

Die oben gezeigte Konfiguration „Urban Mapper“ verfügt über die folgenden Parameter:

Tab. 2: Eigenschaften der Penta DigiCAM „Urban Mapper“

Bildformat (Nadir-Blick)	28.200 × 11.500
Bildformat (Oblique-Blick)	11.600 × 8.700
Oblique Winkel	42°
Brennweite (Nadir & Oblique)	90 mm
Min. Bildfolgezeit	0,85 sec
Pixel Pitch	4,6 µm
Min. Belichtungszeit	1/1600 sec
NIR Kanal	Optional durch das 8'te Kameramodul

Daraus folgt z.B. für einen Flug mit 160 kn und einer GSD von 5 cm, dass eine Vorwärtsüberdeckung von mehr als 87% rechnerisch möglich ist. Bei diesen Bedingungen ist mit einer unkompenzierten Bildverwischung von einem Pixel auch nicht mit nachteiligen Effekten durch Bildverwischung zu rechnen. Je nach Flugbedingungen können also auch GSDs von weniger als 5 cm bzw. Überdeckungen von über 90% realisiert werden.

4.1.3 User Interface

Anders als bei Einzelkamarasystemen wie in Abb. 3 nutzt der Kameraoperator beim „Urban Mapper“ einen 20-Zoll-Touchscreen um das System zu steuern. Dabei stehen dem Nutzer ver-

schiedene Ansichten zur Verfügung. Die Abbildungen 7 und 8 zeigen zwei der möglichen Ansichten.



Abb. 5: Beispiel für eine Ansicht im „Urban Mapper“ Nutzerinterface.

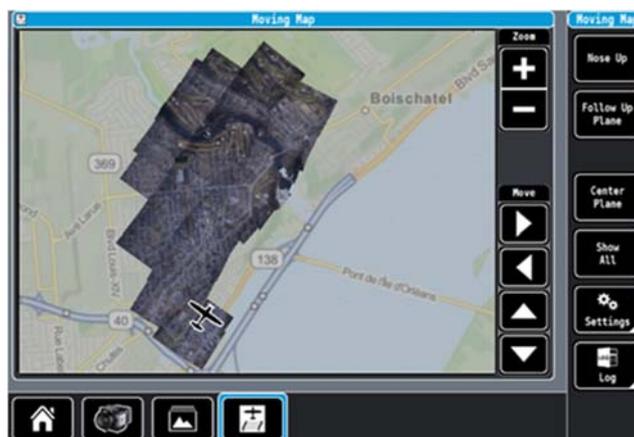


Abb. 6: „Moving Map“ Ansicht im „Urban Mapper“ Nutzerinterface.

5 Zusammenfassung

Eine hohe Bildwiederholrate und die realisierbaren Belichtungszeiten zusammen mit einer hohen Empfindlichkeit der Kameramodule, wurden als besonders wichtig für einen flexiblen Einsatz von modularen Luftbildkameras identifiziert. Die Auswirkung dieser Parameter auf die Einsetzbarkeit für bestimmte Projekte wurde anhand eines Beispiels für das neue IGI DigiCAM 100 Kameramodul gezeigt.

Die Eigenschaften des DigiCAM 100 Moduls erlauben den Aufbau des Systems „Urban Mapper“. Der „Urban Mapper“ vereint die hohe Streifenbreite und Auflösung einer reinen Nadir-Luftbildkamera mit der Fähigkeit, hochwertige Schrägluftaufnahmen zu generieren. Damit können alle Ausgangsdaten für ein hochwertiges, optimal texturiertes Stadtmodell in einem Flug mit einem einzigen Kamerasystem erzeugt werden.

UAS Photogrammetry of Homogenous Snow Cover

YVES BÜHLER¹, ANDREAS STOFFEL¹, MARC ADAMS², RUEDI BÖSCH³ & CHRISTIAN GINZLER³

Abstract: Photogrammetry of the snow cover is expected to be difficult, because of its homogenous, low-contrast surface, hindering the automated detection of matching points to generate digital surface models. On the other hand there is rising request for high spatial resolution snow depth and snow surface type mapping as well as snow avalanche event documentation. Unmanned aerial systems (UAS) enable flexible, efficient and economic data acquisition, even within inaccessible alpine terrain. The efficient and economic generation of high-quality digital surface models (DSM) of snow covered surfaces would be a major step forward for many applications in snow hydrology and hydropower generation, avalanche research and warning, for winter tourism as well as for alpine ecology investigations.

We investigate the performance of UAS-based structure-from-motion (SfM) photogrammetry on very homogenous snow surfaces at the test site Tschuggen at 2,000 m a.s.l. close to Davos, Switzerland under diffuse illumination conditions, caused by a completely overcast sky, a situation that could be described as “worst case” for photogrammetry. We investigate the benefit of near infrared information ($\lambda > 830$ nm), compared to imagery acquired within the visible part of the electromagnetic spectrum ($\lambda = 400 - 700$ nm) and a combination of these two. We evaluate the accuracy of the different digital surface models (DSMs) qualitatively and quantitatively by applying differential Global Navigation Satellite System measurements with an expected accuracy better than 10 cm in x, y and z directions. The results of this study enable an assessment of the potential and the limitations of SfM photogrammetry for applications on snow-covered terrain.

1 Introduction

Snow type, snow depth and their spatiotemporal distribution in alpine regions are key parameters for many applications ranging from hydropower, drinking water supply, hydrological modeling, snow avalanche and flooding hazard mitigation, winter tourism, road and railway maintenance to ecological applications such as alpine or arctic habitat monitoring. In combination with snow density models, snow depth measurements can be calculated to snow water equivalent values (JONAS et al. 2009), describing the amount of water that is stored in the snowpack. Even though the spatial variability of alpine snow cover is known to be very high (SCHWEIZER et al. 2008; LOPEZ-MORENO & NOGUES-BRAVO 2006), ranging from nearly zero to several meters within short distances of a few meters.

Currently, snow depth is operationally measured using automated weather stations (AWS) or observers in the field (EGLI 2008) for various applications such as the Swiss avalanche warning or flood forecasting. The main issue of these measurements is that they are all acquired at point

¹ WSL Institute for Snow and Avalanche Research SLF, Flüelastrasse 11, 7270 Davos Dorf, Switzerland, E-Mail: [buehler, stoffel]@slf.ch

² Austrian Research Centre for Forests (BFW), Rennweg 1, 6020 Innsbruck, Austria, E-Mail: marc.adams@bfw.gv.at

³ Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland, E-Mail: [ruedi.boesch, christian.ginzler]@wsl.ch

locations and are not spatially continuous. This is a key issue for most applications, which are based on snow depth information as input, because the snow cover is spatially variable in particular within alpine terrain (SCHWEIZER et al. 2008).

To overcome this limitation, airborne (ALS) and terrestrial laser scanning (TLS) was successfully applied to map snow depth mapping over small catchments (e.g. DEEMS et al. 2013; GRÜNEWALD et al. 2010; MOTT et al. 2010; PROKOP 2008). With this method, high spatial resolution of better than 1 m and vertical accuracies better than 0.1 m can be achieved. However, data acquisition is time-consuming and often costly (BÜHLER et al. 2015a). UAS-based laser scanning has also been introduced is not yet fully operationally implemented. First devices exist but are very costly in particular considering the high risk of UAS-operation under alpine conditions, i.e. high winds and gusts, low temperature, lack of landing spots et cetera.

Photogrammetry can offer an economic and flexible alternative to laser scanning. First attempts to map snow depth photogrammetrically, using scanned aerial imagery, were already made 50 years ago (SMITH et al. 1967) and the topic was investigated in detail by CLINE (1993, 1994). However, their results suffer from image saturation and insufficient reference data leading them to the conclusion that photogrammetry has much potential but is not yet accurate enough for large-scale snow depth mapping. LEDWITH & LUNDEN (2010) used scanned aerial imagery to derive digital elevation models over glaciated and snow-covered areas in Norway. They report a mean accuracy of 2.8 m in comparison to differential Global Navigation Satellite System (dGNSS) transects, which is clearly too low for meaningful snow depth mapping in alpine regions.

With the current improvement in sensor technology and photogrammetric data processing, in particular the structure from motion (SfM) approach (KOENDERINK & VAN DOORN 1991; POLLEFEYS et al. 2004; WESTBOY et al. 2012; Crandall et al. 2013), digital photogrammetry developed to an efficient and highly accurate tool for surveying applications. First investigations on snow prove its applicability for snow depth mapping with sensors on manned aircrafts (LEE et al. 2008; BÜHLER et al. 2015a; NOLAN et al. 2015) to cover large areas and UAS (VANDER JAGT et al. 2015; BÜHLER et al. 2016; DE MICHELE et al. 2016; HARDER et al. 2016) to cover smaller areas with high spatial resolution. The reported accuracies range from approximately 0.30 m to 0.05 m root mean square error (RMSE). These recent results underline the potential of the proposed technology for spatially continuous snow depth mapping with high vertical accuracies that have the potential to complement or even substitute the currently applied point HS measurements. However, the legal regulations to fly UAS vary substantially for different countries or even states and counties. The regulations have to be checked carefully before data acquisitions with UAS can be planned. The regulations for Switzerland are listed under <https://www.bazl.admin.ch/bazl/en/home/good-to-know%20/drones-and-aircraft-models.html> and are comparably user friendly.

In this study we test SfM photogrammetry in a “worst case” scenario with very homogenous new snow cover and diffuse illumination caused by a closed cloud cover.

2 Test site, data acquisition and experimental setup

2.1 Test site Tschuggen, Davos GR, Switzerland

The test site Tschuggen is located at the bottom of the Flüela valley at an elevation of 1,940 m a.s.l., close to the timberline. This spot is well accessible even during the winter season, because the Flüela pass road is regularly cleared until this point. The high-alpine valley bottom features both quite flat alpine meadows and hilly alpine terrain. The main land cover is a mixture of bushes (mainly alpine rose, juniper and erica) containing steep rocky outcrops and sparse larch and pine trees. Only moderate HS variability can be expected at this site in an average winter season because it is usually not exposed to high winds. The mean slope angle of the test site is 19° ranging from 0 to 83° (BÜHLER et al. 2016).

On 8 January 2016 we had particular conditions: this particular date was immediately after a 0.10 m of fresh snow fall and the sky was completely overcast, leading to diffuse illumination (Fig. 1). We expect the combination of these two conditions to be “worst case” for photogrammetry on snow, as it leads to minimal contrast. We could visually not make out any contrast in most snow-covered areas. Therefore we expect to be able to investigate the limits of photogrammetry on snow with this dataset.



Fig. 1: The Falcon 8 flying at Tschuggen on January 8 2016. The sky is overcast and the new snow cover shows nearly no contrast.

2.2 UAS Falcon 8

As described in (BÜHLER et al. 2016), the Ascending Technologies (AscTec) Falcon 8 octocopter (Fig. 1) is equipped with a customized Sony NEX-7 camera. The Falcon 8 has been in serial production since 2009 and can be customized with different sensor systems. The system weighs 2.3 kg (incl. camera) and can be transported to remote locations fully assembled in a special

backpack, a prerequisite for most alpine applications. Onboard navigation sensors (Global Navigation Satellite System GNSS, Inertial Measurement Unit IMU, barometer and compass) and an adaptive control unit permit high positional accuracy and stable flight characteristics even in challenging conditions found in alpine environment with limited accessibility, low temperatures, high wind speeds and complex terrain.

The applied Sony NEX-7 system camera features a 24 MP APS-C CMOS sensor and is equipped with a small and lightweight Sony NEX 20mm F/2.8 optical lens (81 g). By removing the built-in near infrared filter, the camera sensor is also sensitive within the near infrared part of the spectrum. This allows us to mount the lens with different filters: visible colors (RGB, $\lambda = 550 - 770$ nm), undefined NIR (without filter, combined visual and NIR information) and near infrared only (NIR830, $\lambda > 830$ nm). The near-infrared sensitivity is expected to have advantages on snow-covered surfaces (BÜHLER et al. 2015b). The camera is connected to the Falcon 8 by a gimbal with active stabilization and vibration damping and is powered by the UAS battery. The viewfinder of the camera is transmitted to the ground control station as video signal and the basic camera functions such as the exposure time can be controlled from the ground.

2.3 Experimental setup

We acquired imagery in RGB, near infrared undefined (NIR) and near infrared ($\lambda > 830$ nm, NIR830) wavelengths immediately after each other (total of 104 images per data acquisition). Blurred pictures were eliminated manually. The illumination conditions were nearly constant. An image overlap of 70 % along-track and across-track was chosen, which is based on our experience, a good compromise between the amount of time required for data acquisition and quality of the resulting DSM. The flight elevation above ground was approximately 100 m, the air temperature was -6°C and the mean wind speed was 25 km/h, with wind gusts of up to 35 km/h from southeastern direction (measured at the automated weather station Flüela, 4 km from Tschuggen).

For the absolute orientation, 10 reference points (RPs), distributed over the entire test site, were used to georeference the RGB and NIR imagery. The RPs, blue crosses sprayed on the snow, were measured with a Trimble GeoXH differential GNSS with an expected accuracy of better than 0.10 m. As the RPs are not visible within the NIR830 imagery, because the spray seems to absorb the same amount of radiation as the surrounding snow, so we had to rectify the NIR830 imagery using RPs visually identified within the RGB orthophoto. We applied 11 checkpoints (CPs), which were not used for the rectification process (Fig. 2).

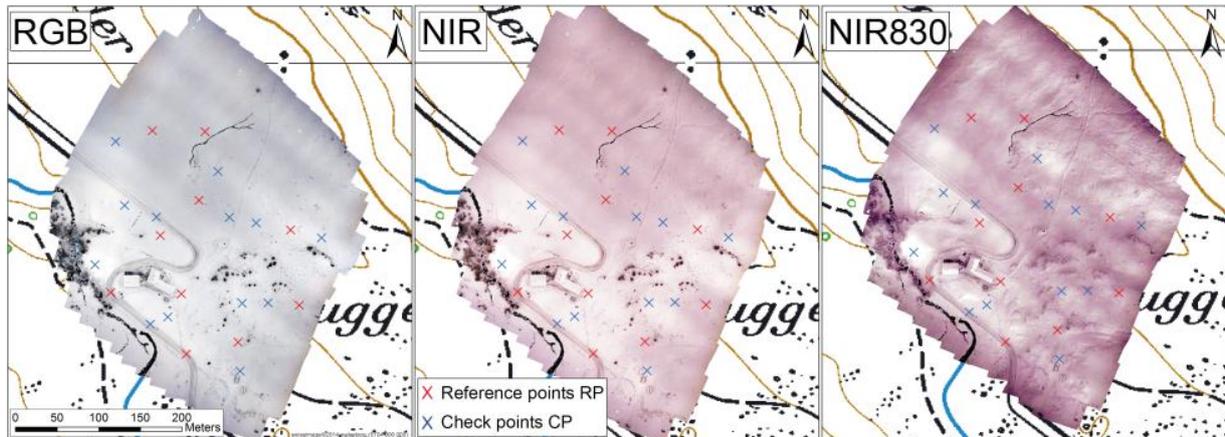


Fig. 2: RGB, NIR and NIR830 orthophotos of the test site Tschuggen with the applied reference (RP) and check points (CP).

The imagery was processed using Agisoft PhotoScan Pro v. 1.2.4. The imagery was aligned automatically and we used the default moderate filter to generate the point cloud. The DSMs were exported with a spatial resolution of 0.10 m using the PhotoScan interpolation algorithm. We classified the point clouds using the classification tool implemented in PhotoScan Pro into ground points and noise/unclassified. As we want to compare the different input imagery we keep all processing parameters constant.

3 Results and discussion

3.1.1 Orthorectification

The achieved average accuracy of the orthorectification process (calculated by the Photoscan Pro software) is 0.17 m ($x = 0.08$ m, $y = 0.08$ m, $z = 0.13$ m) for the RGB imagery and 0.16 m ($x = 0.09$ m, $y = 0.07$ m, $z = 0.11$ m) for the NIR imagery. In the NIR830 imagery, the sprayed points are completely invisible due to similar reflection characteristics as the surrounding snow, we only realized this fact during this study. Therefore we have to reference the NIR830 imagery applying reference points taken from the RGB orthophoto, close to the original RPs. The achieved accuracy based on the RGB reference points are 0.16 m ($x = 0.08$ m, $y = 0.11$ m, $z = 0.07$ m). These orientation accuracies are all in the same range, however we observe a small shift between the NIR830 and the other two datasets.

3.1.2 Qualitative DSM evaluation

We already generated DSMs on March 11 2015 under clear sky conditions and on a snow cover that did not get covered by new snow recently. Under such conditions only very few outlier and no holes occur within the DSM (Fig. 3). The achieved RMSE values of the UAS snow depth map compared to manual avalanche probe measurements is better than 0.10 m on short grass and rocks (Bühler et al. 2016). These results lead to the conclusion, that high-quality surface models can be generated on snow, if the illumination conditions are good (clear sky, high contrast) and the snow cover is not extremely homogenous.

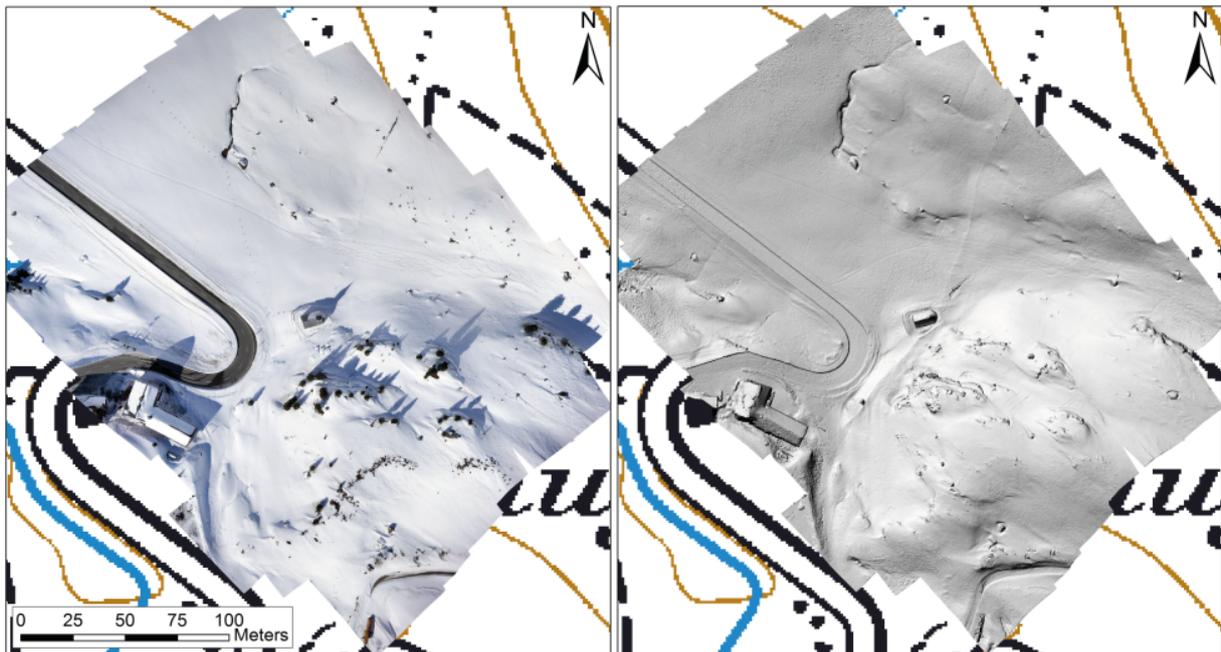


Fig. 3: Reference data acquisition on March 11 2015 under clear sky conditions several days after the last snowfall. Nearly no noise can be identified in the hillshade of the DSM (right) except at the very edge where the image overlap gets insufficient.

In contrast to the data from March 11 2015, the DSM generated from the data acquired on January 8 2016 show obvious errors within the homogenous parts of the snow cover. Large interpolated areas, outliers up to several meters and noisy sections occur. The RGB and NIR data result in similar amounts of corrupted surfaces. But the NIR830 data result in considerably less corrupted area and therefore in a better DSM result, even though it is also not completely free of noisy errors (Fig. 4).

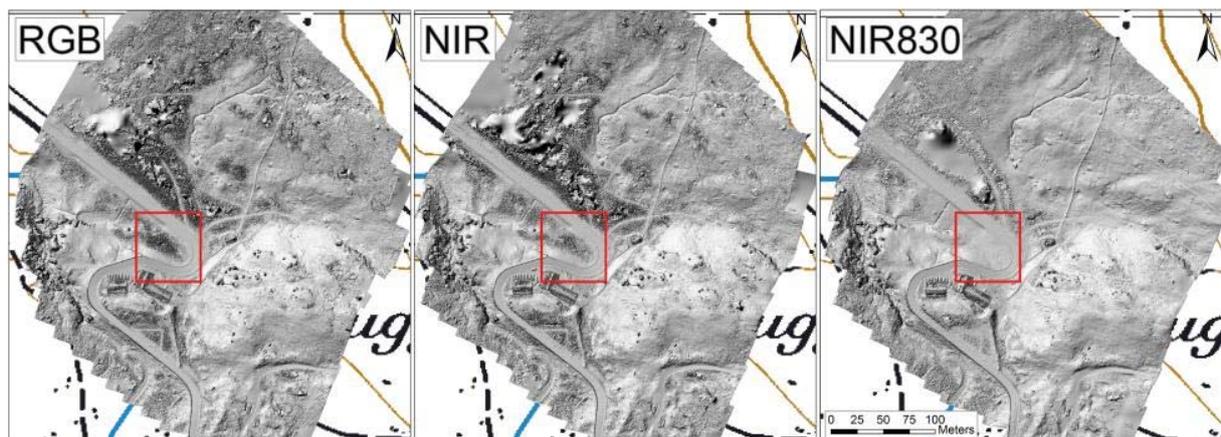


Fig. 4: Hillshades of the DSMs generated from the data acquired under “worst case” conditions. The homogenous snow cover results in very noisy surfaces with amplitudes of up to several meters or even holes, which have to be interpolated. This is most obvious within the northwest part of the test site above the road. The hilly section in the southeast part of the test site shows a quite smooth surface in all bands.

The effect of the different wavelength gets better visible by zooming into a very homogenous snow cover spot at the U-turn curve of the road. To visualize the effects we draw a line profile across this snow spot (Fig. 5). The profile and the hillshade views reveal that the quality of the DSM improves considerably from RGB over NIR to NIR830. This result is surprising on the first sight as the visual impression is best for RGB and worst for NIR830 imagery. However if the imagery is enhanced, the NIR830 shows more contrast and details within the homogenous snow cover as it is already demonstrated by BÜHLER et al. 2015b.

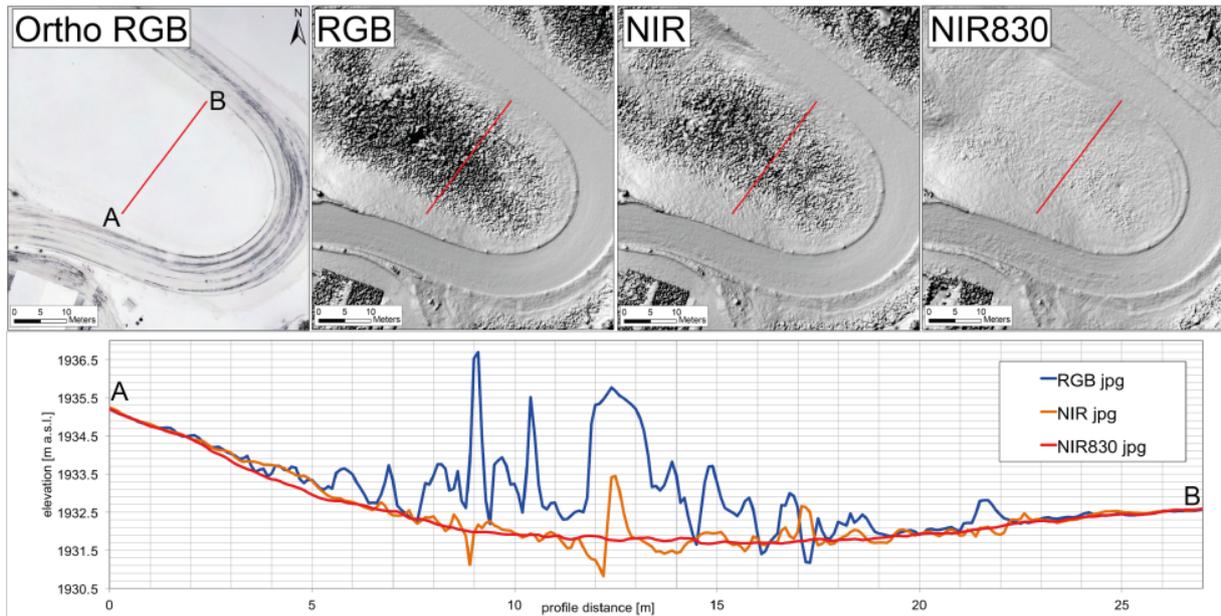


Fig. 5: Profiles across a homogenous snow spot for the different unfiltered DSMs.

The qualitative 3D analysis of the point clouds around the chapel in the middle of the test site reveals that large outliers of several meters occur within the homogenous snow covered parts. We classify the point clouds into ground points, noise and other points (Fig. 6). By only using the ground points for DSM generation, the large outliers are eliminated. However there are still noisy errors with amplitude in the range of 0.5 m within the ground points. Furthermore not all corrupted points are correctly classified, in particular where the point density is low, leading to large areas with wrong elevation values, interpolated within low quality points (Fig. 7).

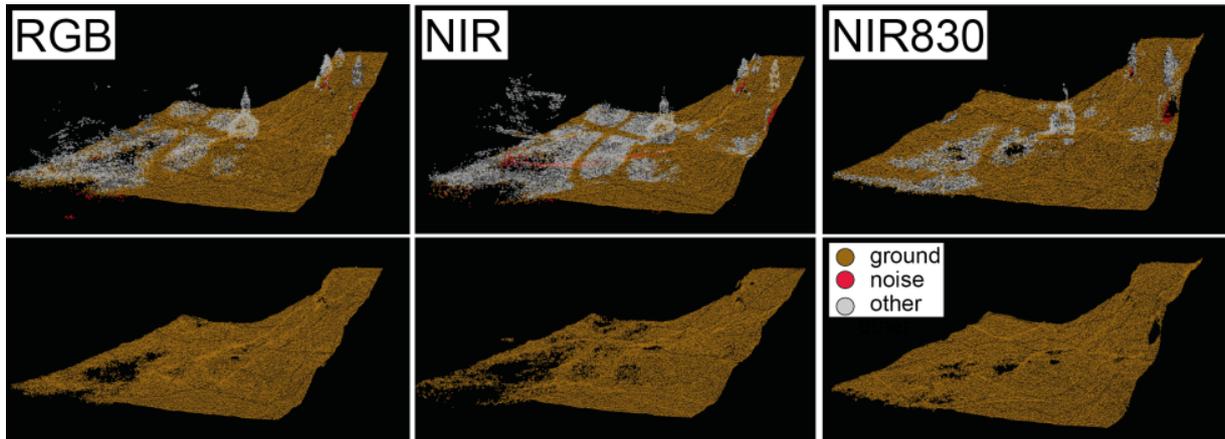


Fig. 6: Point clouds of the different data sets around the chapel in the middle of the test site. The lower tab shows only the points classified as ground.

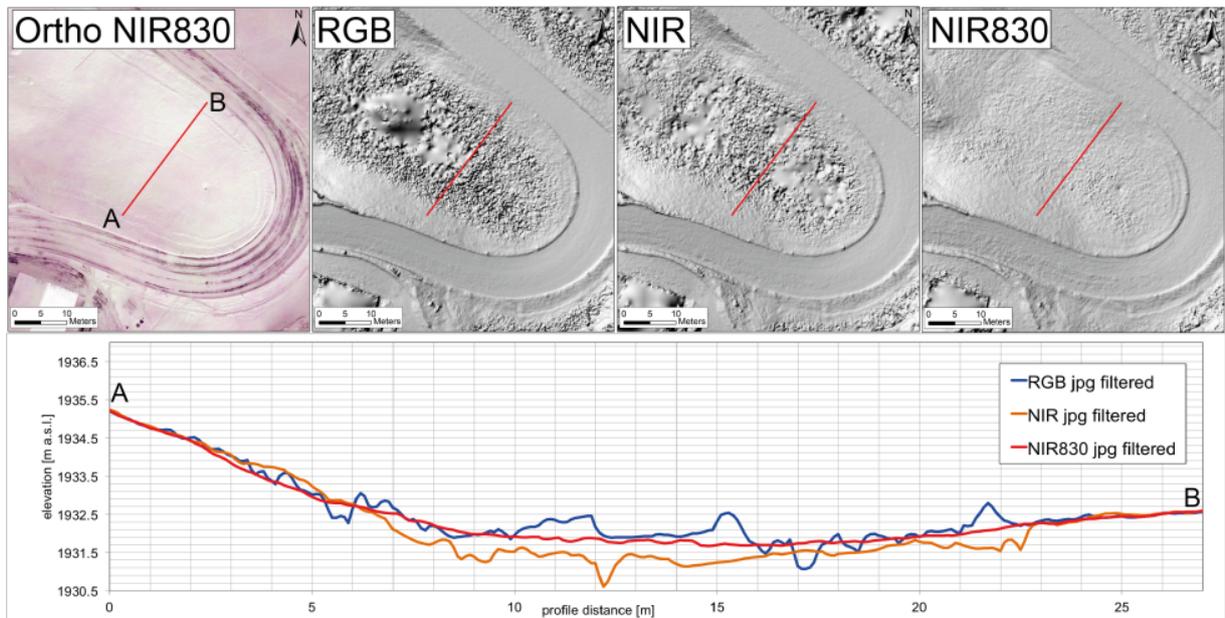


Fig. 7: Profiles across a homogenous snow spot for the different filtered DSMs.

Even though the large outliers are eliminated, big differences in DSM quality still remain within the different bands. The quality of the NIR830 data is again considerably better than the NIR and RGB DSMs. These results clearly indicate the additional value of near infrared bands for surface model generation in particular within homogenous snow spots.

3.1.3 DSM precision at the check point (CP) locations

We analyze the elevation values at the 11 CP locations comparing them to the GNSS elevation (accuracy < 10 cm) to quantitatively compare the different DSMs. Here the picture is different from the qualitative assessment: the RMSE of the NIR830 DSM is approximately 30% higher than for RGB and NIR (Tab 1.) However, the number of points gets higher as more details on the

snow cover are detectable in the near-infrared band. Additionally more points are produced in RGB and in NIR that are classified as noise or other points than ground. These noise points as well as unclassified outliers lead to the very noisy surfaces of the unfiltered DSMs (Fig. 5, Fig. 6).

We suppose that the higher RMSE of the NIR830 DSM is caused by the different orthorectification. As the sprayed crosses are invisible within the NIR830 imagery we had to apply different reference points, which lead to a slight shift and probably cause the larger deviation from the GNSS elevation at the checkpoint locations. As we had to deploy the RPs and CPs before the flight, our ski tracks disturbed the homogenous snow cover and generated contrast. PhotoScan Pro uses this contrast to successfully match points without outliers. Therefore the RMSE values are very low at the checkpoint locations. This result indicates that the results of RGB and NIR are good if there is enough contrast present. However it limits the quantitative quality assessment to areas with high contrast, which was not the main aim of this study.

Tab. 1: Summarized errors of the 11 different DSMs compared to the GNSS elevation at the check point locations and description of the different point clouds and their classification.

	Mean error (bias)	Standard deviation error	RMSE	N° of point total [mio]	ground points [%]	noise [%]	unclassified [%]
RGB jpg	-0.02 m	0.12	0.18 m	68.96	92.24	0.56	7.20
NIR jpg	-0.03 m	0.11	0.16 m	71.37	83.40	0.45	16.15
NIR830 jpg	-0.09 m	0.19	0.27 m	76.80	93.83	0.19	5.98

4 Conclusions

Recent investigations by BÜHLER et al. (2015), NOLAN et al. (2015), VANDER JAGT et al. (2015), BÜHLER et al. (2016), DE MICHELE et al. (2016) and HARDER et al. (2016) showed that the longstanding assumption, photogrammetry is not able to generate high quality digital surface models on snow covered terrain, is incorrect under favorable illumination conditions and over weathered snow surfaces. Photogrammetry can successfully be applied to map snow depth with high spatial resolution even within high alpine terrain.

To investigate the limitations of photogrammetry on snow we explore a “worst case” scenario: homogenous new snow cover and diffuse illumination caused by a completely overcast sky. Under these conditions, the human eye is not able to detect contrast over large parts of the snow cover. The resulting DSMs, calculated based on the RGB and the NIR imagery are, as expected, corrupted by noisy errors with amplitudes up to 10 m and large holes where no points can be matched by the commercial software Agisoft PhotoScan Pro. But if we apply near infrared bands (NIR), the results get better and if we apply near infrared input (NIR830) only, the resulting DSMs are significantly better (Fig. 5). Errors still occur over the most homogenous snow patches but large parts can be mapped surprisingly well. This result supports the assumption by BÜHLER et al. (2015b) that near infrared bands bring significant additional value for point matching over snow covered areas.

The results of this study indicate that digital photogrammetry is able to generate high quality digital surface models and orthophotos even under suboptimal illumination conditions and over very homogenous new snow cover if near infrared bands are applied. The extremely homogenous parts of the snow cover cannot be mapped satisfactorily but large part of the test site can. This opens the door for various applications where spatial continuous snow cover and snow depth information is requested.

5 References

- BÜHLER, Y., MARTY, M., EGLI, L., VEITINGER, J., JONAS, T., THEE, P. & GINZLER, C., 2015a: Snow depth mapping in high-alpine catchments using digital photogrammetry. *The Cryosphere* **9**, 229-243.
- BÜHLER, Y., MEIER, L. & GINZLER, C., 2015b: Potential of operational, high spatial resolution near infrared remote sensing instruments for snow surface type mapping. *Geoscience and Remote Sensing Letters, IEEE*, **12** (4), 821-825.
- BÜHLER, Y., ADAMS, M. S., BÖSCH, R. & STOFFEL, A., 2016: Mapping snow depth in alpine terrain with unmanned aerial systems (UAS): potential and limitations. *The Cryosphere* (in press).
- CLINE, D.W., 1993: Measuring alpine snow depths by digital photogrammetry: Part 1. conjugate point identification. *Proceedings of the Eastern Snow Conference*, Quebec City.
- CLINE, D.W., 1994: *Digital Photogrammetric Determination of Alpine Snowpack Distribution For Hydrologic Modeling*. Colorado State University, CO, USA.
- DEEMS, J.S., PAINTER, T.H., & FINNEGAN, D.C., 2013: Lidar measurement of snow depth: A review. *Journal of Glaciology* **59**, 467-479.
- DE MICHELE, C., AVANZI, F., PASSONI, D., BARZAGHI, R., PINTO, L., DOSSO, P., GHEZZI, A., GIANATTI, R. & DELLA VEDOVA, G., 2016: Using a fixed-wing UAS to map snow depth distribution: an evaluation at peak accumulation. *The Cryosphere* **10**, 511-522.
- EGLI, L., 2008: Spatial variability of new snow amounts derived from a dense network of Alpine automatic stations. *Annals of Glaciology* **49**, 51-55.
- GRÜNEWALD, T., SCHIRMER, M., MOTT, R., & LEHNING, M., 2010: Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment. *The Cryosphere* **4**, 215-225.
- HARDER, P., SCHIRMER, M., POMEROY, J. & HELGASON, W., 2016: Accuracy of snow depth estimation in mountain and prairie environments by an unmanned aerial vehicle. *The Cryosphere Discuss.*, 1-22.
- JONAS, T., MARTY, C. & MAGNUSSON, J., 2009: Estimating the snow water equivalent from snow depth measurements in the Swiss Alps. *J. Hydrol.* **378**, 161-167.
- KOENDERINK, J.J. & VAN DOORN, A.J., 1991: Affine structure from motion. *Journal of the Optical Society of America, A, Optics and Image Science* **8** (2), 377-385.
- LEDWITH, M. & LUNDÉN, B., 2010: Digital photogrammetry for air-photo-based construction of a digital elevation model over snow-covered areas - Blamannsisen, Norway. *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography* **55** (4), 267-273.

- LEE, C.Y., JONES, S.D., BELLMAN, C.J. & BUXTON, L., 2008: DEM creation of a snow covered surface using digital aerial photography. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **37** (B8).
- LOPEZ-MORENO, J.I. & NOGUES-BRAVO, D., 2006: Interpolating local snow depth data: An evaluation of methods. *Hydrological Processes* **20**, 2217-2232.
- MOTT, R., SCHIRMER, M., BAVAY, M., GRÜNEWALD, T., & LEHNING, M., 2010: Understanding snow-transport processes shaping the mountain snow-cover. *The Cryosphere* **4**, 545-559.
- NOLAN, M., LARSEN, C. & STURM, M., 2015: Mapping snow depth from manned aircraft on landscape scales at centimeter resolution using structure-from-motion photogrammetry. *The Cryosphere* **9**, 1445-1463.
- POLLEFEYS, M., VAN GOOL, L., VERGAUWEN, M., VERBIEST, F., CORNELIS, K., TOPS, J. & KOCH, R., 2004: Visual modeling with a hand-held camera. *International Journal of Computer Vision* **59**, 207-232.
- PROKOP, A., 2008: Assessing the applicability of terrestrial laser scanning for spatial snow depth measurements. *Cold Regions Science and Technology* **54**, 155-163.
- SCHWEIZER, J., KRONHOLM, K., JAMIESON, J.B. & BIRKELAND, K.W., 2008: Review of spatial variability of snowpack properties and its importance for avalanche formation. *Cold Reg. Sci. Technol.* **51**, 253–272.
- SMITH, F., COOPER, C. & CHAPMAN, E., 1967: Measuring Snow Depths by Aerial Photography. *Proceedings of the Western Snow Conference*, Boise, Idaho, USA.
- VANDER JAGT, B., LUCIEER, A., WALLACE, L., TURNER, D. & DURAND, M., 2015: Snow Depth Retrieval with UAS Using Photogrammetric Techniques, *Geosciences*, 5, 264-285, 2015.
- WESTOBY, M.J., BRASINGTON, J., GLASSER, N.F., HAMBREY, M.J. & REYNOLDS, J.M., 2012: ‘Structure-from-Motion’ photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **179**, 300-314.

Geoprodukte im Aufwind – die GeoApp.UAS

DORIS HERMLE¹ & IRMGARD RUNKEL¹

Zusammenfassung: Einsatzbereiche und Möglichkeiten von UAS nehmen stetig zu und damit auch die Erschließung bisher nicht genutzter Bereiche. Gründe dafür sind geringer werdende Anschaffungskosten und die leichtere Handhabung von UAS. Damit ergeben sich einerseits neue Einsatzbereiche in bisher GIS-fernen Märkten, andererseits die Forderung nach einer unkomplizierten Prozessierung, ohne Einbußen an Qualität und Verarbeitungszeit. Die cloudbasierte Lösung, die GeoApp.UAS von GEOSYSTEMS, erleichtert die komplexe Prozessierung der UAS-Daten innerhalb eines automatisch ablaufenden Workflows zu Orthomosaik, digitalem Oberflächenmodell und Punktwolke. Zudem können die Ergebnisse komprimiert und als Webdienst bereitgestellt werden. Zwei gegensätzliche Beispiele stellen die Breite der Anwendungsmöglichkeiten und die einfache Handhabung dar.

1 Einleitung

Unbemannte fliegende Systeme (UAS, unmanned aerial systems) sind flexibel und schnell einsetzbar, die klaren Vorteile der raschen Informationsgewinnung über große Flächen, schwer zugängliche oder gar gefährliche Gebiete sind bekannt und erklären den wachsenden Markt (BEER 2016, AMATO 2015). Bildaufnahmen sind vordergründig sehr einfach geworden, die Herausforderung für eine sinnvolle und professionelle Verwendung liegen in der Qualität und Geschwindigkeit, diese Daten kundenfokussiert zu prozessieren und bedarfsgerecht auszuliefern. Hier ist nach wie vor Fachwissen und Fachsoftware gefragt. Dies bremst die offensichtliche Flexibilität von UAS-Bilddaten aus. Dieser Artikel stellt mit der GeoApp.UAS eine Lösung vor, die eine einfache, schnelle Verarbeitung der Daten und Verbreitung von Ergebnissen bietet – analog zur Leichtigkeit ein UAS zu fliegen.

2 Vorstellung GeoApp.UAS

Der Einsatz von UAS zur raschen Gewinnung räumlicher Informationen aus der Luft nimmt in den letzten Jahren stark zu. Neue Einsatzbereiche und Möglichkeiten werden entdeckt, auch der Bedarf und die Nachfrage nach Bildinformation für vertikale und horizontale Auswertungen steigen. Die Zahl der verkauften Drohnen nimmt kontinuierlich zu, bei gleichzeitig geringer werdenden Anschaffungskosten. Mehr und mehr Firmen integrieren UAS-Daten in ihre Geschäftsmodelle. Damit einhergehend sehen sie sich mit einer hohen Anzahl von Bilddaten im Bereich horizontal-räumlicher Geodatenerfassung und mit zunehmenden Anforderungen und Erwartungshaltungen ihrer Kunden konfrontiert.

Die Ableitung von Geoprodukten aus UAS-Daten soll nicht nur schnell und günstig erfolgen, sondern automatisiert oder - vermehrt - von nicht-fachkundigen Personen ausgeführt werden

¹ GEOSYSTEMS GmbH, Riesstraße Nr. 10, D-82110 Germering, E-Mail: d.hermle@geosystems.de

können. Daher ergibt sich ganz klar der Anspruch nach robuster und einfacher Handhabung, als auch Prozessierung. Daraus leitet sich folgender Bedarf ab: Die Bereitstellung von Anwendungen, welche systematisch in einem automatisch ablaufenden Batch-Prozess die Effizienz der Bilddatenprozessierung erhöhen.

Die GeoApp.UAS ist die cloud-basierte Prozessierung von Geobasisprodukten aus generischen UAS Daten. Es werden Orthomosaik, digitale Höhenmodelle und Punktwolken abgeleitet, ohne dass dafür photogrammetrisches Fachwissen beim Nutzer erforderlich ist. Grundvoraussetzung hierfür ist eine Befliegung unter Berücksichtigung der zwingenden Parameter, wie man sie aus der Luftbilddatengewinnung kennt, wie zum Beispiel der lückenlose Bildverband mit ausreichender Längs- und Querüberdeckung, sowie GPS-information über die Bildposition zum Aufnahmezeitpunkt.

3 Methodische Beschreibung der GeoApp.UAS

Entwickelt von der GEOSYSTEMS GmbH, ist die GeoApp.UAS eine Anwendung, die als Smart M.App online auf der Hexagon Plattform abonniert werden kann. Diese Plattform nutzt die Serverumgebung von AWS (Amazon Web Service). Die fertigen Endprodukte können in der Cloud weiterverarbeitet, heruntergeladen oder als OGC-Webdienst (Streamingdienst) bereitgestellt werden.

Die Verarbeitungsschritte zur Ableitung von Geoprodukten sind auf Anforderungen von UAS-Daten abgestimmt (vgl. Abb. 1):

1. Berechnung und Verknüpfung gemeinsamer Bildpunkte
2. Identifikation der Kameraposition für jedes Bild und Feintuning der Kamerakalibrierung
3. Berechnung einer Interims-Punktwolke
4. Oberflächenberechnung durch polygonale 3D-Vermaschung
5. Berechnung dichter Punktwolke mit RGB-Kodierung im LAS-Format
6. Ableitung von Orthomosaik (sog. Trueorthomosaik) und Oberflächenmodell
7. Optional, direkte Komprimierung zu ECW und JP2
8. Verwalten in einem Geodatenmanagementsystem
9. Bereitstellen als Webdienst (OGC-konform oder als Streamingdienst)

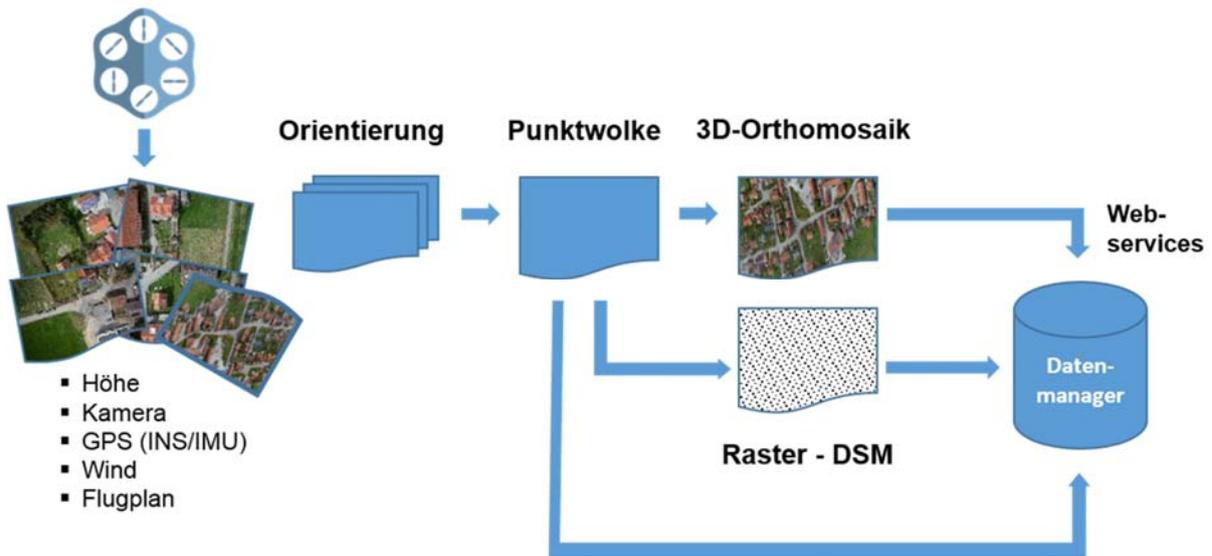


Abb. 1: Schema der Ableitung von Punktwolke, DSM und 3-D Orthomosaik aus UAS-Daten mit der GeoApp.UAS.

Die Ergebnisse können

- als Datensatz im geeigneten Format (*.tif, *.img, *.jp2, *.ecw und *.las) zur lokalen Verwendung (z.B. GIS) heruntergeladen werden,
- innerhalb der Hexagon Plattform mit einem geeigneten webbasierten Analysetool, z.B. dem M.App Reader oder M.App X, interaktiv ausgewertet und annotiert werden oder
- als OGC-konformer Webservice bereitgestellt werden.

Voraussetzung für eine serverzentrierte Auswertung ist, dass die Bild- und Orientierungsdaten auf den Server bereitgestellt werden. Hierfür stellt die Hexagon Smart M.App Plattform mit der M.App Chest ein einfaches Upload Tool zur Verfügung. Für die vollautomatische Prozessierung greift die GeoApp.UAS auf diese Daten zu und führt alle nötigen Berechnungsschritte nahtlos und vollautomatisch aus, ohne dass ein erneutes Starten des Prozesses von Nöten ist. Gerechnet wird dabei innerhalb einer leistungsstarken Serverumgebung mit dem AWS (Abb. 2). Die Ergebnisse stehen nach der Berechnung in der M.App Chest zur weiteren Verwendung bereit.

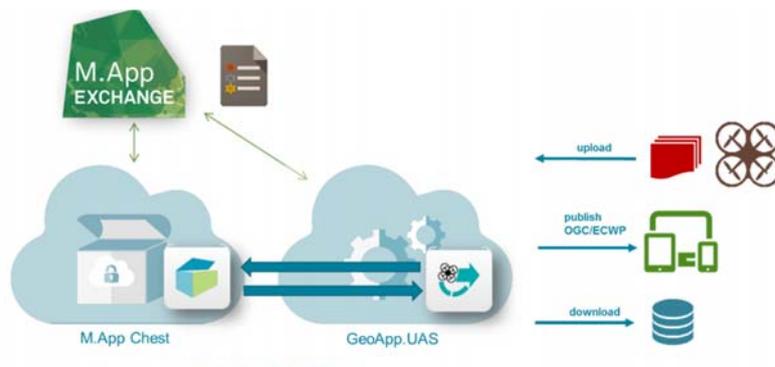


Abb. 2: GeoApp.UAS Workflow: Bildaufnahmen mit einem UAS, anschließendes hochladen der Daten in die Cloudumgebung, anpassen des Orientierungsformats und automatisches prozessieren zu Geodaten (Punktwolke, Orthomosaik, DSM) mit der Weitergabe, Download und Einbindung durch einen OGC Service.

Die GeoApp.UAS minimiert die für die Prozessierung nötigen Parameter und falls möglich, sind diese mit praxisorientierten Voreinstellungen vorbelegt. Eingebunden in die Welt der Hexagon Smart M.Apps, können auf der Geoplattform weitere Apps abonniert werden. Somit hat ein GIS Neuling mit der GeoApp.UAS als Basis die Möglichkeit, zusätzliche Bausteine anzuhängen, um eine für ihn fachspezifische Kombination zu schaffen (Abb. 3).

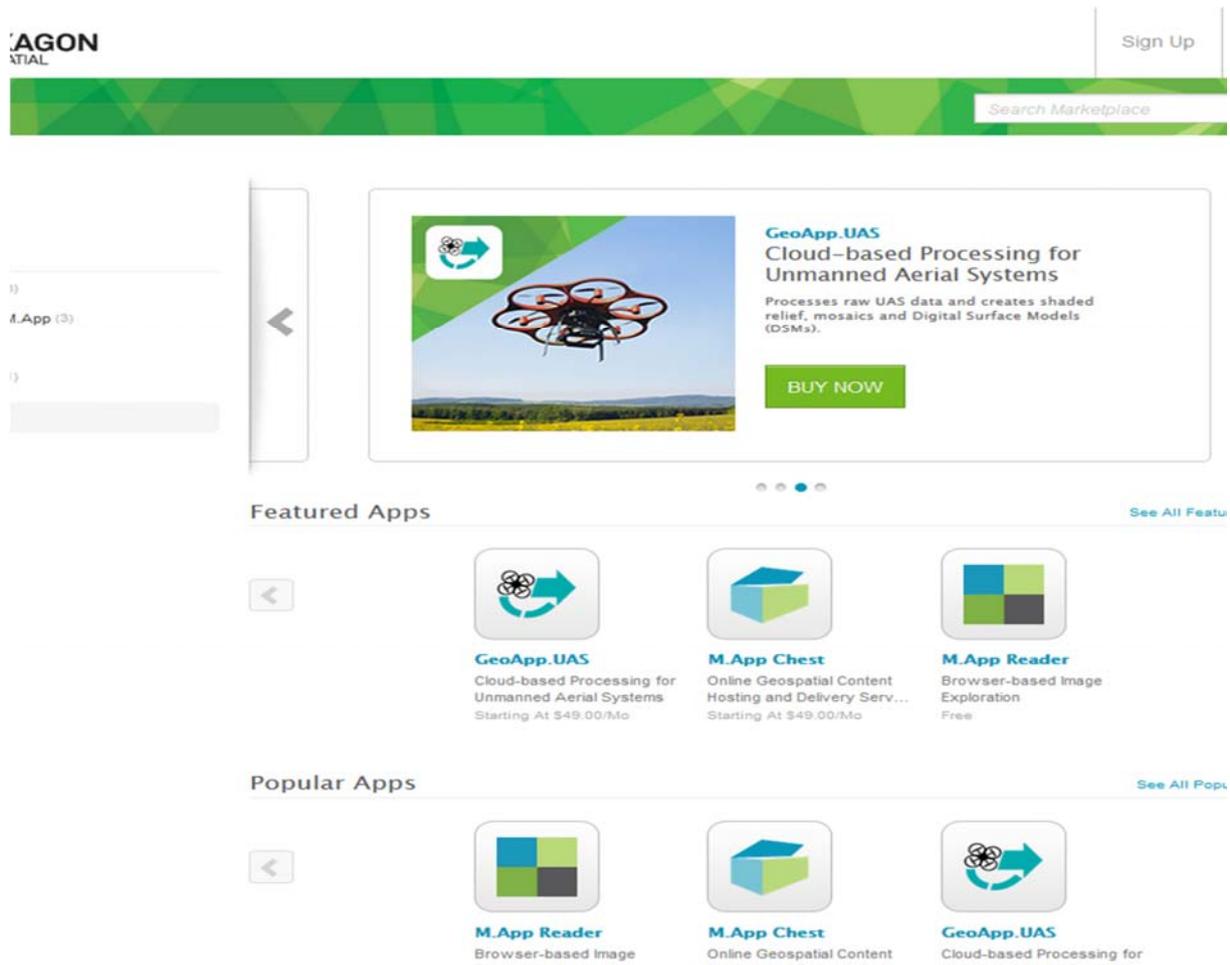


Abb. 3: Der Hexagon Marketplace: die Geoplattform, auf welcher verschiedene Apps als Bausteine gewählt und abonniert werden können.

Für Geschäftsprozesse, die einen Upload der Daten nicht erlauben, kann der GeoApp.UAS-Workflow auf einem gesicherten Server innerhalb eines Unternehmens betrieben werden, eine kundenspezifische Anpassung an die vorhandene Umgebung innerhalb einer eigenen Enterprise-Lösung ist möglich. GEOSYSTEMS bietet mit IMAGINE UAV zusätzlich die Option einer Desktop-Version, die skalierbar auf ERDAS IMAGINE zur Verfügung steht und weitere Möglichkeiten für Experten und Spezialanwendungen bietet (Abb. 4).

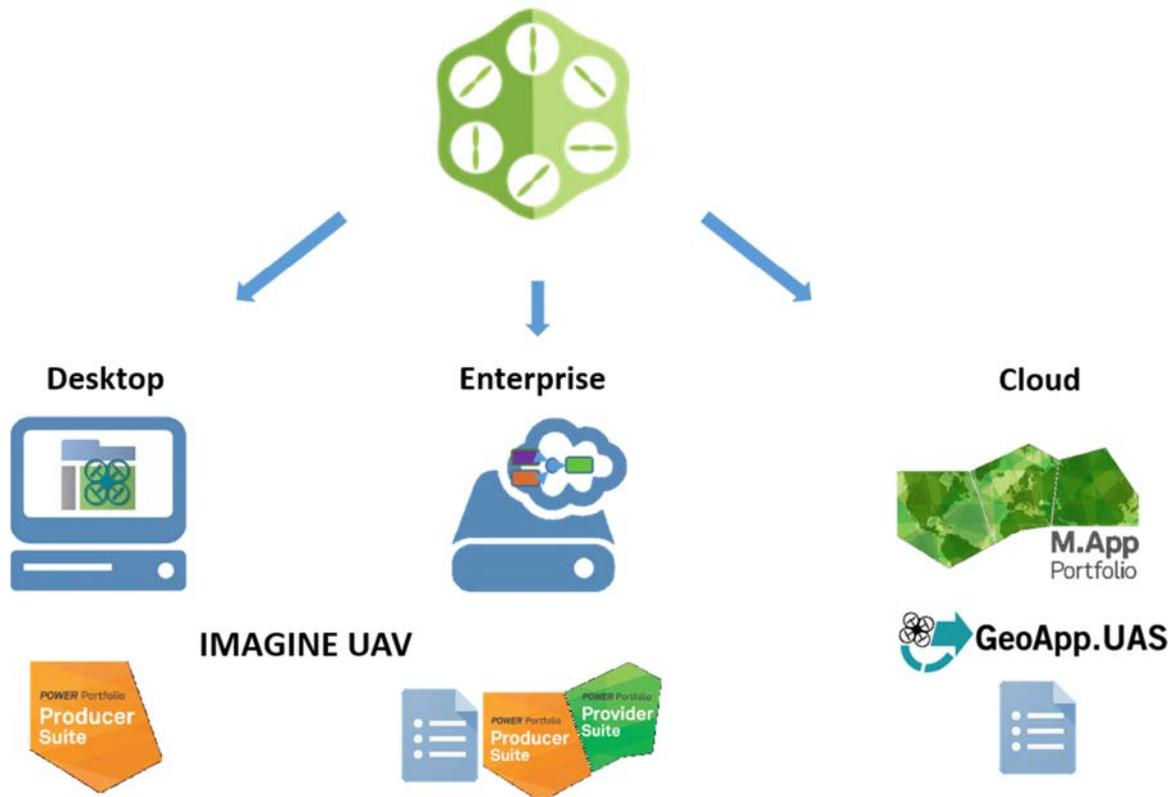


Abb. 4: Die drei UAS-Lösungen von GEOSYSTEMS für die Bearbeitung von UAS Aufnahmen: ERDAS IMAGINE UAV als Desktopanwendung, als Enterprise Lösung mit eigener Serverumgebung oder als Cloud Lösung mit der GeoApp.UAS auf der Hexagon Smart M.App Plattform.

4 Vorteile der GeoApp.UAS

Die Nutzung von UAS-Bilddaten in unterschiedlichsten Geschäftsbereichen legt nahe, dass die zur Prozessierung nötige Fachexpertise und Fachsoftware dem Nutzer durch eine einfache und robuste Bedienung und Prozessierung abgenommen wird. Die neuen Herausforderungen für die Prozessierung von UAS Daten sind eine unterbrechnungsfreie und damit zügige Verarbeitung des erhöhten Datenaufkommens, ohne dabei Qualitätseinbußen, wie beispielsweise geringe Qualität der Punktwolke, hinnehmen zu müssen. Gleichzeitig soll dies einfach und unkompliziert erfolgen.

Die Kernfunktionen der GeoApp.UAS sind der photogrammetrisch robuste, effiziente und automatisch durchgeführte Arbeitsablauf, der gleichzeitig flexibel ist durch die Skalierbarkeit bei Datengröße sowie Genauigkeit, und damit benutzerspezifisch angepasst werden kann. Der Arbeitsablauf kann standortunabhängig, von überall aus mit Internetzugang gestartet und überwacht werden. Die Wahl bleibt auf Seiten des Anwenders, ob Punktwolke, Orthomosaik und Höhenmodell gerechnet werden sollen oder ob nur eines der drei Produkte generiert wird. Eine optionale Datenkompression der Bilddaten und Punktwolken reduziert merklich die Datenmenge und entlastet die Datenspeicherung (weniger Plattenplatz, weniger Kosten). Die Bereitstellung

als Webdienst ist für viele Nachnutzungen der geforderte Service, sei es für die Integration in Web-GIS, in Geoportale oder auch in Leitzentralen.

Für den generischen Ansatz, alle UAV-Typen verarbeiten zu können, die mit proprietären Aufnahmespezifikationen kommen ist es nötig, eine universelle Schnittstelle für Orientierungen anzubieten. Mit dem *Orientation Editor* können alle derzeit bekannten Orientierungsformate, zum Beispiel GPX-Tracks, EXIF-Informationsdaten, GPS/GNSS RTK-Daten verwendet werden. Berücksichtigt werden die Rotationswinkel (RPY/OPK), sie sind aber nicht zwingend notwendig, i.d.R. reichen die Positionsinformationen aus XYZ-Koordinaten aus.

Ein wesentlicher Vorteil der Prozessierung in der Cloud ist die Hardwareunabhängigkeit. Zum einen ist man nicht an die Verfügbarkeit eines Computers und dessen Rechenkapazität gebunden, zum anderen können gleichzeitig mehrere Berechnungen durchgeführt werden und damit rascher Daten für die Weitergabe bereitgestellt werden. Durch die unterbrechungsfreie, automatische Berechnung mit der Möglichkeit der zeitgleich multiplen Prozessierung, ergibt dies für den Anwender einen optimierten und zeiteffizienten Workflow. Zugleich ist die Ableitung der UAS Daten skalierbar, was die Menge und Größe der Eingangsdaten und die Qualitätsstufe der gewünschten Ergebnisse betrifft. Dies hat letztendlich eine Auswirkung auf die Rechnerleistung, damit auf die Rechenzeit und die Kosten. Da der Nutzer für die reine Prozessierungszeit zahlt (es werden unterschiedlich Paketierungen angeboten), hat er volle Kostenkontrolle und klare Kostenübersicht. Durch diese Flexibilität an die jeweiligen Bedürfnisse, ist der GeoApp.UAS Workflow kosteneffektiv. Zusätzliche Vorteile ergeben sich aus der Reduzierung der Infrastruktur, hohe Anschaffungskosten für leistungsstarke Rechner sowie der Entfall derer Wartung. Des Weiteren kann der Nutzer durch die Cloudprozessierung flexibel auf die Auftragslage reagieren: Rechenprozesse werden kontinuierlich und gegebenenfalls parallel abgearbeitet, womit Leerlauf- und Wartezeiten entfallen.

Ein anderer positiver Effekt ist die Reduzierung von Verwaltung und Administration, da keine Kosten für Lizenzierung und Softwarewartung, Einschulung oder Personalkosten für einen Experten anfallen. Die Cloudumgebung gewährleistet Sicherheit, Redundanz und Backup der Daten und ist somit ausfallsicher. Insgesamt erleichtert diese einfache Handhabung den Umgang mit UAS-Daten für die Geodatenerstellung und verschafft Befliegungsunternehmen Eigenständigkeit und Unabhängigkeit von Unterauftragnehmern.

5 Kommende Herausforderungen im B2B-Bereich

Raumrelevante Themen finden zunehmendes Interesse in bisher GIS-fernen, vertikalen Märkten. Gerade die Flexibilität von UAS-Aufnahmen erregt hier besonderes Interesse und rückt zunehmend in den Fokus deren Geschäftsprozesse. Das Potential von UAS-Aufnahmen wird gesehen und will genutzt werden, doch der Hemmschuh ist oftmals die Komplexität der nötigen Bildverarbeitung und Photogrammetrie, die für ein augenscheinlich einfaches Ergebnis oder Produkt notwendig ist. Somit erschließt sich mit der GeoApp.UAS ein neuer Kreis an Interessenten und Nutzern, der jedoch keine bis wenige Grundkenntnisse der Geoinformatik mit sich bringt. Diese Herausforderungen betreffen Märkte, die wiederum andere Märkte bedienen: es zielt auf den b2b-Bereich ab. Auf dem eigenen Fachgebiet Experten, brauchen diese ein einfaches und unkompliziertes Werkzeug, das darüber hinaus flexibel und kosteneffektiv

arbeitet. Die einfache Struktur der GeoApp.UAS kommt dem nach, da weder Expertenwissen, Einschulung noch lange Einarbeitungszeit notwendig ist. Da neben dem GIS-Experten die Infrastrukturkosten für Hardware, Software, Lizenzierung und Datenmanagement entfallen, ist nicht nur der Einstieg leicht, sondern es sind auch die anfänglichen Kosten gering. Folglich können neue Ergebnisse entstehen, mit einer hohen Qualität was das Messen und Erfassen betrifft.

6 Anwendungsbeispiele

Im Folgenden werden zwei Beispiele aus dem Alltag vorgestellt, bei denen die GeoApp.UAS zum Einsatz kommt. Einerseits eine Studie, bei der terrestrische Aufnahmen berechnet werden, andererseits die Vorstellung eines neuen Einsatzbereiches mit klassischen UAS Nadiraufnahmen durch einen Laien.

6.1 Beispiel 1: Forstwirtschaft: Präzise Erfassung von Wildschäden

Wildschäden durch Wildschweine an Wiesen und Weiden verursachen hohe Forderungen von Seiten der Landwirte, wobei Waldbesitzer, Jagdgenossenschaften und Jagdpächter die Kompensationszahlungen aufgrund der Wildschadensersatzpflicht leisten müssen. Die konventionelle Begutachtung erfolgt zu Fuß mit der Foto- und Videoaufnahme, sowie mit Hubschraubern. Diese Methoden sind kostenintensiv, ungenau und oft aufwändig. Der alternative Einsatz einer Drohnenbefliegung hat mehrere Vorteile. Neben der genauen Verortung und effektiven Protokollierung können die resultierenden Geodaten für eine bessere Abschätzung der betroffenen Flächen genutzt werden und dies bei weniger Aufwand und zu wesentlich geringeren Kosten.



Abb. 5: Wildschaden an einer Wiese durch Wildschweine

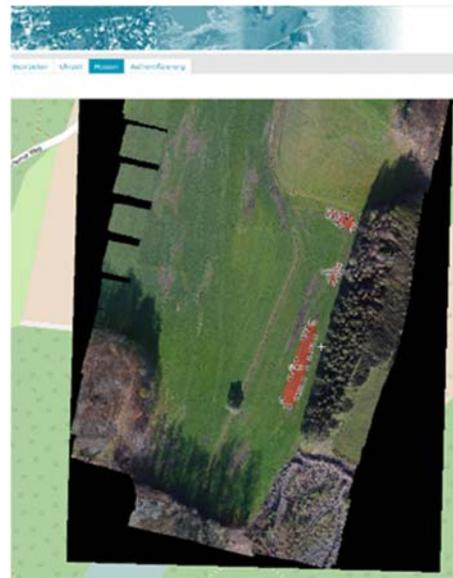


Abb. 6: Orthomosaik als Ergebnis des GeoApp.UAS Workflows. Die betreffenden Felder sind mit roten Polygonen versehen.

Mithilfe einer preisgünstigen Drohne wurde ohne große Vorkenntnisse vom Waldbesitzer selbst die betroffene Fläche überflogen, direkt im Anschluss daran die UAS-Daten auf in die Cloud geladen und mit der GeoApp.UAS ein Orthomosaik erzeugt. Somit konnten die Schäden lagegenau erfasst und korrekt berechnet werden (Abb. 6).

6.2 Beispiel 2: Baustellenüberwachung, Beispiel A96

Eine Studie zur Prozessierung terrestrischer Bilddaten mit der GeoApp.UAS wurde an einer Baustelle an der A96 durchgeführt. Sie zeigt eine temporäre Umgehungsstraße über welche zwei Brücken führen. Für diese Untersuchung wurde eine Digitalkamera des Types ‚Canon EOS 400D DIGITAL‘ verwendet und mit dem ‚IMAGE VECTOR 3D‘ Sensor der Firma RedCatch ausgestattet. Der Sensor wird durch SBAS unterstützt und erhält damit eine Lagegenauigkeit von bis zu 2 m.

Insgesamt wurden um die Baustelle kreisförmig zu Fuß insgesamt 115 Aufnahmen aufgenommen und final 88 Bilder (*.jpg, XYZ, RPY, circa 400 MB) für die Berechnung verwendet. Die Berechnungsdauer für die niedrige Qualitätsstufe betrug 21 Minuten, für die hohe Qualitätsstufe 3 Stunden, 16 Minuten.

Trotz dass die Aufnahmebedingungen durch die starke Befahrung erschwert waren und einige der Bilder Aufnahmen von Autos enthalten, ist das Ergebnis einwandfrei und ohne negative Unregelmäßigkeiten (durch den Verkehr).

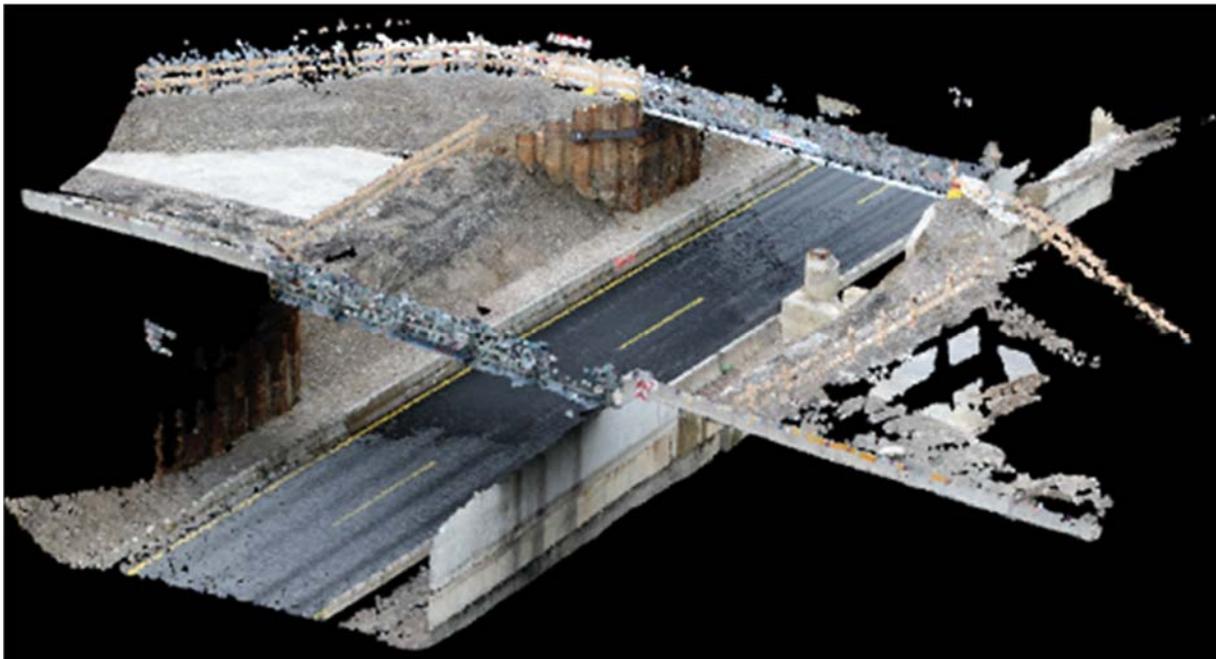


Abb. 7: Punktwolke in hoher Prozessierungsqualität. Aufnahmen wurden kreisförmig gemacht, wobei parallel zur Fahrbahn die gegenüberliegende Seite aufgenommen wurde. Zusätzlich wurden Aufnahmen entlang der zwei Brücken getätigt (erkennbar durch die zwei grauen Mauern in der Höhe).

7 Ausblick

Die Erschließung der möglichen Anwendungsbereiche und Einsatzzwecke von UAS entwickelt sich rasant fort und wird zunehmend auch für Nicht-GIS-Experten attraktiv. Dort gibt es ein weites Feld an Möglichkeiten, Geodaten durch Drohnenaufnahmen gesammelt mit anderen Fachbereichen zu verknüpfen.

Die Prozessierung von Video-Daten als Geo-Videodatenstrom birgt ein großes Potential und wird von der Nutzerseite gefordert. Darüber hinaus ist der Fokus auf die Weiterentwicklung von ‚Nahe-Echtzeit‘-Übertragungssysteme zu richten. Sie haben einen breiten Einsatzbereich durch die sofortige Informationsgewinnung. Beispiele dafür sind Überwachung von Wilderern in Afrika, Monitoring von Pipelines, Identifikation von Brandherden und Ausbreitung großer Flächenbrände, der Einsatz bei Naturgefahren wie etwa Hochwässern, Hangrutschungen oder der Erfassung von Sturmschäden an Bäumen (SCHULZ 2016, MEIER & FRANK 2014). In vielen dieser Fälle ist für die sofortige Abschätzung der Ausmaße oder für die spätere Beweisführung eine lagegenaue Verortung unabdingbar. Interessant wird die Anwendung von UAS-Daten im Bereich multitemporaler Veränderungen, zum Beispiel bei Bauentwicklungen. Dazu können Fortschritte erfasst, berechnet und überprüft werden, wobei diese Daten auch im Zuge von Sachstands- und Beweissicherung von wertvollem Nutzen sind.

Neben optischen Sensoren werden bereits auch Thermalsensoren eingesetzt, welche schon bei der Bergrettung oder bei Feuerwehren Anwendung finden und sich weiterentwickeln werden. Zukünftige Einsatzmöglichkeiten für unzugängliche, gefährliche Gebiete finden sich zum einen im Themenkomplex um das Aufspüren und Verorten von Landminen, basierend auf Radarsensorik zur Erfassung und Identifizierung von Objekten unter der Erdoberfläche. Einen anderen interessanten Einsatz für die Verwendung von UAS ist im Bereich der Überwachung und Kartierung radioaktiv verseuchter Gebiete.

In der Diskussion sind derzeit auch Überlegungen, mit einer Kombination von höchst aufgelösten Satellitenbilddaten und kleinräumig aufgenommenen Drohnen Daten neue Wege, zum Beispiel für die landesweite Abdeckung von Bilddaten zu gehen. Satellitendaten mit einer räumlichen Auflösung von 30-60 cm sind bereits vergleichsweise günstig zu beziehen und besitzen eine temporal sehr gute Abdeckung. Dazu könnten Bilddaten von UAS ein Gegenpart werden. Sie ragen durch ihre räumlich sehr gute Auflösung im Subdezimeterbereich sowie der sehr flexiblen Einsatzfähigkeit und damit zeitlich recht spontanen Datenaufnahme hervor. Dies könnte sich auf die Verwendung des klassischen Luftbildes auswirken. Zwar bieten sie eine gute räumliche Auflösung, doch die geringe temporale Abdeckung und die Inflexibilität bei gleichzeitig recht teuren Produkten mindert deren bisherige Vorteile gegenüber Konkurrenzprodukten stark.

8 Fazit

Geoprodukte, gewonnen aus UAS-Daten, sind im Aufwind – definitiv! Die klaren Vorteile was deren Flexibilität und rasche Informationsgewinnung betrifft, sind hinreichend bekannt. Dieser Artikel stellte vor, wie diese Flexibilität auch für die Weiterverarbeitung, hin zu Geoprodukten, aufrechterhalten werden kann. Die GeoApp.UAS besticht durch ihre simple Handhabung in der automatischen Ableitung von Punktwolke, Orthomosaik und digitalem Geländemodell aus UAS-

Aufnahmen. Durch den generischen Ansatz der Anwendung, können die Orientierungsdaten letzterer flexibel gelesen und verwendet werden, womit man unabhängig vom Hersteller und Typ des UAS ist.

Ein weiterer, klarer Vorteil liegt bei der Prozessierung: die Anwendung läuft auf einem Server, das heißt in der Cloud als Hexagon Smart M.App auf dem Marketplace. Somit entfallen laufende Kosten für die Software, Wartung und Einarbeitung, zudem ist der Nutzer ungebunden was Hardwarevoraussetzungen betrifft und kann jederzeit und überall die Anwendung starten und einsehen. Mit dem zugrundeliegenden Amazon Web Server kann sicher, schnell und parallel prozessiert werden. Es stehen innerhalb des Hexagon Marketplace unterschiedliche Abonnementpakete zur Verfügung.

Die angeführten Beispiele beschreiben die vielfältigen Einsatzmöglichkeiten der GeoApp.UAS: sie reichen vom Laien bis hin zum GIS-Experte mit seinen speziellen Anforderungen. Dies spricht auch Experten anderer Fachbereiche an und resultiert in der Gewinnung neuer Erkenntnisse durch die Verknüpfung verschiedener Disziplinen. Dabei erleichtert die einfache Herangehensweise den Zugang, da keine großen Hürden bewältigt werden müssen, sei es auf Seiten des Anwenderwissens noch was Hard- und Software betrifft.

9 Literaturverzeichnis

- AMATO, A., 2015: dronelife.com: The Latest Drone Numbers From CB Insights' Future of Frontier Tech Report, 28.8.2015, URL: <http://dronelife.com/2015/08/28/the-latest-drone-numbers-from-cb-insights-future-of-frontier-tech-report/>, letzter Zugriff, 18.3.2016.
- BEER, K., 2016: US-Luftfahrtbehörde: Zahl der Drohnen wird stark steigen. heise online, 25.3.2016, URL: <http://heise.de/-3151772>, letzter Zugriff, 29.3.2016.
- MEIER, G. & FRANK, S., 2014: Dokumentation und Überwachung einer Rutschung mittels UAV (Unmanned Aerial Vehicle). Geomatik Schweiz (10), 449-452.
- SCHULZ, C., 2016: Drohnen – Bedrohung oder Bereicherung? Editorial, der adler **4**, 3-10.

Bild-basierte 3D-Rekonstruktion texturarmer Kleinwiederkäuerknochen – erste Leistungsuntersuchungen

DARIA HOLLENSTEIN¹ & STEPHAN NEBIKER¹

Zusammenfassung: Im Rahmen eines interdisziplinären Forschungsprojekts sollen feldtaugliche Verfahren zur automatischen Klassifizierung von archäologischen Knochenfunden entwickelt werden. In einem ersten Schritt werden kostengünstige Lösungsansätze zur bildbasierten Erfassung und 3D-Rekonstruktion ausgewählter Skelettelemente untersucht. Im Zentrum dieses Beitrags steht die 3D-Rekonstruktion von Skelettelementen kleiner Wiederkäuer. Diese Objekte stellen aufgrund ihrer zum Teil geringen Größe und mangelnder Textur besondere Anforderungen. Für die Untersuchungen wurden Studienobjekte mit einer handelsüblichen DSLR-Kamera erfasst und mittels Structure-from-Motion-Ansätzen bildbasiert rekonstruiert. Dabei wurden die Resultate unterschiedlicher Aufnahmekonfigurationen mit Referenzmodellen aus Daten eines Streifenlichtscanners verglichen. Die bildbasierten 3D-Modelle erfüllen mit Abweichungen, die sich weitestgehend im Submillimeter-Bereich bewegen, grundsätzlich die Genauigkeitsanforderungen, um für morphologische Analysen und als Referenz für die Artenbestimmung herangezogen zu werden. Die verwendete Basis-Aufnahmekonfiguration kann durch kontrastverstärkende Maßnahmen und das Aufnehmen zusätzlicher Bilder pro Drehtellerumdrehung im Hinblick auf die resultierende Objektgenauigkeit optimiert werden.

1 Einleitung

Die morphologische Artenbestimmung archäologischer Tierknochenfunden erfordert herkömmlicherweise einen ausreichenden Bestand an modernen Referenzskeletten aus einer anatomischen Vergleichssammlung. Nicht immer stehen in der Praxis entsprechende Vergleichsskelette zur Verfügung: Zum Teil decken die zugänglichen Sammlungen nicht das benötigte Artenspektrum oder eine ausreichende Varietät innerhalb der gefragten Spezies ab, oder die Artenbestimmung muss im Feld erfolgen, ohne dass eine Vergleichssammlung konsultiert werden kann.

Mit dem Aufkommen hochauflösender Verfahren für die 3D-Digitalisierung und der Entwicklung entsprechender Verarbeitungs- und Analysesoftware, besteht die Möglichkeit 3D-Daten von Skelettelementen oder ganzen Skeletten aus Vergleichssammlungen ortsunabhängig verfügbar zu machen. Das Potential neuer Digitalisierungstechnologien für die Bereitstellung von 3D-Modellen von Vergleichssammlungsstücken für die Archäozoologie wurde bisher erst in Einzelprojekten genutzt. Vor allem mit Hilfe von vergleichsweise teuren Scannern wurden wichtige Einzelobjekte (MARSHALL UNIVERSITY 2005) oder Sammlungen vollständiger Skelette (DIGIMORPH STAFF 2002; INSTITUTE OF BIOTECHNOLOGY, UNIVERSITY OF HELSINKI 2003; NIVEN et al. 2009) in 3D digitalisiert und virtuell bereitgestellt. So sind zum Teil umfassende Kataloge vollständiger Einzelskelette für bestimmte Artenspektren in hoher Qualität entstanden (BETTS et al. 2011; COLLEGE OF THE HOLY CROSS et al. 2011). Für die archäozoologische Bestimmungspraxis werden aber ganze Reihen einzelner Skelettelemente einer bestimmten

¹ Fachhochschule Nordwestschweiz FHNW, Institut Vermessung und Geoinformation IVGI, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: [daria.hollenstein, stephan.nebiker]@fhnw.ch

Spezies benötigt. Von einer flächendeckenden Digitalisierung vorhandener Vergleichsbestände, was letztlich eine ortsunabhängige Artenbestimmung erlauben würde, kann noch nicht die Rede sein. Mit ein Grund dafür ist möglicherweise, dass die bisher angewandten Digitalisierungsverfahren meist auf teuren Sensortechnologien beruhen, deren Anwendung Expertenwissen voraussetzt.

Im Rahmen eines interdisziplinären Forschungsprojekts sollen feldtaugliche Verfahren zur automatischen Klassifizierung von archäologischen Knochenfunden entwickelt werden. In einem ersten Schritt werden kostengünstige, bildbasierte Lösungsansätze für die 3D-Rekonstruktion ausgewählter Skelettelemente untersucht. Da für das Vorhaben eine größere Reihe moderner Tierknochen digitalisiert werden soll, ist ein Verfahren nötig, das bei hoher Aufnahmeeffizienz 3D-Modelle in zweckdienlicher Auflösung und Genauigkeit liefert.

Bildbasierte Verfahren, wie die 3D-Rekonstruktion mit Structure-from-Motion (SfM)-Ansätzen, erfüllen diese Anforderungen grundsätzlich (KERSTEN & LINDSTAEDT 2012; KOUTSOUDIS et al. 2013; OMELANOWSKY et al. 2013; LERMA & MUIR 2013; REMONDINO et al. 2014; KOUTSOUDIS et al. 2015) und sind zusammen mit einem Drehteller gut geeignet für die Digitalisierung kleiner, leichter Objekte. Allerdings stellt die 3D-Rekonstruktion moderner Skelettelemente aus anatomischen Vergleichssammlungen besondere Herausforderungen (vgl. KOUTSOUDIS et al. 2015): die Oberfläche moderner Vergleichsskelette ist meist hell und texturarm. Bei SfM-gestützten Verfahren basiert die Rekonstruktion der Kameraorientierung auf der Erkennung homologer Bildpunkte für die Verknüpfung der Bilder. Dazu sind Erkennungsmerkmale auf der Objekt Oberfläche unabdinglich. Die Güte und Zahl der gefundenen Bildverknüpfungspunkte hat einen positiven Einfluss auf den Erfolg der Bildorientierung und auf die Genauigkeit der Objektrekonstruktion (KOUTSOUDIS et al. 2015). Für die SfM-basierte Objektrekonstruktion wäre also eine texturreiche Oberfläche mit starken Kontrasten ideal. Die Anwendung von Verfahren, welche direkt auf die Oberfläche einwirken um dem Problem mangelnder Texturierung zu begegnen, ist jedoch bei der Digitalisierung von Objekten aus anatomischen Vergleichssammlungen aus kuratorischen Gründen nicht zulässig.

Im Folgenden werden die Resultate einer Vorstudie präsentiert, in welcher die Tauglichkeit des SfM-Ansatzes für die bildbasierte 3D-Rekonstruktion moderner Kleinwiederkäuerknochen (Größe z.T. < 40 mm) untersucht wurde. Damit die resultierenden 3D-Modelle den Zwecken morphologischer Analysen und als Referenz für die Artenbestimmung genügen, müssen sie sowohl hinsichtlich der Auflösung wie auch hinsichtlich der Genauigkeit Qualitätsansprüche im Submillimeterbereich erfüllen. Um diese Anforderungen zu erfüllen, wurde daher abhängig von der Objektgröße mit unterschiedlichen Abbildungsmaßstäben gearbeitet. Neben den grundsätzlichen Tauglichkeitstests wurden verschiedene Parameter der Aufnahmekonfiguration und ihr Einfluss auf die Aufnahmeeffizienz und die resultierende Objektgenauigkeit untersucht: Dazu wurde eine Basis-Aufnahmekonfiguration hinsichtlich gewählter Hintergrundgestaltung und Lichtquelle, Anzahl aufgenommener Bilder und Abbildungsmaßstab variiert. Die resultierenden 3D-Rekonstruktionen wurden mit Referenzmodellen aus Streifenlichtscanning-Daten verglichen. Eingangs werden die Untersuchungsgegenstände, die Sensoraufstellung und die Bildverarbeitung für die 3D-Rekonstruktion dargelegt. Es folgt eine Beschreibung der Basis-Aufnahmekonfiguration und ihrer Variationen für die experimentelle Datenaufnahme. Schließlich werden die

Resultate mit dem Referenzmodell aus Streifenlichtscanning-Daten verglichen und die gewonnenen Erkenntnisse für eine optimierte Aufnahmekonfiguration diskutiert.

2 Ausgangslage

2.1 Verwendete Studienobjekte

Für die Studie wurden zwei Skelettelemente unterschiedlicher Größenordnung herangezogen (Abb. 1). Das größere Objekt ist ein Humerus (150.8 mm Länge) der Ziege, das kleinere ist ein Zehenknochen – Phalang I – (40.0 mm Länge) vom Schaf. Beide Objekte verfügen über eine helle Oberfläche mit geringfügiger Texturierung. Vom morphologische Standpunkt her sind nicht nur die Gesamtgeometrie des Objektes sowie die Form von Gelenkrollen und Fortsätzen interessant, sondern auch Merkmale wie das Vorhandensein und die Position kleiner Foramina (Öffnungen für Blutgefäße), deren Grösse am Ziegen-Humerus ca. 2 mm beträgt, oder der Verlauf feiner Knochenkämme.

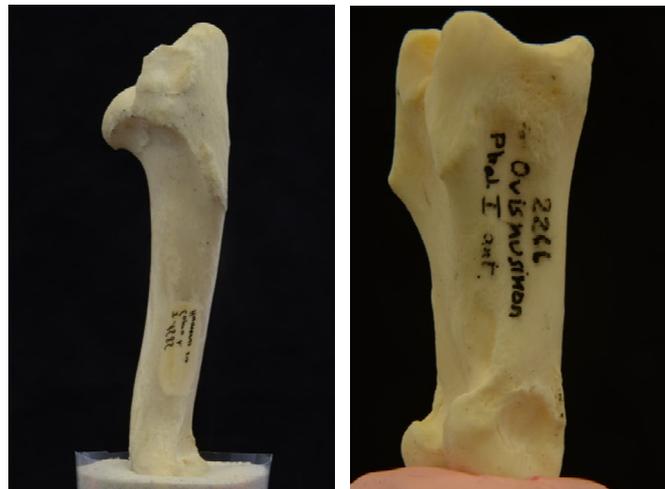


Abb. 1: Studienobjekte – Humerus (max. Länge 151.0 mm) (links), Phalang I (max. Länge 40.0 mm) (rechts)

2.2 Grundlegende Merkmale der Sensoraufstellung

Für die Aufnahme der Bilder wurde eine handelsübliche DSLR-Kamera, vom Typ Nikon D7000 (16.2 Megapixel, 23.6 mm × 15.6 mm Sensorgröße) mit einem Nikkor AF-S 18-105 mm Zoom-Objektiv in Kombination mit dem Seitz Roundshot VR Drive als automatischer Drehteller verwendet (Abb. 2 links). Das Skelettelement wurde mittels Plastilin oder kleinem Sandkasten vertikal auf dem Drehteller positioniert. Die Drehtelleroberfläche wurde mit einer gemusterten Auflage bedeckt. Das Muster dient einerseits als Maßstab für die Skalierung der 3D-Modelle und andererseits liefert es zusätzliche Merkmale für die Verknüpfung von Bildpaaren. Für die Vollerfassung des Objekts sind mindestens zwei Einstellungen, respektive Bildreihen, nötig: Einmal wird das Skelettelement mit dem proximalen Gelenkende nach oben auf dem Drehteller positioniert und einmal mit dem distalen Gelenkende (Abb. 2 Mitte & rechts).

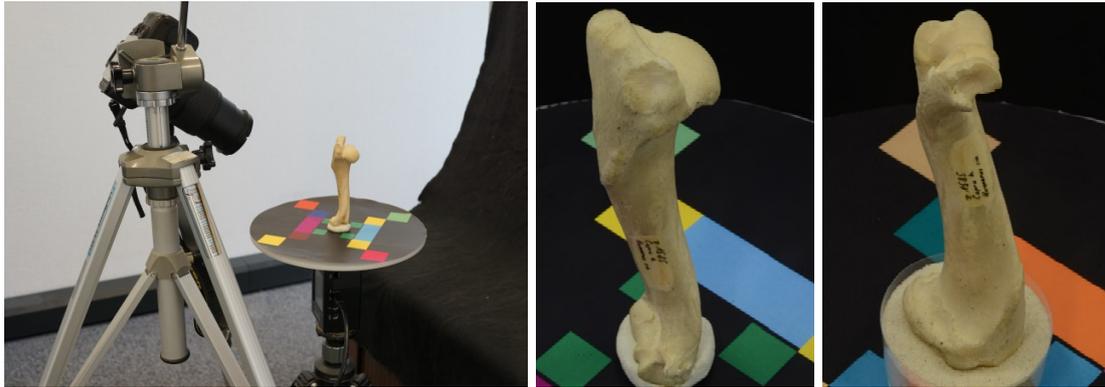


Abb. 2: Sensoraufstellung (links), Skelettelementpositionierung mit dem proximalen Gelenkende nach oben (Mitte) und mit dem distalen Gelenkende nach oben (rechts).

In Anlehnung an die Resultate anderer Studien (KOUTSOUDIS et al. 2013; REMONDINO et al. 2012) wurde auf eine vorgängige Kalibrierung der Kamera verzichtet. Die Kameraparameter wurden über die Selbstkalibrierung mit der verwendeten Software bestimmt. Dies auch weil die Kalibrierung der Kamerakonstanten bei der Digitalisierung einer grösseren Reihe von Tierknochen nicht praktikabel wäre. Die Kamerakonstante müsste für jedes Objekt der Messreihe separat kalibriert werden, da aufgrund der geringen Tiefenschärfe in diesem Distanz- und Abbildungsmaßstabbereich für jedes Objekt neu fokussiert werden muss, damit das gesamte Objekt scharf abgebildet wird.

Auf eine rigorose Kontrolle der Lichtverhältnisse wurde beim Experiment ebenfalls verzichtet, da in der Praxis eine Kontrolle nicht möglich sein wird und mit einer Mischung aus Natur- und Kunstlicht – konkret Deckenbeleuchtung – gearbeitet werden muss. Das Resultat sind leichte Belichtungsunterschiede zwischen den einzelnen Bildern. Wo im folgenden Kapitel 3D-Modelle direkt miteinander verglichen werden, wurden die Bilder für die Rekonstruktionen in direkter Folge und mit derselben Belichtungszeit aufgenommen, sofern nicht explizit anders vermerkt. Die Aufnahmeverhältnisse sind somit vergleichbar.

2.3 Datenprozessierung

Die Bilddaten wurden mit der SfM-Software Agisoft Photoscan Version 1.6.0 zu 3D-Oberflächenmodellen verarbeitet. Bei allen Rekonstruktionen wurde demselben Verarbeitungsprozess gefolgt (Ausnahme Modell D, vgl. 3.2.4) und die Einstellung der Programm-Parameter wurde für alle Modelle beibehalten:

Zuerst wurden jeweils alle Bilder zusammen orientiert, welche für die Rekonstruktion eines Knochenmodells herangezogen wurden. Das so berechnete Modell wurde anhand des Maßstabs auf der Drehtellerauflage skaliert. Anschließend wurden dichte Punktwolken für jede einzelne Bildreihe generiert und manuell bereinigt. Dann wurden die dichten Punktwolken der einzelnen Bildreihen miteinander vereinigt. Aus der vereinigten Punktwolke wurde das 3D-Oberflächenmodell abgeleitet. In Geomagic Qualify Version 12 wurden Ausreißer bereinigt.

2.4 3D-Modellevaluation mittels Referenzmodellen

Für die Evaluierung der resultierenden 3D-Netze wurden Referenzmodelle des Humerus und des Phalangen mit einem hochauflösenden GOM ATOS I Streifenlichtscanner (spez. Genauigkeit <

50 μm) angefertigt. Die Referenzmodelle werden in der vorliegenden Studie als fehlerfreie Werte angenommen. Die Registrierung der bildbasierten 3D-Rekonstruktionen mit den Referenzmodellen erfolgte in 3D Reshaper 2015 MR1 und CloudCompare Version 2.6.2 in zwei Schritten: Einer groben Ausrichtung anhand dreier manuell gesetzter Punkte folgte eine bestgeeignete Registrierung mittels Methode der kleinsten Quadrate. Danach erfolgte in derselben Software der Oberflächenvergleich zwischen bildbasierter Rekonstruktion und Referenzmodell, das heißt die Berechnung von Distanzkarten.

3 Datenaufnahme

3.1 Basis-Aufnahmekonfiguration

In der Basis-Aufnahmekonfiguration (Abb. 2 links, Tabelle 1: Modelle A1, B1, C1) betrug der Winkel der Kamerakonstanten auf die Drehtelleroberfläche ca. -45° . Für die Aufnahmen des Humerus betrug der geringste Abstand zwischen Kameraobjektiv und Skelettelement ca. 300 mm. Bei diesem Abstand und einer Brennweite von 105 mm wird der Humerus längs des Sensors in etwa bildfüllend aufgenommen. In dieser Konfiguration wurden jeweils zwei Vollumdrehungen des Drehtellers mit je 36 Bildern im Abstand von 10° aufgenommen: eine Umdrehung für die Objektpositionierung mit dem proximalen Gelenkende nach oben und eine mit dem distalen Gelenkende nach oben. Position und Fokus der Kamera wurden zwischen diesen beiden Einstellungen unverändert gelassen. Aus dieser Aufstellung ergibt sich eine Vollerfassung des Objekts. Als Hintergrundfarbe wurde schwarz gewählt.

Ausgehend von der Basis-Aufnahmekonfiguration wurden in den Experimenten A-D jeweils Parameter variiert (Tabelle 1), um zu untersuchen, ob damit einerseits eine Verbesserung der Modellgenauigkeit oder eine höhere Aufnahmeeffizienz bei vergleichbarer Modellqualität erzielt werden kann. Wo im Kapitel 4 Resultate unterschiedlicher Aufnahmekonfigurationen miteinander verglichen werden, sind jeweils ein gleichbleibender Abstand zum Objekt und dieselbe Brennweite ohne Neufokussierung zwischen diesen Konfigurationen gegeben.

3.2 Variationen der Basis-Aufnahmekonfiguration

3.2.1 Hintergrundfarbe

In Experiment A wurde die Verwendung unterschiedlicher Hintergrundfarben getestet. Ziel der Untersuchung war das Finden einer Balance zwischen Kontrastverstärkung und Minimierung der Belichtungsdauer. Für das Gelingen der Bildorientierung mittels SfM-Verfahren ist der Hintergrund idealerweise kontrastlos und unterscheidet sich im Grauwert deutlich vom Objekt und den Farben des Maßstabmusters. Gleichzeitig sollte die Grauwertvarianz innerhalb des Objektbereichs möglichst groß sein. Diese Faktoren begünstigen den Erfolg des Matching-Algorithmus beim Identifizieren homologer Punkten auf dem Objekt in den verschiedenen Bildern. Ein dunkler Hintergrund (schwarz oder dunkelgrün) entspricht im vorliegenden Fall am ehesten diesen Anforderungen. Andererseits gilt, je dunkler der Hintergrund, desto länger die Belichtungszeit. Neben der Basis-Aufnahmekonfiguration mit schwarzem Hintergrund (Modell A1) wurde die Verwendung von Rot (Modell A2) und Hellgrün (Modell A3) getestet. Die Belichtungszeit lässt sich bei Verwendung von Rot oder Hellgrün um eine, respektive zwei

Belichtungsstufen reduzieren im Vergleich zu Schwarz. Je nach Lichtverhältnissen bedeutet das eine Einsparung von über einer halben Minute pro Vollumdrehung auf dem Drehteller.

Tab. 1: Experimente A1-D mit den entsprechenden Parametern (N/K: Natur- und Kunstlicht; N: Naturlicht)

Modell	Objekt	Einstellung	Brennweite [mm]	Distanz [mm]	Base-to-height ratio	Konvergenzwinkel	Pixelgröße im Objektraum [mm]	Bilder/Reihe	Hintergrundfarbe	Beleuchtung	Belichtungszeit
A1	Humerus	-45°	105	300	0.175	10.0°	0.014	36	schwarz	N/K	1.3
A2	Humerus	-45°	105	300	0.175	10.0°	0.014	36	rot	N/K	1.6
A3	Humerus	-45°	105	300	0.175	10.0°	0.014	36	hellgrün	N/K	2
B1	Humerus	-45°	105	320	0.175	10.0°	0.015	36	schwarz	N/K	1.3"
B2	Humerus	-45°	105	320	0.175	10.0°	0.015	36	schwarz	N	3"
C1	Humerus	-45°	105	320	0.175	10.0°	0.015	36	schwarz	N/K	1.3"
C2	Humerus	-45°	105	320	0.116	6.7°	0.015	54	schwarz	N/K	1.3"
C3	Humerus	-45°	105	320	0.088	5.0°	0.015	72	schwarz	N/K	1.3"
D	Phalang I	-45°/0°	173	100/130	0.116	6.7	0.003/0.004	54	schwarz	N/K	1"

3.2.2 Beleuchtung

In Experiment B wurde die Lichtquellenkombination aus Natur- und Kunstlicht (Modell B1) in der Basis-Aufnahmekonfiguration mit Naturlicht als einzige Belichtungsquelle (Modell B2) verglichen. Der Vorteil des reinen Naturlichts sind stärkere Kontraste auf der Objektoberfläche. Mit eingeschaltetem Deckenlicht hingegen verringern sich die Belichtungsunterschiede zwischen den einzelnen Bildern und die Belichtungszeit verkürzt sich.

3.2.3 Anzahl Bilder

Für Experiment C wurde die Anzahl Bilder pro Modellrekonstruktion variiert. Gegenüber der Grundeinstellung von 36 Bildern (Modell C1) wurde die Anzahl auf 54 (Modell C2) und 72 Bilder (Modell C3) erhöht. Eine größere Anzahl Bilder resultiert in breiteren Überlappungsbereichen zwischen den Bildern. Das Basis-Höhen-Verhältnis ändert sich von 0.175 bei 36 Bildern zu 0.116 bei 54 respektive 0.088 bei 72 Bildern (Tabelle 1). Bei kleinerem Basis-Höhen-Verhältnis können potentiell mehr oder bessere Verknüpfungspunkte für die Orientierung der Bilder gefunden werden. Andererseits verschlechtert sich dadurch die aus der Kameraaufstellung resultierende Strahlenschnittgeometrie und der Aufwand für die Prozessierung steigt erheblich.

3.2.4 Geringe Objektgröße (Phalang I)

Kleinere Skelettelemente, wie der erste Phalang (40 mm Länge) (Modell D), können mit der beschriebenen Basis-Aufnahmekonfiguration nicht befriedigend rekonstruiert werden, da sie an der Naheinstellgrenze des Objektivs in keiner Richtung des Sensors auch nur annähernd

bildfüllend sind. Um den Abbildungsmaßstab mit einfachen Mitteln zu vergrößern, wurden dem Nikkor AF-S 18-105 mm Zoom-Objektiv bei einer Brennweite von 105 mm Makrozwisehringe von einer Gesamtlänge von 68 mm vorgelagert. Damit kann der Phalang in der Ansicht in Längsrichtung des Sensors bildfüllend aufgenommen werden. Die kürzeste Distanz zwischen Objektiv und Skelettelement in dieser Aufstellung beträgt 100 mm in der -45° -Winkel-Draufsicht und 130 mm in der Horizontalansicht. Neben den beiden Einstellungen der Basis-Aufnahmekonfiguration, mit denen aufgrund der begrenzten Tiefenschärfe nicht mehr das ganze Objekt, sondern nur noch das proximale, respektive das distale Gelenkende aufgenommen werden können, wurden zusätzlich Bilder aus einer horizontalen Einstellung auf das Objekt für die 3D-Rekonstruktion verwendet. In allen drei Einstellungen wurden 54 Bilder pro Vollumdrehung des Drehtellers aufgenommen. Für Modell D konnte nachdem die Bilder sämtlicher Reihen miteinander orientiert waren direkt die dichte Punktwolke über alle Bilder gerechnet werden, da die Fixierung des Objekts (Plastilin) nicht in den Bildausschnitten der -45° -Einstellungen enthalten ist.

4 Resultate

Die Resultate aus den Experimenten A-D werden in Tabelle 2 präsentiert. Einige Feststellungen könne über die einzelnen Experimente hinaus für alle rekonstruierten 3D-Modelle gemacht werden: Grundsätzlich erfüllen alle 3D-Modelle weitestgehend die formulierten Genauigkeitsanforderungen im Submillimeter-Bereich (Tabelle 2): Der mittlere Bildmessfehler (mean reprojection error) beträgt je nach Modell und Bild zwischen 0.9 und 1.4 Pixel beim Humerus und zwischen 1.7 und 2.6 Pixeln beim Phalangen. Die grössten Abweichungen zum Referenzobjekt betragen je nach Modell zwischen 1.0 und 1.6 mm. Die Standardabweichung (σ) zur mittleren Abweichung vom Referenzobjekt beträgt maximal 0.16 mm. Bildbasierte Verfahren sind also durchaus geeignet für die 3D-Rekonstruktion texturarmer Kleinwiederkäuferknochen. Auch ohne anspruchsvolle Requisiten (vgl. KOUTSOUDIS et al. 2015) ist für geringfügig texturierte Tierknochen mit einer Aufstellung, wie sie als Basis-Aufnahmekonfiguration in Kapitel 3.1 beschrieben wurde, eine hohe Genauigkeit erzielbar. Modell A1 z.B. weicht über 96.8% der Punkte bzw. der Oberfläche um maximal ± 0.3 mm vom Referenzmodell ab (Abb. 3, links). Abweichungen von mehr als ± 0.1 mm zum Referenzobjekt treten bei allen Modellen in erster Linie im Bereich der Gelenkenden auf (vgl. Abb. 3, links) und zwar dort, wo das Modell auf den Bilddaten lediglich einer Einstellung basiert, weil der entsprechende Bereich nur durch eine Bildreihe abgedeckt ist. Die grössten Abweichungen gegenüber dem Referenzmodell treten in konkaven Bereichen der Skelettelementoberfläche auf. Werden die Modelle im Hinblick auf die formulierten morphologischen Anforderungen einer visuellen Prüfung unterzogen ist festzustellen, dass gerade die genannten Abweichungen in konkaven Bereichen der Knochenoberfläche mitunter problematische Auswirkungen haben: Die Foramina z.B. sind auch bei Modellen mit geringen Abweichungen gegenüber dem Referenzmodell nur undeutlich zu erkennen (Abb. 3, rechts).

Auch wenn Vergleiche zwischen Modellen aus unterschiedlichen Experimenten aufgrund der schlecht kontrollierbaren Rahmenbedingungen schwierig sind, zeichnet sich ein Zusammenhang

zwischen der Pixelgrösse im Objektraum (Tabelle 1) und der Modellgenauigkeit (Tabelle 2) ab (vgl. A1 und D mit B1 und C1).

Ferner ist die Genauigkeit der Orientierung der verschiedenen Bildreihen zueinander entscheidend. Aber mit der verwendeten Software ist bzgl. der vorgestellten Experimente nur bedingt eruiert, was eine exakte Orientierung der Bildreihen zueinander ausmacht.

Tab. 2: Kenngrößen der Oberflächenmodelle aus Experimenten A-D

Modell	Mittlerer Bildmessfehler der Bildorientierung [mm]	Mittlere Abweichung zum Referenzmodell [mm]	σ zur mittleren Abweichung zum Referenzmodell [mm]	größte Abweichungen zum Referenzmodell [mm]	Modelloberfläche mit Abweichung zum Referenzmodell ≤ 0.1 mm [%]	Modelloberfläche mit Abweichung zum Referenzmodell ≤ 0.3 mm [%]	Modelloberfläche mit Abweichung zum Referenzmodell ≤ 1 mm [%]
A1	1.1	0.018	0.098	1.10/-0.93	74.3	96.8	100.0
A2	1.0-1.3	-0.067	0.117	1.08/-0.97	58.8	94.4	100.0
A3	1.0-1.2	0.085	0.160	1.26/-1.58	53.4	84.6	99.9
B1	0.9-1.1	0.054	0.134	1.48/-1.15	61.9	91.7	99.9
B2	0.9-1.2	-0.047	0.080	1.28/-0.48	71.8	98.7	100.0
C1	0.9-1.1	0.073	0.157	1.63/-1.29	57.5	85.1	99.8
C2	1.1-1.3	0.027	0.102	1.50/-1.01	74.1	96.4	100.0
C3	1.1-1.4	-0.049	0.103	1.27/0.53	68.1	96.4	99.9
D	1.7-2.5	0.082	0.070	0.62/-0.22	68.3	99.4	100.0

4.1 Hintergrundfarben (A1-A3)

Experiment A (Tabelle 2) zeigt, dass mit deutlichen Qualitätseinbußen zu rechnen ist, wenn zugunsten verkürzter Belichtungszeiten, eine hellere Hintergrundfarbe, wie Rot oder Grün, gewählt wird. Die auf den Gesamtrekonstruktionsprozess gesehen relativ geringe Zeiteinsparung scheint aufgrund dieses Experimentes nicht lohnenswert.

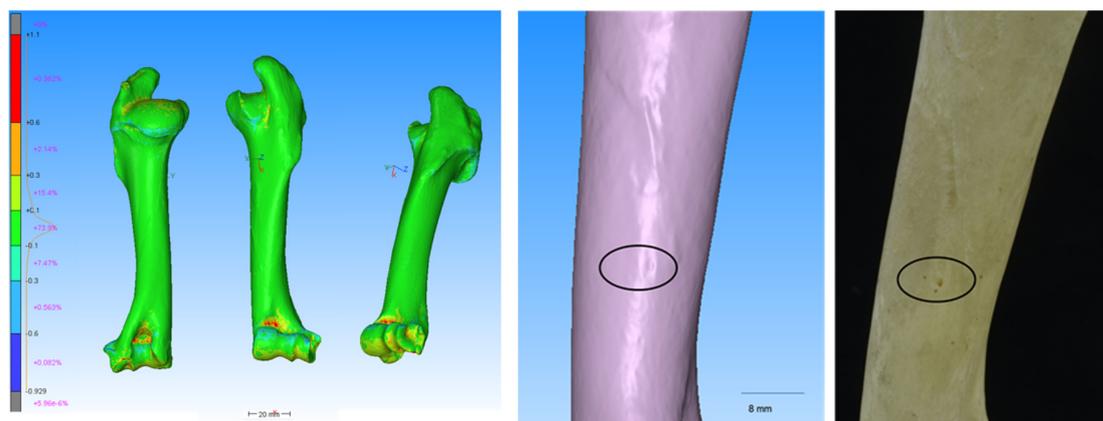


Abb. 3: Model A1 gefärbt nach Distanzen zum Referenzmodell in mm (links), Sichtbarkeit des Foramen am Humerusschaft auf Model A1 und im Bild des Untersuchungsgegenstandes (rechts)

4.2 Beleuchtung

Die ausschließliche Verwendung von Naturlicht als Beleuchtungsquelle (Modell B2), also die Kontrastverstärkung auf der Objektoberfläche mittels Ausschalten des Deckenlichts, ergibt im direkten Vergleich mit der Basis-Aufnahmekonfiguration (Modell B1) eine Verbesserung der Modellgenauigkeit. Dieser Befund ist abhängig von den natürlichen Lichtverhältnissen im Moment der Aufnahme. Zum Zeitpunkt der Aufnahmen für dieses Experiment war durch Ausschalten des Deckenlichts eine deutliche Kontrastverstärkung auf dem Objekt erkennbar (Abb. 5). Dies ist nicht bei allen Witterungsverhältnissen und Tageszeiten der Fall. Für diesen Befund ist insofern eine bedingte Reproduzierbarkeit zu erwarten. Experimente A1 und D1 (Tabelle 2) machen zudem deutlich, dass das Ausschalten des Deckenlichts keine Voraussetzung für eine hohe Modellgenauigkeit ist. Sehr gute Ergebnisse sind auch mit eingeschaltetem Deckenlicht möglich.



Abb. 4: Aufnahme des Studienobjekts bei Kunstlicht- (links) und Naturlicht-Verhältnissen (Mitte); Oberflächenmodell D (rechts)

4.3 Anzahl Bilder

Mit 54 Bildern pro Einstellung (C2) ergab sich in Experiment C (Tabelle 2) eine deutliche Verbesserung gegenüber der Verwendung von 36 Bildern (C1). Wenn statt der 54 Bilder, 72 Bilder verwendet werden, verbessert sich die Genauigkeit nur noch geringfügig. Neben dem fehlenden Vorteil hinsichtlich Genauigkeitszuwachs steigt die Verarbeitungsdauer für die Modellrekonstruktion bei 72 Bildern erheblich: Sie beträgt deutlich mehr als das Doppelte der Rekonstruktion aus 36 Bildern.

4.4 Geringe Objektgröße

Der Phalang (Modell D) konnte unter Verwendung von Makro-Zwischenringen mit einer vergleichsweise hohen Genauigkeit (Tabelle 2) (mittlere Abweichung = 0.08 mm, $\sigma = +/-0.07$ mm) zum Referenzmodell rekonstruiert werden. Die errechnete Punktwolke, respektive das abgeleitete Modell, weisen aber Rauschen auf (Abb. 4, rechts). Die Verwendung von 72 Bildern pro Reihe, sowie die Verwendung einer Drehtellerauflage, die im Zentrum mehr Musterung aufweist, könnten hier allenfalls zusätzlich Verbesserung bringen.

5 Fazit und Ausblick

Mit den Experimenten A1-E konnte gezeigt werden, dass mit kostengünstigen bild-basierten Verfahren 3D-Rekonstruktionen texturarmer Kleinwiederkäuerknochen mit einer Genauigkeit im Submillimeterbereich ohne weiteres realisierbar sind. Der für diese Experimente angewandte Bildverarbeitungsprozess ist jedoch relativ aufwändig, da aufgrund der verwendeten Plastilin- und Sandfixierung des Knochens nicht direkt alle Bildreihen zu einem Modell verrechnet werden können, sondern die dichten Punktwolken separat für jede Reihe gerechnet, bereinigt (Entfernen des Plastilin- respektive Sandbereiches) und anschliessend vereinigt werden müssen. Die Verwendung einer Fixierung des Skelettelementes, welche eine solche Nachbearbeitung überflüssig macht, würde einerseits eine Effizienzsteigerung bedeuten und andererseits würden Ungenauigkeiten bei der Vereinigung der dichten Punktwolken vermieden. Ein anderer Ansatz wäre – analog zum Vorgehen bei Modell D – die Wahl eines Bildausschnittes, der die Objektfixierung nicht enthält.

Bei kommenden Experimenten könnten zudem Objektivverzerrungsparameter durch eine vorgängige Kalibrierung bestimmt und in der Ausgleichung berücksichtigt werden. Die vorgängige Bestimmung der Verzerrungsparameter könnte Verbesserungen der Modellgenauigkeiten bringen. Ferner sollte der Einfluss des Abbildungsmassstabs systematischer untersucht werden. Im Falle des Humerus ist eine Vergrösserung nur möglich, wenn nicht mehr das ganze Objekt im Bild ist. Dies führt allerdings zu erheblichem Mehraufwand bei den Bildaufnahmen (fünf statt zwei Bildreihen) und bei der Modellberechnung. Eine Kosten-Nutzen-Evaluation wäre hier nötig. Schliesslich muss in weiteren Experimenten die Reproduzierbarkeit der Modellgenauigkeit bei gegebener Distanz zwischen Kamera und Objekt für andere Humeri und Phalangen überprüft werden, die in der Oberflächenbeschaffenheit von den Untersuchungsgegenständen abweichen.

6 Danksagung

Wir bedanken uns bei der Abteilung Integrative Prähistorische und Naturwissenschaftlichen Archäologie der Universität Basel, namentlich Prof. Jörg Schibler, für die zur Verfügung gestellten Kleinwiederkäuerknochen. Weiter danken wir dem Institut für Medizinal- und Analysetechnologie der Fachhochschule Nordwestschweiz, das freundlicherweise seinen GOM ATOS Scanner zur Verfügung gestellt und bei der Anfertigung der Referenzmodelle Hand geboten hat.

7 Literaturverzeichnis

BETTS, M.W., MASCHNER, H.D.G., SCHOU, C.D., SCHLADER, R., HOLMES, J., CLEMENT, N. & SMUIN, M., 2011: Virtual zooarchaeology: building a web-based reference collection of northern vertebrates for archaeofaunal research and education. *Journal of Archaeological Science* **38** (4), 755-762.

- COLLEGE OF THE HOLY CROSS, MUSEUM OF COMPARATIVE ZOOLOGY & HARVARD UNIVERSITY, 2011: Aves 3D. A three dimensional database of avian skeletal morphology. <http://aves3d.org/> (aufgerufen im August 2015).
- INSTITUTE OF BIOTECHNOLOGY, UNIVERSITY OF HELSINKI, 2003. MorphoBrowser. Verfügbar unter: <http://morphobrowser.biocenter.helsinki.fi/> (aufgerufen im August 2015).
- DIGIMORPH STAFF, 2002: Digital Morphology - A unique biological visualization library @ DigiMorph.Org. <http://digimorph.org/aboutdigimorph.phtml> (aufgerufen im April 2016).
- LERMA, J.L. & MUIR, C., 2014: Evaluating the 3D documentation of an early Christian upright stone with carvings from Scotland with multiples images. *Journal of Archaeological Science* **46**, 311-318.
- KERSTEN, T.P. & LINDSTAEDT, M., 2012: Generierung von 3D Punktwolken durch kamera-basierte low-cost Systeme – Workflow und praktische Beispiele. Schriftenreihe des DVW, Band 69, Terrestrisches Laserscanning 2012 (TLS2012), Beiträge zum 121. DVW Seminar am 13. und 14. Dezember 2012 in Fulda, Wissner-Verlag, Augsburg, 25-46.
- KOUTSOUDIS, A., VIDMAR, B., IONNAKIS, G. ARNAOUTOGLU, F. PAVLIDIS, G. & CHAMZAS, C., 2015: Using noise function-based patterns to enhance photogrammetric 3D reconstruction performance of featureless surfaces. *Journal of Cultural Heritage* **16** (5), 664-670.
- KOUTSOUDIS, A., VIDMAR, B. & ARNAOUTOGLU, F., 2013: Performance evaluation of a multi-image 3D reconstruction software on a low-feature artefact. *Journal of Archaeological Science* **40**, 4450-4456.
- MARSHALL UNIVERSITY 2005: PaleoView3D – an interactive 3D fossil museum. <http://paleoview3d.marshall.edu/> (aufgerufen im April 2016).
- NIVEN, L., STEELE, T. E., FINKE, H., GERNAT, T. & HUBLIN, J.-J., 2009: Virtual skeletons: using a structured light scanner to create a 3D faunal comparative collection. *Journal of Archaeological Science* **36**, 2018-2023.
- OMELANOWSKY, D., KERSTEN, T., & LINDSTAEDT, M., 2013: Untersuchungen von Low-Cost Systemen zur 3D-Rekonstruktion kleiner Objekte. *Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage 2010*, Luhmann, T. & Müller, C. (Hrsg.), VDE Verlag GmbH, Berlin/Offenbach, 217-228.
- REMONDINO, F., DEL PIZZO, S., KERSTEN, T.P. & TROISI, S., 2012: Low-cost and open-source solutions for automated image orientation a critical overview. *Lecture Notes in Computer Science* **7616**, 40-54.
- REMONDINO, F., SEPRA, M.G., NOCERINO, E., MENNA, F. & NEX, F., 2014: State of the Art in High Density Image Matching. *The Photogrammetric Record* **29** (146), 144-166.

Kombination unterschiedlicher Photogrammetrischer Ansätze zur Rekonstruktion zerstörter Kulturdenkmäler

WISSAM WAHBEH¹ & STEPHAN NEBIKER¹

Zusammenfassung: In diesem Beitrag dokumentieren wir Untersuchungen und Resultate zur bildbasierten 3D-Rekonstruktion des Baaltempels, eines der wichtigsten Syrischen Kulturdenkmäler. Der Fokus der Untersuchungen liegt auf der kombinierten Nutzung frei verfügbarer touristischer Bildaufnahmen und professioneller Panoramaaufnahmen. Der Rekonstruktionsprozess umfasst die automatische Bildorientierung unter Einbezug geometrischer Bedingungen für die Panoramaaufnahmen und anschliessendes Multi-View Dense Image Matching. Zudem zeigen wir den Nutzen von Panoramabilddaten in der Co-Registrierung und Georeferenzierung schwach verknüpfter Bildblöcke und Punktwolken auf. Als Resultat des Rekonstruktionsprozesses mit den kombinierten Bilddatenquellen resultieren weitgehend vollständige und detaillierte 3D-Modelle mit einer guten relativen und absoluten Genauigkeit.

1 Einleitung

In der Syrischen Stadt Palmyra, ca. 250 km nordöstlich von Damaskus gelegen, befinden sich die monumentalen Überreste einer mächtigen Stadt, die eines der wichtigsten kulturellen Zentren der Antike war. Palmyra befand sich am Schnittpunkt mehrerer Zivilisationen und war eine pulsierende Oasenstadt, als sie Mitte des 1. Jahrhunderts n. Chr. als Teil der Provinz Syria unter Römische Kontrolle geriet. Palmyra gewann laufend an Bedeutung als Stadt an der Handelsroute, welche Persien, Indien und China mit dem Römischen Reich verband. In der Stadt finden sich einzigartige Beispiele monumentaler Grabmäler, in welchen sich Griechisch-Römische Kunst mit einheimischen Elementen und Persischem Einfluss zu einem unverwechselbaren Stil vereinen. Ausserhalb der Stadtmauern finden sich Überreste eines Römischen Aquädukts und einer riesigen Nekropole.

Palmyra ist das Symbol der Syrischen Archäologie und ein UNESCO-Weltkulturdenkmal. Als Folge des Bürgerkriegs in Syrien wurde Palmyra 2013 in die Rote Liste des gefährdeten Welterbes der UNESCO eingetragen. Die Terrororganisation Islamischer Staat (IS) zerstörte in den von ihm besetzten Gebieten im Nordosten Syriens und im Westen Iraks zahlreiche Kulturdenkmäler. In Palmyra zerstörte der IS die bekanntesten Monumente: den Tempel des Bēl, auch bekannt als Baaltempel, den Triumphbogen (Hadrianstor) an der Prachtstrasse der Stadt sowie den Baalschamin-Tempel. Der grosse Baaltempel (Abb. 1) gilt als wichtigster Sakralbau des 1. Jahrhunderts n. Chr. im Osten mit einer einzigartigen Architektur (KHAN ACADEMY, 2016). Der Tempel befindet sich im Zentrum eines quadratischen architektonischen Bezirks mit einer Seitenlänge von ca. 205 m. Dieser Bezirk wurde wiederum von einer Säulenhalle (Portikus) umgeben. Der Tempel selbst hat ein massives Fundament, welches eine gestufte

¹ FHNW Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation IVGI, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: [wissam.wahbeh, stephan.nebiker]@fhnw.ch

Plattform trägt. Auf der obersten Plattform (Stylobat) mass der Tempel 55×30 m und der Hauptraum (Cella) mit der Kultstatue wies eine Höhe von 14 m und horizontale Dimensionen von $39.45 \text{ m} \times 13.86 \text{ m}$ auf.



Abb. 1: Einige der für die Rekonstruktion verwendeten touristische Aufnahmen des Baaltempels

Touristische Fotos von Kulturdenkmälern sind sehr nützlich, indem sie deren geometrischen und spektralen Eigenschaften zu früheren Zeitpunkten festhalten. Mit der systematischen Zerstörung Syrischer und Irakischer Kulturdenkmäler durch den IS steigt der Bedarf nach Lösungen zur Dokumentation sowie zur virtuellen Rekonstruktion dieser Objekte als Basis für Visualisierungen und für allfällige zukünftige Rekonstruktionen. Die virtuellen 3D-Rekonstruktionen sind auch ein wichtiges Statement, dass zumindest das virtuelle Gedächtnis auch durch derartige barbarische Zerstörungen nicht gelöscht werden kann.

Die photogrammetrische Rekonstruktion ist die einzig verfügbare Lösung in Fällen wie diesem, in welchen das Monument zerstört wurde und keine systematischen 3D-Aufnahmen bspw. mittels Laserscanning existieren. Ein ähnliches Problem stellte sich nach der Zerstörung der grossen Buddha-Statuen von Bamiyan, Afghanistan durch die Taliban im Jahre 2001. Im Anschluss daran wurde eine photogrammetrische 3D-Rekonstruktion mit nur drei metrischen Aufnahmen aus dem Jahre 1970 durchgeführt (GRÜN et al. 2003). Heute, nach über einem Jahrzehnt rasanter Entwicklungen in der Photogrammetrie und Computer Vision sind automatisierte 3D-Rekonstruktionen aus grossen Sammlungen an metrischen und sogar nicht-metrischen Aufnahmen, bspw. von Touristen, möglich.

Aktuelle Forschungsbestrebungen zur bildbasierten 3D-Rekonstruktion von urbanen Räumen oder Kulturdenkmälern können in zwei Hauptströmungen unterteilt werden: genaue 3D-Rekonstruktionen aus (semi-) professionellen Bilddaten einerseits und hoch skalierbare 3D-Rekonstruktionen aus webbasierten Foto-Sammlungen andererseits. KOUTSOUDIS et al. (2014), bieten eine gute Übersicht über Algorithmen auf der Basis von Structure-from-Motion (SfM) und Dense Multi-View Reconstruction (DMVR) und den erzielbaren Resultaten bei Verwendung von (semi-) professionellen DSLR-Kameras. KERSTEN & LINDSTAEDT (2012) und SANTAGATI et al. (2013) untersuchen das Genauigkeitspotential bildbasierter 3D-Rekonstruktionen durch Vergleiche mit Referenzdaten aus terrestrischen Laserscans (TLS). Demgegenüber zeigen beispielsweise AGARWAL ET AL. (2011) das Potential einer vollautomatischen 3D-Rekonstruktion urbaner Räume oder grosser Monumente aus ungeordneten Foto-Sammlungen im Web. Die beiden Hauptströmungen werden einerseits ergänzt durch Untersuchungen wie etwa von REMONDINO et al. (2012) zu (kostengünstigen) Bildorientierungslösungen mit einer kritischen Diskussion zu schwachen Netzgeometrien, wie diese bei öffentlichen Bilddaten oft

unvermeidbar sind. Die nachfolgenden Untersuchungen bauen zudem auf Methodiken zur Beurteilung von Dense Image Matching Resultaten auf (CAVEGN et al. 2015).

In diesem Artikel stellen wir zunächst die verfügbaren touristischen und professionellen Bilddaten vor und diskutieren typische Limitierungen touristischer Aufnahmen. In Kapitel 3 diskutieren wir drei unterschiedliche Strategien zur 3D-Extraktion. In Kapitel 4 beschreiben wir einen Ansatz zur Co-Registrierung mehrerer partieller Punktwolken, die typischerweise aus öffentlichen Bilddaten entstehen. Schliesslich zeigen wir die Nutzung sphärischer Photogrammetrie als Mittel zur Co-Registrierung und Georeferenzierung des Gesamtmodells.

2 Datensätze

2.1 Bilddaten

Für die Rekonstruktion des Baaltempels standen zwei Bilddatenquellen zur Verfügung: frei zugängliche touristische Fotos sowie Panoramaaufnahmen, die mit professioneller Ausrüstung erfasst worden waren.

Die touristischen Aufnahmen wurden im Web gesucht, wovon ein Grossteil auf flickr.com gefunden wurde. In der Bildsuche wurden Bildtags in verschiedenen Sprachen verwendet, die einen thematischen oder örtlichen Bezug zum Objekt haben könnten. Als besonders wertvoll erwies sich auch die "Reverse Image Search"-Funktion von Google. Diese liefert Suchergebnisse basierend auf der inhaltlichen und semantischen Ähnlichkeit (z.B. gleiche Website) mit einem Eingabebild und lieferte zusätzliche Fotos ohne thematische Tags. In einer anschliessenden Filterung wurden Bilder ohne EXIF-Header und einer Bilddimensionen unter 1200×1600 Pixel entfernt, womit noch 180 geeignete touristische Fotos verblieben (vgl. Abb. 2 & 3). Diese umfassten 112 Aussenaufnahmen und 68 Innenaufnahmen des Tempels über einen Zeitraum von 2005 bis 2011 und mit einer grossen Bandbreite an Brennweiten.

Die professionellen Panoramaaufnahmen waren von G. Fangi auf einer Exkursion durch Syrien im Jahr 2010 erfasst worden. Von den total 20 Panoramen, bestehend aus insgesamt 551 Einzelaufnahmen, decken 13 Panoramen den Aussenraum sowie 7 Panoramen den Inneraum des Tempels ab (vgl. Abb. 2 & 3). Die Bilder wurden mit einer DSLR-Kamera Canon EOS 450D mit einem 28 mm Objektiv und einem Nodaladpater erfasst, womit ein gemeinsames Projektionszentrum für alle Aufnahmen eines Panoramastandorts realisiert wird. Um die Bildgeometrie der Originalbilder zu erhalten und um Umprojektions- bzw. Stitching-Fehler zu vermeiden, wurden für die nachfolgenden Untersuchungen ausschliesslich die Original-Einzelaufnahmen verwendet. Zudem unterstützen typische SfM-Softwarepakete, wie etwas AgiSoft PhotoScan, ausschliesslich das perspektivische Lockkamera- sowie neuerdings auch Fischaugenmodelle, nicht aber typische Panoramaprojektionen. Aus der Panoramamesskampagne standen 74 der Aussenaufnahmen sowie 162 der Innenaufnahmen für den Rekonstruktionsprozess zur Verfügung.

2.2 Netzgeometrie und Bilderfassungsmuster

Die Qualität der Netzgeometrie ist ein Schlüsselfaktor für genaue und zuverlässige 3D-Rekonstruktionen (ALSADIK et al. 2014). Im Falle touristischer Aufnahmen ist die Netzgeometrie allerdings kaum beeinflussbar. Typische Bilderfassungsmuster von Touristen auf städtischem Massstab wurden beispielsweise von KÁDÁR & GEDE (2013) anhand der Geotags von frei

zugänglichen Bildsammlungen im Internet untersucht. In unseren Untersuchungen interessieren die Erfassungsmuster auf Stufe Sehenswürdigkeit bzw. Kulturdenkmal und deren Auswirkungen auf den Rekonstruktionsprozess und die Resultate. Touristen interessieren sich oft viel stärker für berühmte bzw. herausragende Teile eines Kulturdenkmals als für weniger bekannte oder weniger leicht zugängliche Teile. Eine vertiefere Analyse und Diskussion dieser Aspekte finden sich in WAHBEH & NEBIKER (2016).



Abb. 2: Standorte und Blickrichtungen der Aussenaufnahmen (touristische Fotos: grün; Panoramas: blau)

Abbildung 2 zeigt die Aufnahmestandorte und Blickrichtungen der Aussenaufnahmen des Baaltempels und Abbildung 3 diejenigen der Innenaufnahmen. Ein visueller Vergleich zeigt die Komplementarität der Panoramaaufnahmen (blau) und der touristischen Fotos (grün/grün) sowie die unvollständige Abdeckung durch letztere.



Abb. 3: Standorte und Blickrichtungen der Innenaufnahmen im Grundriss des erzeugten 3D-Modells (touristische Fotos: grün; Panoramas: blau)

3 3D-Extraktion

Wie bereits in früheren Arbeiten aufgezeigt wurde, nimmt die Zahl an quelloffenen und kommerziellen Softwarepaketen für die Bündelorientierung und für die dichte Bildzuordnung stetig zu (SANTAGATI et al. 2013; KOUTSOUDIS et al. 2014). Aufgrund früherer Vergleiche verschiedener Lösungen wurde für die folgenden Untersuchungen Agisoft PhotoScan verwendet. Ein wichtiger Grund war die Möglichkeit, in der Bündelausgleichung geometrische Bedingungen einzuführen. In unserem Fall war insbesondere die Option zur Fixierung eines gemeinsamen Projektionszentrums für alle Aufnahmen eines einzelnen Panoramas sehr wertvoll. Zudem hatten andere Autoren, wie etwa REMONDINO et al. (2012) aufgezeigt, dass PhotoScan im Vergleich mit anderen Softwarepaketen sowohl bezüglich Genauigkeit als auch Zuverlässigkeit gut abschnitt, auch bei ungünstigen Netzgeometrien.

3.1 Bildorientierung und Bildzuordnung

Die verwendeten zwei Bilddatensätze waren beide nicht mit dem Ziel einer genauen und detaillierten 3D-Rekonstruktion erfasst worden. Beide Datensätze haben jedoch ihre spezifischen Stärken, die in der Folge ausgenutzt werden sollen. So bieten einerseits die touristischen Aufnahmen eine sehr dichte und hoch aufgelöste Abdeckung besonders wichtiger Teile der Denkmäler. Nachteile sind eine oft schwache Netzgeometrie, eine unvollständige Abdeckung sowie die Verwendung zahlreicher nicht-metrischer Kameras. Andererseits wurden die Panoramaaufnahmen von bewusst gewählten, gut verteilten Standorten aus mit einer fixen Brennweite aufgenommen. Die grossen Basislinien und die daraus resultierenden grossen Schnittwinkel sind zwar ideal für eine interaktive Panoramaphotogrammetrie, aber ungeeignet für SfM-Ansätze.

Um die Beiträge der einzelnen Datensätze zur automatischen 3D-Rekonstruktion beurteilen zu können, wurden die nachfolgenden drei Szenarien A) – C) untersucht. Zu diesem Zweck wurden 3D-Punktwolken für jedes Szenario berechnet und anschliessend verglichen.

- Szenario A) Ausschliessliche Verwendung der touristischen Aufnahmen
- Szenario B) Touristische Aufnahmen in Kombination mit Panorama-Einzelbildern **ohne** geometrische Bedingungen
- Szenario C) Touristische Aufnahmen in Kombination mit Panorama-Einzelbildern **mit** geometrischen Bedingungen

3.2 Generierte Punktwolken

Die Aussenhülle des Tempels konnte für alle drei Szenarien in einem einzigen Berechnungsdurchlauf rekonstruiert werden (Abb. 4, blau). Demgegenüber konnten die Innenaufnahmen nicht als gesamtes und in einem Durchlauf orientiert und berechnet werden, da kaum touristische Aufnahmen existieren, welche die beiden sich gegenüberliegenden spektakulären nördlichen und südlichen Kammern verknüpfen würden.

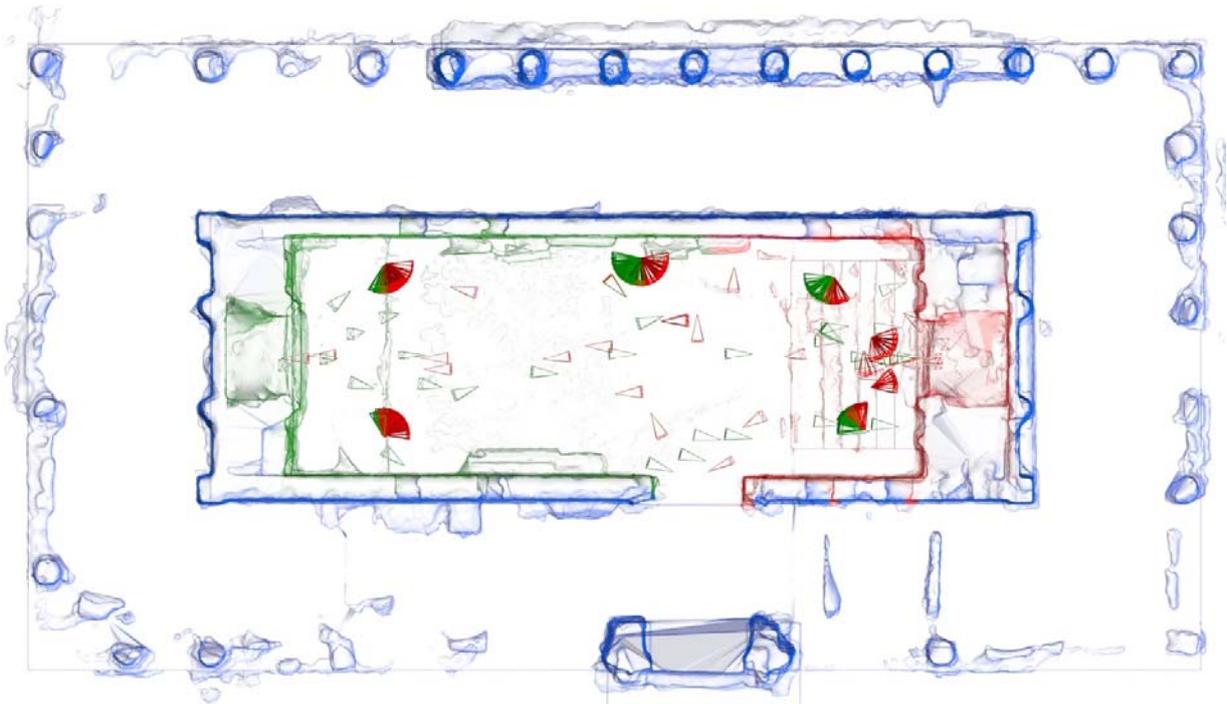


Abb. 4: Grundriss des Baaltempels eingefärbt nach den drei rekonstruierten Teilbereichen bzw. den drei Punktwolken: Innenraum Nordseite (grün); Innenraum Südseite (rot) und Aussenraum (blau).

Als Folge davon wurden die Innenaufnahmen in zwei separate, sich überlappende Gruppen unterteilt und pro Gruppe eine Punktwolke abgeleitet. Die erste Gruppe enthält die Aufnahmen der nördlichen Kammer (Abb. 4, links/grün), die zweite diejenigen der südlichen Kammer (Abb. 4, rechts/rot). In den Szenarien B und C wurden die Panorama-Aufnahmen mit einbezogen. Der Einbezug gemeinsamer Panoramaaufnahmen in diesen Szenarien kann in der späteren Co-Registrierung ausgenutzt werden. Aus dem umfassendsten Szenario C mit den meisten Bilddaten ergaben sich folgende drei Punktwolken: eine Punktwolke für den nördlichen Innenraum mit ca. 8.5 Mio. Punkten und für den südlichen mit 9.5 Mio. Punkten sowie eine Punktwolke der Aussenhülle mit ca. 6.5 Mio Punkten.

3.3 Analyse der Ergebnisse

Um den Einfluss der einzelnen Bilddatentypen und der geometrischen Bedingungen zu analysieren, wurden die bildbasierten Punktwolken miteinander verglichen. Da wegen der Zerstörung des Tempels kein unabhängiger Referenzdatensatz zur Verfügung stand, wurden die Ergebnisse aus Szenario C) in der Folge als Referenz verwendet. Dieses Szenario umfasst alle verfügbaren Bilddaten und wurde mit geometrischen Bedingungen prozessiert, welche für alle Einzelaufnahmen eines Panoramas jeweils ein gemeinsames Projektionszentrum schätzen. Damit werden die Redundanz und die Stabilität des Bildblocks signifikant erhöht. Für die folgenden Experimente wurde die Punktwolken des südlichen Innenraums aus den drei Szenarien verglichen (Abb. 4, rechts/rot). Alle drei Punktwolken wurden vorgängig über gemeinsame natürliche Merkmale co-registriert. Der ungefähre Modellmassstab wurde über

bekannte Masse aus archäologischen Plandaten eingeführt. Abbildungen 5 und 6 zeigen die Vergleiche der Szenarien A und B mit der Referenzpunktwolke aus Szenario C.

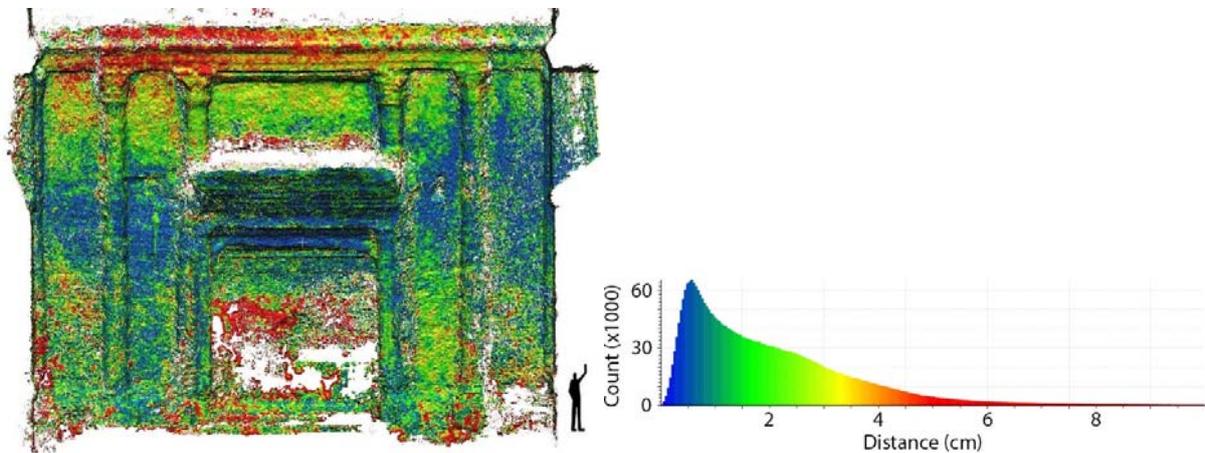


Abb. 5: Abstandsanalyse zwischen Punktwolke A (nur touristische Aufnahmen) und C (touristische Aufnahmen in Kombination mit Panorama-Einzelbildern mit geometrischen Bedigungen)

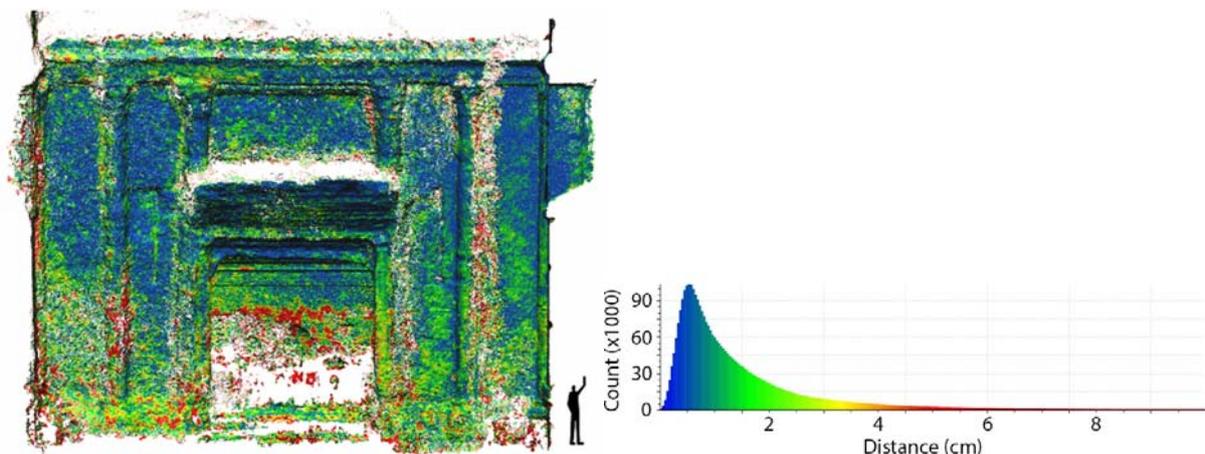


Abb. 6: Abstandsanalyse zwischen Punktwolke B (touristische Aufnahmen in Kombination mit Panorama-Einzelbildern ohne geometrischen Bedigungen) und C (Kombination mit geometrischen Bedigungen)

Abbildung 5 illustriert die Abstände zwischen der Punktwolken A und C der südlichen Innenwand. Der mittlere Abstand beträgt 2.2 cm und die Standardabweichung 3.3 cm. In diesem Vergleich sind vor allem am oberen Rand der Wand bzw. des Bildblocks noch signifikante systematische Effekte zu erkennen, die auf nicht kompensierte Fehler in der Orientierung der touristischen Aufnahmen schliessen lassen. Abbildung 6 zeigt den identischen Vergleich zwischen Punktwolke B und C mit einem mittleren Abstand von 1.5 cm und einer Standardabweichung von 1.4 cm. Dies lässt darauf schliessen, dass das Rauschen der extrahierten Puktwolke durch die Einführung der Panorama-Aufnahmen um mehr als einen Faktor 2 reduziert werden konnte.

4 Co-Registrierung mit Panoramazentren als virtuelle Passpunkte

Die Co-Registrierung der beiden inneren Punktwolken stellte einige Herausforderungen wegen fehlender Passpunkte und einer geringen Punktdichte in den überlappenden Bereichen an den West- und Ostwänden. Versuche zur Nutzung gemeinsamer natürlicher Merkmale ergaben keine befriedigenden Resultate. Die Lösung bestand darin, die Transformationsparameter zwischen den beiden Bildnetzen zu schätzen und diese für die Transformation der Punktwolken zu verwenden. Dabei dienten die fünf Panoramazentren, die in beiden Bildblöcken mitgeschätzt worden waren, als gemeinsame 'virtuelle' Passpunkte. Eine 3D-Helmerttransformation zwischen den beiden Punktgruppen ergab einen mittleren 3D-Transformationsfehler von ca. 12 cm mit einer minimalen Restklaffe von 5 cm für das zentrale Panorama (Abb. 7) und einer maximalen von 17 cm an beiden entfernten Panoramastandorten. Dieser mittlere Transformationsfehler entspricht ca. 0.3% der maximalen Dimension der Halle von ca. 40 m, was in Anbetracht der erschwerten Umstände als zufriedenstellend betrachtet werden darf.

5 Sphärische Photogrammetrie als topographische Datenbank

5.1 Sphärische Panorama-Photogrammetrie

Sphärische Panorama-Photogrammetrie (SPP) verwendet sphärische Panoramen als Bildbasis. Diese werden vorgängig aus überlappenden Einzelbildern erzeugt, die idealerweise mittels Nodaladapter zentrisch erfasst und anschliessend in einer sphärischen Projektion "gesticht" bzw. fusioniert werden (SZELINSKY & SHUM 1997). Hauptvorteile der SPP sind ein grosses Gesichtsfeld von bis zu $360^\circ \times 180^\circ$ in Kombination mit einer hohen Auflösung. Objektkoordinaten werden in der Regel via Vorwärtseinschnitt mit interaktiven Bildbeobachtungen in mind. zwei Panoramen bestimmt. SPP wird erfolgreich zur interaktiven 3D-Rekonstruktion architektonischer Kulturdenkmäler eingesetzt (PISA ET AL. 2010). Um den Mangel an Stereoskopie auszugleichen und trotzdem komplexe Oberflächen modellieren zu können, werden neue Modellierungsansätze untersucht. So werden beispielsweise die orientierten sphärischen Bildpanoramen in eine CAD-Umgebung projiziert (e.g. D'ANNIBALE et al. 2009). Für eine detailliertere Betrachtung der sphärischen Photogrammetrie sei auf FANGI & NARDINOCCHI (2013) verwiesen.

5.2 Sphärische Photogrammetrie zur Georeferenzierung von Punktwolken

Im vorliegenden Fall, wo keine Passpunkte für eine einheitliche Georeferenzierung der einzelnen Bildblöcke bzw. der daraus resultierenden 3D-Punktwolken vorliegen, bietet sich die sphärische Panorama-Photogrammetrie als Ersatzlösung an. Dabei dient die SPP einerseits zur Schätzung eines einheitlichen Panoramabildverbandes über einen Netzausgleich und andererseits zur Bestimmung von neuen 3D-Passpunkten. Zur absoluten Orientierung und möglichst genauen Skalierung des lokalen SPP-Netzes wurden die Hauptachsen des Tempels und existierende Plandaten des Tempelinnenraums ($39.45 \text{ m} \times 13.86 \text{ m}$) beigezogen.



Abb. 7: Das zentrale Panorama im Innenraum des Baaltempels.

Das photogrammetrische Netzwerk der Panoramaaufnahmen des Aussenraums ist in Abb. 8 dargestellt. Es besteht aus 13 sphärischen Panoramen verteilt rund um den Tempel und wurde mittels 1320 Bildbeobachtungen mit einem σ_0 von 0.0051 radians orientiert.

```
Number observations= 1320
Number of unknowns = 298
      redundancy = 1022
Number unk.points  =  82
Number unk.0 bear. =  13
SIGMA NAUGHT=    .005079
```

* σ_0 in radians

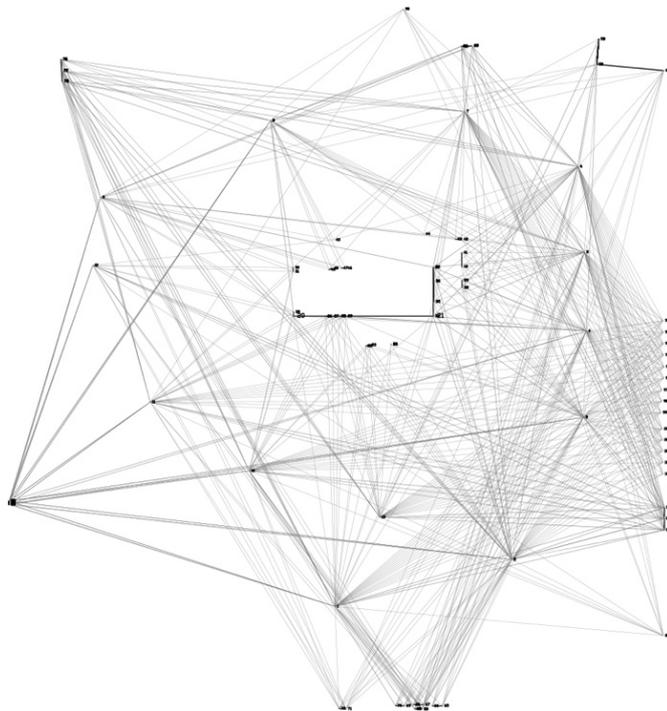


Abb. 8: Äusseres photogrammetrisches Netz der Panoramaaufnahmen

Das photogrammetrische Netz der Panoramen im Inneraum wird in Abb. 9 dargestellt. Es besteht aus insgesamt sieben Panoramen, wovon fünf den ganzen Innenraum abdecken und zwei Detailaufnahmen der südlichen Kammer liefern. Die Bündelausgleichung besteht aus 692 Beobachtungen und weist ein σ_0 von 0.0033 radians auf.

```

Number observations= 692
Number of unknowns = 363
      redundancy = 329
Number unk. points = 113
Number unk.0 bear. = 6
  SIGMA NAUGHT= .003338

```

* σ_0 in radians

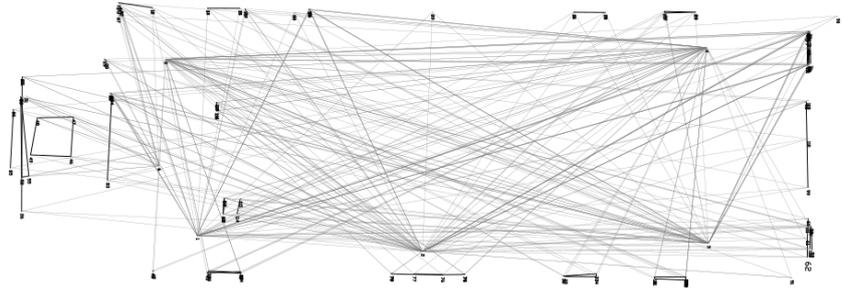


Abb. 9: Netzwerk der Panoramen im Innenraum

Im Anschluss an die Bündelorientierung liegen die Verknüpfungspunkte aus diesen sphärischen Bildblöcken als 3D-Objektpunkte vor, die als dünn besetzte 3D-Punktwolken (FANGI & NARDINOCCHI 2013) verwendet oder zu polygonalen Gittermodellen weiter verarbeitet werden können (FANGI & WAHBEH 2013). Im Baaltempel dienten diese Verknüpfungs- bzw. Neupunkte aus dem Netzausgleich zur Skalierung und Georeferenzierung der generierten Punktwolke in ein einheitliches Koordinatensystem.

6 Resultate

Die Ergebnisse der bildbasierten 3D-Rekonstruktion des Baaltempels aus einer Kombination von frei zugängliche Touristen-Aufnahmen und professionellen Panorama-Aufnahmen werden in Abbildung 10 und 11 illustriert.



Abb. 10: Das definitive 3D-Modell des Baaltempels als trianguliertes Oberflächenmodell der fusionierten Punktwolke (rechts) und Modell mit Texturen aus den Inputbilddaten (links)

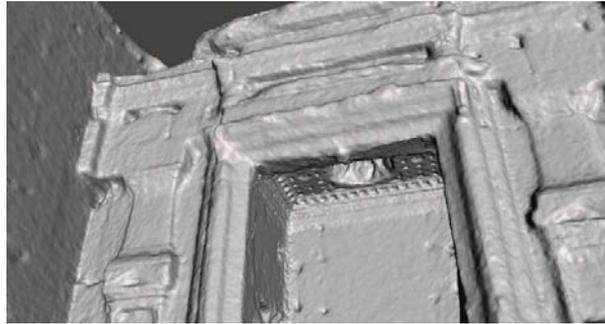


Abb. 11: Detailansicht des erzeugten Oberflächenmodells für einen Ausschnitt der nördlichen Kammer.

Mit dem vorgestellten Ansatz konnte eine weitgehend vollständige, dichte 3D-Punktwolke erzeugt werden. Ausnahmen bilden einige horizontale Flächen an Maueroberflächen und an einigen Stellen des Bodens. Die Dichte des erzeugten Oberflächenmodells und damit der Detaillierungsgrad variieren stark. Auf Grund der typischen touristischen Bilderfassungsmuster werden die 'interessanten' Teile des Monuments sehr detailreich und realitätstreu abgebildet. Dies gilt ganz besonders in Kombination mit den perfekt co-registrierten Texturdaten. Auf Grund der vorgestellten Untersuchungen, hat das resultierende 3D-Modell eine hohe metrische Qualität mit einer relativen Genauigkeit im Bereich von 2-3 cm und einer absoluten Genauigkeit über die gesamte Objektausdehnung von 10-15 cm, was in etwa 0.3% der Tempelausdehnung entspricht.

7 Fazit und Ausblick

In diesem Beitrag haben wir verschiedene Methoden und Szenarien für die bild-basierte 3D-Rekonstruktion von zerstörten oder nicht mehr zugänglichen Kulturdenkmälern vorgestellt und am Beispiel des zerstörten Baaltempels im Syrischen Palmyra praktisch untersucht. Unsere Untersuchungen umfassten einerseits eine Rekonstruktion aus ausschliesslich frei verfügbaren touristischen Bildaufnahmen aus dem Web sowie Rekonstruktionen in Kombination mit professionellen Panoramaaufnahmen. Die Rekonstruktion nur mit touristischen Aufnahmen lieferte eine gute relative Genauigkeit und einen hohen Detaillierungsgrad für die populären Teile des Monuments. Auf Grund der unvollständigen Abdeckung dieser Aufnahmen konnte allerdings der Tempel nur unvollständig rekonstruiert werden. Die Kombination von touristischen Fotos und professionellen Panoramaaufnahmen stellte sich demgegenüber als sehr mächtig heraus. Sie erlaubte eine fast vollständige 3D-Rekonstruktion des gesamten Baaltempels mit einer guten relativen und absoluten Genauigkeit und einem sehr hohen Detaillierungsgrad für die wichtigen Teile des Monuments. Die Unterstützung von geometrischen Bedingungen in der Bündelorientierung zur Schätzung eines gemeinsamen Projektionszentrums aller Aufnahmen eines jeweiligen Panoramastandorts stellte sich als sehr wertvoll heraus und resultierte in einer signifikanten Qualitätssteigerung. Um ein vollständiges, co-registriertes und georeferenziertes 3D-Modell zu erzeugen, mussten innovative Co-Registrierungsansätze gefunden werden. Dabei zeigten sich die Panoramabilder wiederum als sehr wertvoll. Einerseits dienten die gemeinsamen Projektionszentren als virtuelle Passpunkte zur Co-Registrierung von Bildblöcken mit einer geringen Punktdichte in den Überlappungsbereichen. Andererseits konnte mit einem photo-

grammetrischen Netzwerk aus sphärischen Panoramen und mit sphärischen photogrammetrischen Auswertungen das gesamte 3D-Modell korrekt georeferenziert, d.h. skaliert und in ein einheitliches lokales Bezugssystem transformiert werden.

In unserer zukünftigen Arbeit werden wir Methoden untersuchen, welche eine Quantifizierung und Visualisierung der Rekonstruktionsqualität für unterschiedliche Teile bzw. Abschnitte rekonstruierter Monumente ermöglichen. Zudem planen wir, historische Bilder in den Bildorientierungs- und 3D-Extraktionsprozess einzubeziehen mit dem Ziel, die Vollständigkeit und Genauigkeit solcher Rekonstruktionen weiter zu verbessern. Schliesslich hoffen wir, dass diese 3D-Rekonstruktionen eine nützliche Basis für eine allfällige physische Rekonstruktion des Baaltempels in friedlicheren Zeiten bieten könnte.

8 Dank

Wir bedanken uns hiermit bei allen unbekanntenen Personen, die ihre Fotos von Palmyra frei verfügbar im Internet veröffentlichten. Ein weiterer Dank gilt Gabriele Fangi für die Zurverfügungstellung der Panorama-Aufnahmen des Baaltempels.

9 Literaturverzeichnis

- AGARWAL, S., FURUKAWA, Y., SNAVELY, N., SIMON, I., CURLESS, B., SEITZ, S.M. & SZELISKI, R., 2011. Building rome in a day. *Communications of the ACM* **54** (10), 105-112.
- AGISOFT, 2016: AgiSoft PhotoScan User Manual. http://downloads.agisoft.ru/pdf/photoscan-pro_1_0_0_en.pdf
- ALSADIK, B., GERKE, M. & VOSSELMAN, G., 2013: Automated Camera Network Design for 3D Modelling of Cultural Heritage Objects. *Journal of Cultural Heritage* **14** (6), 515-526.
- ALSADIK, B., GERKE, M., VOSSELMAN, G., DAHAM, A. & JASIM, L., 2014: Minimal camera networks for 3d image based modeling of cultural heritage objects. *Sensors* **14** (4), 5785-5804.
- D'ANNIBALE, E. & FANGI, G., 2009: Interactive modelling by projection of oriented spherical panorama. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **38** (5/W1).
- CAVEGN, S., HAALA, N., NEBIKER, S., ROTHERMEL, M. & TUTZAUER, P., 2014: Benchmarking High Density Image Matching for Oblique Airborne Imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (3), 45-52.
- FANGI, G., 2007: The multi-image spherical panoramas as a tool for architectural survey. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **36** (5/C53), 311-316.
- FANGI, G. & WAHBEH, W., 2013: The destroyed minaret of Umayyad Mosque of Aleppo, the survey of the original state. *European Scientific Journal, SPECIAL*, edition **4**.
- FANGI, G. & NARDINOCCHI, C., 2013: Photogrammetric processing of spherical Panoramas. *The Photogrammetric Record* **28** (143), 293-311.

- GRUEN, A., REMONDINO, F., ZHANG, L., 2002: Reconstruction of the Great Buddha of Bamiyan, Afghanistan. *Int. Arch. of Photogrammetry and Remote Sensing* **34** (5), 363-368.
- KÁDÁR, B. & GEDE, M., 2013: Where Do Tourists Go? Visualizing and Analysing the Spatial Distribution of Geotagged Photography. *Cartographica: The International Journal for Geographic Information and Geovisualization* **48** (2), 78-88.
- 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*, (6), 727-740.
- KHAN ACADEMY, 2016: Temple of Bel, Palmyra. www.khanacademy.org/humanities/ancient-art-civilizations/palmyra/a/temple-of-bel-palmyra
- KOUTSOUDIS, A., VIDMAR, B., IOANNAKIS, G., ARNAOUTOGLU, F., PAVLIDIS, G. & CHAMZAS, CH., 2014: Multi-image 3D reconstruction data evaluation. *Journal of Cultural Heritage* **15** (1), 73-79.
- NEBIKER, S., LACK, N. & DEUBER, M., 2014: Building Change Detection from Historical Aerial Photographs Using Dense Image Matching and Object-Based Image Analysis. *Remote Sensing* **6** (9), 8310-8336.
- REMONDINO, F., DEL PIZZO, S., KERSTEN, TH., 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)*, Volume **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.
- SZELISKI, R. & SHUM, H.-Y., 1997: Creating full view panoramic image mosaics and texture-mapped models. *Computer Graphics (SIGGRAPH'97 Proceedings)*, 251-258.
- WAHBEH, W., NARDINOCCHI, C., 2015: *Toward the Interactive 3D Modelling Applied to Ponte Rotto in Rome*. Nexus Network Journal Vol.1 Springer Basel
- WAHBEH, W. & NEBIKER, S., 2016: Combining Public Domain And Professional Panoramic Imagery For The Accurate And Dense 3d Reconstruction Of The Destroyed Bel Temple In Palmyra. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **3** (5).

Segmentation and Localization of Individual Trees from MMS Point Cloud Data Acquired in Urban Areas

MARTIN WEINMANN^{1,2}, CLÉMENT MALLET¹ & MATHIEU BRÉDIF¹

Abstract: In this paper, we address tree segmentation and localization in the scope of the IQmulus Processing Contest IQPC'15. Based on the part of pre-classified 3D point cloud data which corresponds to trees, we present a novel framework which involves a downsampling of the original data, a projection of the downsampled data onto a horizontally oriented plane, a mean-shift-based segmentation of the projected points, a transfer of the segmentation results to the original data, a refinement of the segmentation results via segment-based shape analysis, and a localization of respective tree trunks. The results derived for a benchmark dataset reveal that all individual trees are correctly detected and localized with both acceptable accuracy and reasonable computational effort.

Zusammenfassung: In diesem Beitrag wird die Segmentierung und Lokalisierung von Bäumen im Rahmen des IQmulus Processing Contest IQPC'15 behandelt. Basierend auf jenem Anteil von vorab klassifizierten 3D-Punktwolken, welcher Bäume repräsentiert, wird eine Methodik vorgestellt, die eine teilweise Betrachtung der Originaldaten, eine 2D-Projektion entsprechender Punkte auf eine horizontal ausgerichtete Ebene, eine Segmentierung der projizierten Punkte basierend auf dem Mean-Shift-Verfahren, einen Transfer der Segmentierungsergebnisse zu den Originaldaten, eine Verbesserung der Segmentierungsergebnisse basierend auf einer Analyse der Segmentform und eine Lokalisierung entsprechender Baumstämme umfasst. Die für einen Referenz-Datensatz erreichten Ergebnisse zeigen auf, dass alle darin enthaltenen Bäume mit ausreichender Genauigkeit und angemessenem Rechenaufwand erkannt und lokalisiert werden können.

1 Introduction

The automated analysis of 3D point cloud data is of great interest in photogrammetry and remote sensing. Amongst a variety of research topics, the detection of individual trees from 3D point cloud data has been paid particular attention in recent years. In the context of forest applications, for instance, airborne laser scanning data has been used at tree-level to delineate individual crowns, to estimate wood volumes and biomass, or to identify the respective tree species. In contrast, mobile mapping systems allow a ground-based acquisition of significantly denser 3D point cloud data which are meanwhile used to derive tree cadasters in urban environments, e.g. for applications in city modeling and city planning.

In the IQmulus Processing Contest IQPC'15, point cloud analysis has recently been addressed with a special track on tree separation and classification (GORTE et al. 2015). Thereby, the focus is on 3D point cloud data acquired with mobile mapping systems and, hence, mainly data

¹ Université Paris-Est, IGN, SRIG, MATIS, 73 avenue de Paris, 94160 Saint-Mandé, France, E-Mail: [martin.weinmann, clement.mallet, mathieu.bredif]@ign.fr

² Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Englerstraße 7, 76131 Karlsruhe, Germany, E-Mail: martin.weinmann@kit.edu

acquired along roads and in urban environments. While the first step consists in a classification into 3D points corresponding to trees and other 3D points, the second step consists in separating those 3D points corresponding to the “tree” class into clusters referring to the respective individual trees (Fig. 1).

Since the emphasis of IQPC’15 is on tree separation, we focus on a straightforward and efficient approach for individual tree segmentation and localization in the scope of this paper, so that scalability towards the processing of even larger datasets is given. In contrast to approaches relying on a voxelization of 3D space to reduce the computational burden, we present an approach which considers the data on point-level (which is highly relevant for an accurate delineation of individual trees and also for tree parameter extraction) but increases computational efficiency by conducting a typically rather time-consuming mean-shift-based segmentation only for a specific subspace as well as a downsampled version of the original data. The obtained segmentation results are then transferred to the original data. A subsequent refinement of the segmentation results is carried out via a segment-based shape analysis and yields plausible tree segments. Finally, the tree location is estimated by detecting a respective tree trunk. We demonstrate the performance of our approach for a recent benchmark dataset acquired in an urban environment, and we discuss the achieved results with regard to competing approaches.

The paper is structured as follows. After briefly discussing related work in Section 2, we present our proposed methodology for tree segmentation and localization in Section 3. Subsequently, we provide experimental results in Section 4 and discuss these results in detail in Section 5. Finally, in Section 6, we provide concluding remarks and suggestions for future work.

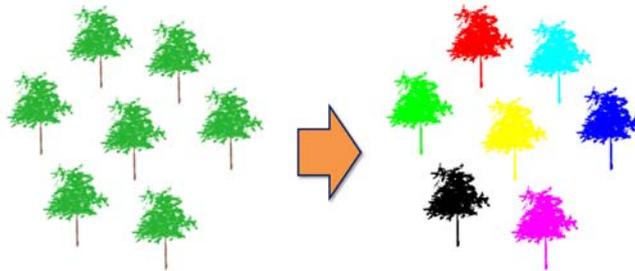


Fig. 1: Schematic illustration for the segmentation of single trees from all 3D points classified as tree points. All those 3D points not belonging to the “tree” class are assumed to have been removed by the previous step of tree classification.

2 Related Work

Point cloud segmentation aims to provide a meaningful partitioning of a set of 3D points into smaller, connected subsets which correspond to objects of interest or to parts of these (MELZER 2007; VOSSELMAN 2013). In the context of our work, we deal with 3D point cloud data acquired with a mobile mapping system, and the derived segments should represent individual trees.

Many approaches for detecting single trees from 3D point cloud data rely on a voxelization of 3D space. In YAO & FAN (2013), it is for instance proposed to derive a 2D accumulation map on a horizontally oriented plane and – based on respective features – to separate natural objects such

as trees from man-made objects. Those 3D points corresponding to natural objects are transferred to a voxel space and then a normalized cut segmentation based on the voxel structure is carried out (REITBERGER et al. 2009). A different strategy for detecting and segmenting individual trees consists in first classifying laser scanning point clouds into “tree” and “non-tree” classes (SIRMACEK & LINDENBERGH 2015), and subsequently focusing on tree individualization. The latter may be achieved by a voxelization of 3D space, deriving connected components and separating the components further if they contain multiple clusters (GORTE et al. 2015). Alternatively, tree individualization may be achieved by a downsampling and retiling of the original 3D point cloud data via voxelization, where a subsequent 2D gridding allows to find local maxima in point density and thus potential tree locations (LINDENBERGH et al. 2015). Based on these tree locations, individual trees are finally segmented via octree-based region growing and thresholding techniques.

Instead of involving a voxelization of 3D space, it has been proposed to perform tree individualization by considering the original data on point-level and calculating geometric descriptors for each 3D point, projecting these descriptors onto a horizontally oriented 2D accumulation map and considering a spatial filtering to obtain individual tree clusters (MONNIER et al. 2012). A further approach for tree individualization on point-level focuses on deriving connected components for those 3D points categorized into a “tree” class, and the connected components are further split via an upward and downward growing algorithm if there are multiple seeds at a height between 0.5m and 1m (GORTE et al. 2015; OUDE ELBERINK & KEMBOI 2014). Furthermore, a direct consideration of the original data on point-level is possible by applying a standard clustering technique such as k -means clustering or hierarchical clustering (GUPTA et al. 2010), or the mean shift algorithm presented in (FUKUNAGA & HOSTETLER 1975). The latter has for instance been applied on 3D point cloud data in (FERRAZ et al. 2012; SCHMITT et al. 2013; YAO et al. 2013; SHAHZAD et al. 2015). However, such an approach can be computationally demanding, particularly for a large number of considered 3D points. To improve computational efficiency, it seems desirable to apply the mean shift algorithm on a 2D projection of a 3D point cloud as e.g. described in (SCHMITT et al. 2015) for 3D point cloud data acquired via tomographic SAR processing. However, the point density will be significantly higher for a ground-based acquisition of 3D point cloud data as e.g. given when using mobile mapping systems. Hence, handling the amount of available data still remains challenging due to the computational burden arising from data processing and data storage.

3 Methodology

Our proposed framework for tree segmentation and localization is illustrated in Fig. 2. It consists of 6 successive steps represented by (1) a downsampling of the original data, (2) a projection of the downsampled data onto a horizontally oriented plane, (3) a mean-shift-based segmentation of the projected points, (4) a transfer of the segmentation results to the original data, (5) a refinement of the segmentation results via segment-based shape analysis, and (6) a localization of respective tree trunks. The input of the framework is given by 3D points which – in a previous step – have been classified to correspond to a tree (Fig. 3), and the output of the framework is

given by individual tree segments and the location of respective tree trunks. More details on the single components are provided in the following subsections.

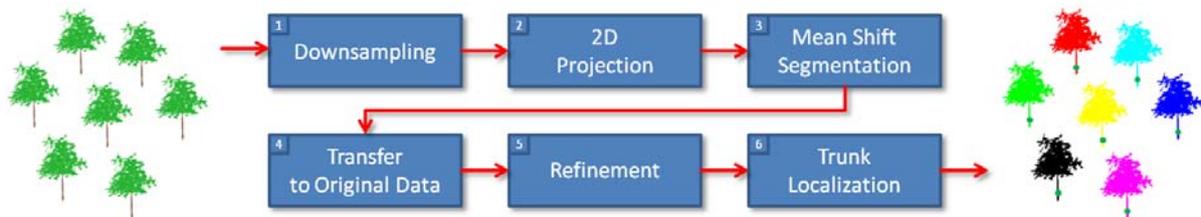


Fig. 2: The proposed framework for segmenting and localizing individual trees from 3D point cloud data acquired with a mobile mapping system: the input data are downsampled and projected onto a horizontally oriented plane; a mean-shift-based segmentation is subsequently carried out in 2D space and results are transferred to the original data; a shape-based refinement yields plausible tree segments and respective trunks are finally localized for these plausible segments.

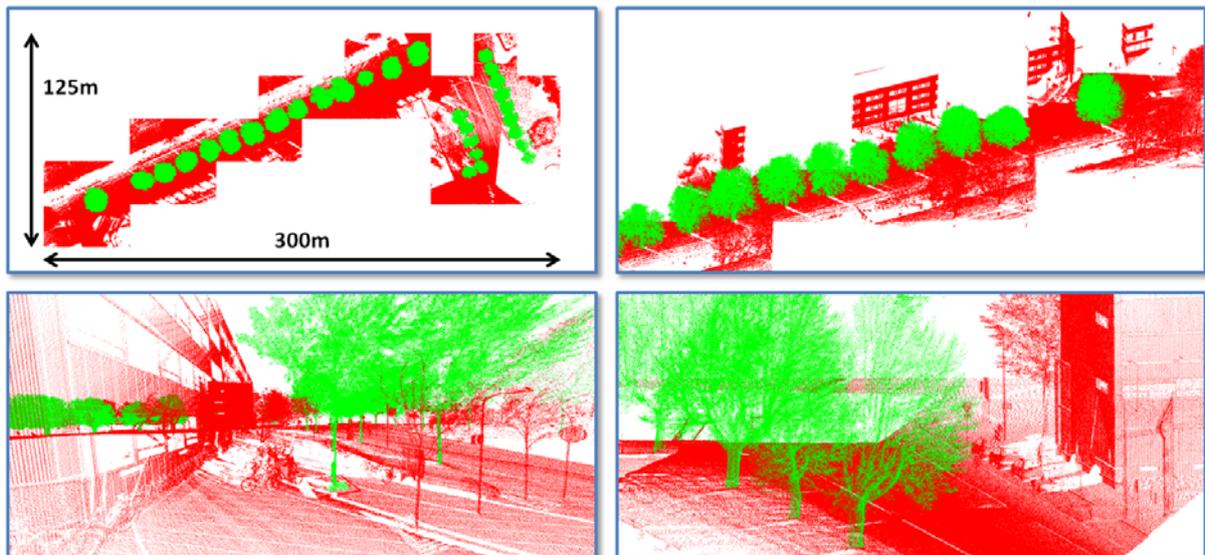


Fig. 3: Pre-classified 3D point cloud data (top row: nadir view and side view; bottom row: more detailed views): 3D points categorized into the “tree” class are indicated in green and serve as input for our framework, whereas all other 3D points are indicated in red and not considered by our framework.

3.1 Downsampling of the Original 3D Point Cloud

Taking into account that MMS point cloud data provides a very dense representation of object surfaces near the acquisition system, it might be desirable to conduct time-consuming tasks only for a downsampled version of the original data to reduce the computational effort with respect to processing time and memory consumption. In this regard, it also becomes obvious that the point density of 3D points corresponding to trees may significantly be decreased while still being able to detect individual trees in the respective 3D point cloud data. Accordingly, we conduct a downsampling by only keeping every k -th point of the original 3D point cloud. In the results, we

will report results for $k = 10$, a parameter selected heuristically. Alternatively, a pruning of this parameter could be conducted based on the local point density (CARAFFA et al. 2015).

3.2 Projection Onto Horizontally Oriented Plane

In contrast to forested areas, urban areas typically provide a larger spacing and less overlap between individual trees due to human intervention in nature and different planning processes. Accordingly, we propose to neglect the occurrence of dominant, co-dominant or dominated trees in urban environments and assume that individual trees may still sufficiently be delineated when only considering a 2D projection of the downsampled 3D point cloud data corresponding to trees onto a horizontally oriented plane.

3.3 Segmentation via Mean Shift Algorithm

In order to derive a meaningful partitioning of the downsampled 3D point cloud data, we consider the respective 2D projections on the horizontally oriented plane and apply the mean shift algorithm (FUKUNAGA & HOSTETLER 1975; CHENG 1995; COMANICIU & MEER 2002; MELZER 2007) for segmentation. Generally, the mean shift algorithm represents an iterative, statistical technique for locating the maxima / modes of a probability density function by only considering discrete data sampled from that probability density function. There is no need to explicitly estimate the underlying probability density function itself, and the mean shift algorithm does neither rely on the assumption of a specific geometric model nor require an initial definition of the number of modes.

In our case, we treat the 2D projections of the downsampled 3D point cloud data as discrete data points sampled from an empirical 2D probability density function. For each data point, the mean shift algorithm iteratively (1) calculates the weighted mean of data points within a window defined by a kernel K (typically an isotropic kernel such as a Gaussian kernel or an Epanechnikov kernel (COMANICIU & MEER 2002)), (2) defines the mean shift vector as the difference between the data point and the weighted mean of data points within the considered window, and (3) moves the data point along the mean shift vector. In this context, the step-size defined by the length of the mean shift vector will be large in areas of low point density, whereas it will be low in areas of high point density. In this way, the mean shift algorithm iteratively performs an adaptive gradient ascent until convergence (up to numerical accuracy). The stationary points correspond to regions of high point density and represent the modes of the underlying distribution of data points. Finally, all data points leading to the same mode are considered as cluster or segment. For our application, the single clusters / segments are expected to represent the individual trees in the considered scene.

The only parameters to be specified are represented by the bandwidth parameters indicating the size of the kernel in each direction. Due to the consideration of data points on the horizontally oriented plane, we may assume the same bandwidth h in all directions, which is intuitively justified since we want to detect individual trees. Furthermore, we may involve prior knowledge about the shape and size of the trees in the considered scene and hence heuristically select the bandwidth parameter. Based on different tests, we heuristically select a bandwidth parameter of $h = 3.8\text{m}$ for the involved dataset.

3.4 Transfer of Segmentation Results to Original Data

To get from the segmentation results for the downsampled version of the original data back to the original 3D point cloud data, we consider an intuitive, simple and straightforward approach. This approach assigns each 3D point of the original 3D point cloud data the segment label of the closest 3D point in the downsampled version of the original data. Thereby, we conduct a nearest neighbor search based on Euclidean distances. As a result, we have transferred the segmentation results from the downsampled data to the original 3D point cloud data.

3.5 Refinement of Segmentation Results

The previous steps lead to a coarse segmentation of the considered 3D point cloud data, yet they do not guarantee that the derived segments are plausible. For our application, we want to have meaningful segments corresponding to individual trees. It becomes obvious that – for a segment corresponding to an individual tree – the top parts of the segment are characterized by foliage, whereas the lower parts are characterized by the tree trunk. We hence focus on checking each derived segment for a trunk-like structure at its lower parts. For this purpose, we consider the lowest 3D point of the respective segment as basis and cut the part up to the diameter of breast height (which may vary between 1.3m and 1.5m depending on the tree type and depending on the country) into horizontally oriented slices of a certain width. More specifically, we select 6 slices covering the lowest 1.5m of each segment, i.e. each slice has a width of 0.25m. Since circle fitting tends to be impractical because of only partly visible trunks and the fact that possibly occurring trunk radii would have to be pre-defined for a model-based approach, we focus on an approximate data-driven solution. We first calculate the mean value m_i of all 2D coordinates per slice i and then define the mean value M across these means as potential location of the tree trunk. Subsequently, we take into account the standard deviation σ of the distance between slice-wise means m_i and the potential location M of the tree trunk. If the standard deviation σ becomes too large (e.g. $\sigma > 0.2\text{m}$) or if the derived clusters comprise a relatively low number N_p of points (e.g. $N_p < 100$ points), we assume that the respective segment does not correspond to a tree. In this case, the segment is either merged with the next-closest plausible segment in case of directly adjacent segments or, otherwise, removed from the set of 3D points which have been classified to correspond to a tree.

3.6 Localization of Tree Trunks

In the last step, we consider each plausible segment for an individual tree and define the location of the corresponding tree trunk to coincide with the respectively estimated mean position M across the 6 lowest slices defined for that segment.

4 Experimental Results

The involved benchmark dataset has been acquired in the vicinity of the campus of TU Delft in the Netherlands with the Fugro DRIVE-MAP system (GORTE et al. 2015). For our experiments, we consider the provided subset comprising 26 tiles, for which results of a previous binary classification are available (Fig. 3). The 26 tiles contain a total number of about 10.13M points of which about 1.78M points (17.6%) are assigned to the “tree” class, and we intend to subdivide

the set of all these 3D points into segments corresponding to individual trees (i.e. we address the Challenge 1 of the IQPC'15 special track on tree separation and classification).

For our experiments, we conduct a downsampling by only keeping every 10th point of the original 3D point cloud. Then, we perform the 2D projection and apply the mean shift algorithm with a bandwidth parameter of $h = 3.8\text{m}$. The results obtained for the benchmark dataset are shown in Fig. 4. With our approach, we have found a total number of 29 trees in the scene and a visual inspection reveals that all the derived segments correspond to individual trees. Thus, the segmentation of individual trees is reached with an accuracy of 100%. A more detailed analysis of the derived results reveals that, due to the localization of trees based on considering the lowest slices of plausible segments, all estimated tree locations coincide with a respective tree trunk. Note that, for the considered benchmark dataset, these estimates are biased towards the position of the sensor, since only a part of the tree trunks is acquired with the mobile mapping system. However, we consider the estimate to be sufficiently accurate for our application.

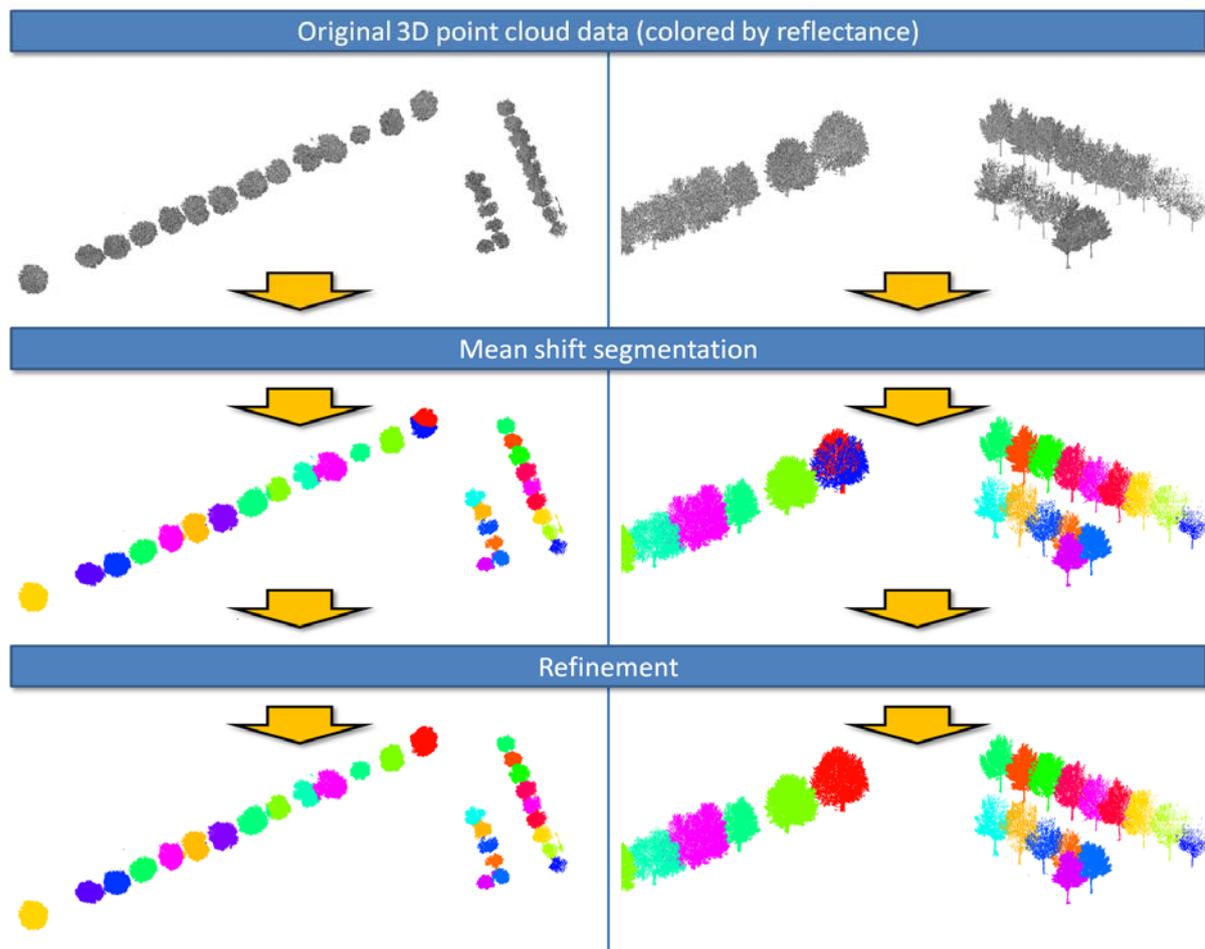


Fig. 4: The involved dataset in nadir view (left column) and in a side view (right column): original 3D point cloud data (top row), segmentation result after applying the mean shift algorithm on a 2D projection of a downsampled version of the original data and the transfer back to the original data (center row), and segmentation result after the shape-based refinement (bottom row).

The required computational effort for the whole methodology comprising downsampling, 2D projection, mean shift segmentation, transfer of segmentation results to original data, refinement and trunk localization is less than 1 minute on a notebook of medium performance (Intel Core i5-2410M, 2.3GHz, 4GB RAM, Matlab implementation).

5 Discussion

As already pointed out in (MELZER 2007) in the context of point cloud segmentation, one of the most important advantages of applying the mean shift algorithm consists in its capability to directly work on the original data. Since this may however be rather time-consuming for large datasets, we apply the mean shift algorithm only on a downsampled version of the original data and subsequently transfer the segmentation results back to the original data. Thereby, we further involve a 2D projection, since the computational effort of a mean shift segmentation in 2D is significantly lower than for a mean shift segmentation in 3D (FERRAZ et al. 2012; SCHMITT et al. 2013). Thus, our strategy reduces the computational burden considerably without significantly affecting the quality of the derived segmentation results.

Besides representing an efficient approach for individual tree segmentation and localization, our approach is rather simple and easy-to-use. It directly works on the given data without relying on a voxelization as e.g. presented in (GORTE et al. 2015; LINDENBERGH et al. 2015), where the voxel size as well as the voxel orientation might strongly influence the respective segmentation results. Considering the derived segmentation results, we may state that our results are sufficiently accurate for the benchmark dataset and that only minor segmentation errors at segment borders are visible if adjacent trees are relatively close to each other. However, the latter also becomes visible in the results presented in (GORTE et al. 2015).

A disadvantage of the proposed approach consists in the fact that it relies on one critical parameter represented by the bandwidth parameter h which has a strong influence on the number of clusters. However, since it has a physical meaning, we have selected a suitable value by involving heuristic knowledge about the expected size and shape of trees in the scene which is in accordance to (SCHMITT et al. 2015). Furthermore, it is obvious that the results of the presented approach strongly depend on the quality of the previous separation between 3D points corresponding to trees and 3D points corresponding to other objects, but this also holds for the competing approaches presented in (GORTE et al. 2015; LINDENBERGH et al. 2015).

6 Conclusions & Future Work

In this paper, we presented a framework for tree segmentation and localization representing an important topic in the scope of the IQmulus Processing Contest IQPC'15. The input to our framework consists in 3D points which – in a previous step – have been classified to correspond to a tree. Our framework is composed of (1) a downsampling of the original data, (2) a projection of the downsampled data onto a horizontally oriented plane, (3) a mean-shift-based segmentation of the projected points, (4) a transfer of the segmentation results to the original data, (5) a refinement of the segmentation results via segment-based shape analysis, and (6) a localization of respective tree trunks. The results derived for a benchmark dataset reveal that all individual

trees are correctly detected and localized with acceptable accuracy and reasonable computational effort. Only minor segmentation errors may be observed at the boundaries of some segments, but these errors are in the same order as in related investigations (GORTE et al. 2015).

For future work, we intend to address the fundamental requirement for our framework, i.e. the appropriate separation of 3D points belonging to trees from 3D points belonging to other objects. Particularly the use of geometric features e.g. as the ones presented in (WEINMANN et al. 2015a,b) seems to be promising for this task, since respective geometric features allow an appropriate separation of 3D points corresponding to foliage, trunk and ground (STUCKER 2015).

7 Acknowledgements

This work was supported by the European Commission's Seventh Framework Programme under grant agreement FP7-ICT-2011-318787 (*IQmulus: A High-Volume Fusion and Analysis Platform for Geospatial Point Clouds, Coverages and Volumetric Data Sets*).

8 References

- CARAFFA, L., BRÉDIF, M. & VALLET, B., 2015: 3D octree based watertight mesh generation from ubiquitous data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (3/W3), 613-617.
- CHENG, Y., 1995: Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17** (8), 790-799.
- COMANICIU, D. & MEER, P., 2002: Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24** (5), 603-619.
- FERRAZ, A., BRETAR, F., JACQUEMOUD, S., GONÇALVES, G., PEREIRA, L., TOMÉ, M. & SOARES, P., 2012: 3-D mapping of a multi-layered Mediterranean forest using ALS data. *Remote Sensing of Environment* **121**, 210-223.
- FUKUNAGA, K. & HOSTETLER, L., 1975: The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory* **21** (1), 32-40.
- GORTE, B., OUDE ELBERINK, S., SIRMACEK, B. & WANG, J., 2015: IQPC 2015 Track: Tree separation and classification in mobile mapping lidar data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (3/W3), 607-612.
- GUPTA, S., WEINACKER, H. & KOCH, B., 2010: Comparative analysis of clustering-based approaches for 3-D single tree detection using airborne fullwave lidar data. *Remote Sensing* **2** (4), 968-989.
- LINDENBERGH, R.C., BERTHOLD, D., SIRMACEK, B., HERRERO-HUERTA, M., WANG, J. & EBERSBACH, D., 2015: Automated large scale parameter extraction of road-side trees sampled by a laser mobile mapping system. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (3/W3), 589-594.
- MELZER, T., 2007: Non-parametric segmentation of ALS point clouds using mean shift. *Journal of Applied Geodesy* **1** (3), 159-170.

- MONNIER, F., VALLET, B. & SOHEILIAN, B., 2012: Trees detection from laser point clouds acquired in dense urban areas by a mobile mapping system. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **1** (3), 245-250.
- OUDE ELBERINK, S. & KEMBOI, B., 2014: User-assisted object detection by segment based similarity measures in mobile laser scanner data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (3), 239-246.
- REITBERGER, J., SCHNÖRR, CL., KRZYTEK, P. & STILLA, U., 2009: 3D segmentation of single trees exploiting full waveform lidar data. *ISPRS Journal of Photogrammetry and Remote Sensing* **64** (6), 561-574.
- SCHMITT, M., BRÜCK, A., SCHÖNBERGER, J. & STILLA, U., 2013: Potential of airborne single-pass millimeterwave InSAR data for individual tree recognition. *Publications of the German Society for Photogrammetry, Remote Sensing and Geoinformation, Volume 22*, Potsdam, Seyfert, E. (Ed.), 33. *Wissenschaftlich-Technische Jahrestagung der DGPF*, 27 February - 1 March 2013, Freiburg, 427-436.
- SCHMITT, M., SHAHZAD, M. & ZHU, X. X., 2015: Reconstruction of individual trees from multi-aspect TomoSAR data. *Remote Sensing of Environment* **165**, 175-185.
- SHAHZAD, M., SCHMITT, M. & ZHU, X. X., 2015: Segmentation and crown parameter extraction of individual trees in an airborne TomoSAR point cloud. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XL** (3/W2), 205-209.
- SIRMACEK, B. & LINDENBERGH, R., 2015: Automatic classification of trees from laser scanning point clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2** (3/W5), 137-144.
- STUCKER, C., 2015: Klassifizierung von Vegetation in einer Laserscan-Punktwolke. *Publications of the German Society for Photogrammetry, Remote Sensing and Geoinformation, Volume 24*, Potsdam, Kersten, T.P. (Ed.), 35. *Wissenschaftlich-Technische Jahrestagung der DGPF*, 16-18 March 2015, Cologne, 370-375.
- VOSELMAN, G., 2013: Point cloud segmentation for urban scene classification. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (7/W2), 257-262.
- 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., URBAN, S., HINZ, S., JUTZI, B. & MALLET, C., 2015b: Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas. *Computers & Graphics* **49**, 47-57.
- YAO, W. & FAN, H., 2013: Automated detection of 3D individual trees along urban road corridors by mobile laser scanning systems. *Proceedings of the International Symposium on Mobile Mapping Technology (MMT) 2013*, 1-6.
- YAO, W., KRZYTEK, P. & HEURICH, M., 2013: Enhanced detection of 3D individual trees in forested areas using airborne full-waveform lidar data by combining normalized cuts with spatial density clustering. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2** (5/W2), 349-354.

Gipuma: Massively Parallel Multi-view Stereo Reconstruction

SILVANO GALLIANI¹, KATRIN LASINGER¹ & KONRAD SCHINDLER¹

Abstract: We describe a method for dense multi-view matching and 3D point cloud reconstruction, which has been developed at the Institute of Geodesy and Photogrammetry of ETH Zurich. The method generates high-quality point clouds from oriented images very efficiently, by massively parallel processing on graphics processing units (GPUs). It is available as open-source code. Technically, the method is an extension of the PatchMatch Stereo algorithm: 3D depth values and surface normal vectors are iteratively propagated across the image and refined, in order to find a maximally photo-consistent depth map and normal field for each view. Photo-consistency is computed in a slanted tangent plane, such that the reconstruction does not suffer from fronto-parallel bias. We extend PatchMatch to 3D scene space, such that photo-consistency can be aggregated over multiple views, which allows for more robust and more accurate depth estimation. Moreover, the sequential propagation is replaced by a local, diffusion-like scheme, such that the computation can be massively parallelised. All computations are local, thus computation time is linear in the image size and inversely proportional to the number of parallel threads. Moreover, memory requirements are also modest, since only four values per pixel must be stored. Experiments on benchmark datasets show that our method delivers point clouds (respectively, surfaces) with high accuracy and completeness, across a range of applications.

1 Introduction

In this work we present a method to reconstruct dense, accurate 3D point clouds from oriented images. 3D shape reconstruction is a fundamental task of photogrammetry. Given that automatic camera orientation and point triangulation are essentially solved - at least for images taken in such a way that they are suitable for dense surface reconstruction - the most challenging step is to find a dense set of corresponding points between different images. The majority of the literature deals with the basic two-view stereo setup (e.g. INTILLE & BOBICK 1994; SCHARSTEIN & SZELISKI 2002; RANFTL et al. 2012; RHEMANN et al. 2011), but obviously using more than two views affords redundancy and is thus beneficial in terms of both accuracy and robustness. More viewpoints also mitigate the occlusion problem and typically yield more complete reconstructions (e.g. CAMPBELL et al. 2008; FURUKAWA & PONCE 2010).

The proposed method is a variant of the *PatchMatch Stereo* algorithm (BLEYER et al. 2011). On one hand, we adapt *PatchMatch* to operate in 3D scene space, such that the method can directly establish correspondence across multiple viewpoints. On the other hand, we modify the algorithm in such a way that it can be parallelised into thousands of computational threads on graphics processing units (GPUs).

¹ ETH Zürich, Photogrammetry and Remote Sensing, Stefano-Francini-Platz 5, CH-8093 Zürich, E-Mail: [silvano.galliani, katrin.lasinger, schindler]@geod.baug.ethz.ch

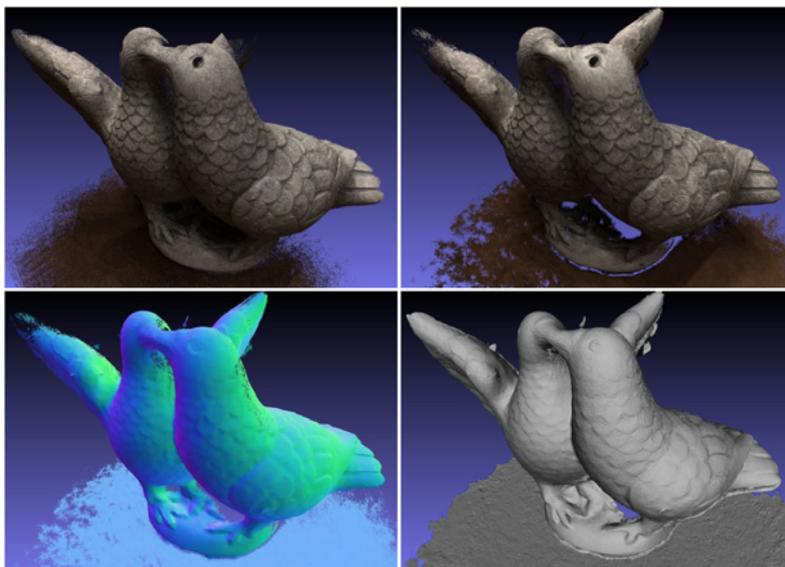


Fig. 1: Example reconstruction. *Top left*: Ground truth point cloud. *Top right*: reconstructed point cloud with texture. *Bottom left*: colour-coded surface normal. *Bottom right*: reconstructed surface.

1.1 Background and Related Work

Dense image matching seeks to establish correspondence densely for all pixels by maximizing photo-consistency. There are two main schools: *local methods*, which derive correspondences only from information in a local neighbourhood, and *global methods*, which formulate an objective function over all pixels. The simplest local method is naive block matching with a window around each pixel. Starting from there, various ideas have been developed to mitigate problems induced by the larger neighbourhood window, in particular the bias towards fronto-parallel surfaces (BLEYER et al. 2011; BURT et al. 1995; DEVERNAY & FAUGERAS 1994; EINECKE & EGGERT 2013; GALLUP et al. 2007), and the fattening of foreground objects along their silhouettes (FUSIELLO et al. 1997; KANADE & OKUTOMI 1994; YOON & KWON 2006). With local methods, smoothness of the reconstructed surface is accounted for only implicitly through the overlap of nearby support windows, or by local filtering of the depth map. Explicit models of smoothness lead to global methods: the correlations between nearby pixels give rise to global objective functions defined over the entire image domain, which can be (approximately) maximised either by discrete labeling (e.g., FELZELSZWALB & HUTTENLOCHER 2006; HIRSCHMÜLLER 2008; KOLMOGOROV & ZABIH 2001) or by variational inference, (e.g. RANFTL et al. 2012). The size of modern digital images calls for matching algorithms with low computational complexity (ideally at most linear in the number of pixels) and low memory footprint, especially in the multi-view case. This has led to a renewed interest in local methods, and on benchmark datasets like DTU (JENSEN et al. 2014) or KITTI (GEIGER et al. 2012) local matchers have reached competitive results (BLEYER et al. 2011; FURUKAWA & PONCE 2010; RHEMANN et al. 2011).

Multi-view matching goes back to at least (OKUTOMI & KANADE 1993), where matching costs (deviations from photo-consistency) are accumulated over different stereo pairs. Another way to exploit multiple views is to iteratively grow a 3D point cloud from reliable seed points (e.g.,

FURUKAWA & PONCE 2010). Computationally efficient multi-view methods are often based on plane sweeping, i.e. moving a plane through 3D object space along its normal and exhaustively comparing the photo-consistency across multiple views for all possible plane positions (e.g. COLLINS 1996; GALLUP et al. 2007; HU & MORDOHAI 2012).

Another way to classify multi-view stereo methods is by the representation they are based on (SEITZ et al. 2006). The 3D scene can be represented by voxels, level-sets, polygon meshes, or depth maps. We point out that depth maps are a point-wise representation, in the sense that they define a 2.5D point cloud, similar to those generated with scanning devices. The remaining three representations additionally solve (at least implicitly) the step from discrete points to surfaces. Here we aim to reconstruct depth maps and fuse them into 3D point clouds, whereas surface fitting is seen as a post-processing step. If needed, we found that the generic Poisson surface reconstruction method (KAZHDAN & HOPPE 2013) works well with our point clouds, which by construction already include per-point surface normals.

2 Gipuma – Depth Map Computation from Multiple Views

The proposed multi-view matcher is based on the *PatchMatch* (BARNES et al. 2009) principle, which employs iterative propagation and refinement of randomly drawn initial values to quickly find a good solution in a large search space without having to test all possibilities. The resulting low memory footprint makes *PatchMatch Stereo* well-suited for large images or memory-constrained environments, including GPUs (BAILER et al. 2012; BAO et al. 2014; HEISE et al. 2013; Zheng et al. 2014). On the contrary a large portion of the computation time and/or memory footprint of many stereo algorithms is due to the fact that they exhaustively compute, and then compare, the cost for every putative disparity value (e.g. FELZENSZWALB & HUTTENLOCHER 2006; HIRSCHMÜLLER 2008; RHEMANN et al. 2011).

2.1 PatchMatch Stereo

We first briefly describe the original *PatchMatch Stereo* algorithm. That method parameterizes the 3D scene by a local tangent plane per pixel, i.e. at every pixel one stores not only a disparity (horizontal parallax), but also a surface normal. These values are initialised randomly for each pixel. Starting at the top left corner one then loops through the image and tests whether it makes sense to extend the local tangent plane to the lower or right neighbour pixel. If propagating the plane lowers the matching cost, then the corresponding disparity and normal replace the previous values. After the update the plane parameters are refined with fast bisection search. Having reached the lower right corner, the propagation is repeated in the opposite direction. Empirically, only two to three such iterations are needed to converge to a good result.

PatchMatch Stereo computes the photo-consistency in the tangent plane in disparity space, defined by the local disparity and normal, to avoid fronto-parallel bias. The method is compatible with different photo-consistency measures. In difficult lighting conditions the Hamming distance between Census signatures (ZABIH & WOODFILL 1994) is often the best choice. For more controlled lighting we found that the measure recommended in the original *PatchMatch Stereo* paper works very well, namely a linear combination of the absolute intensity difference and the absolute difference in gradient magnitude. The weighted average of these two quantities

(truncated for robustness) is computed separately per pixel and summed over the support window, using individual weights per pixel. These weights reflect the intensity difference to the central pixel of the window, assuming that discontinuities in intensity are likely to coincide with surface discontinuities.

Empirically we found that one can speed up the computation without loss in performance by using only every other row and column in the cost computation, as recommended in the context of the “sparse Census transform” (ZINNER et al. 2008). This is particularly valuable because *PatchMatch* needs larger support windows (typically between 11×11 and 25×25 pixels) than some other stereo methods, in order to reliably estimate the surface normal.

2.2 Multi-view Extension

Disparity is only defined between two images rectified to the photogrammetric normal case. We aim to adapt *PatchMatch* to the multi-view setting. We therefore define the tangent planes per pixel of the reference image in Euclidean scene space rather than in disparity space. This not only avoids epipolar rectification (which is a rather awkward procedure for more than two views), but also ensures that the estimated normal reside in 3D scene space, and can be used for later processing steps such as surface reconstruction (KAZHDAN & HOPPE 2013). Most importantly, working in scene space makes it possible to aggregate matching costs directly across multiple views, by warping all support windows with the same 3D tangent plane with the corresponding plane-induced homographies, see Figure 2. For a given reference image we always accumulate the pairwise matching costs to all stereo partners.

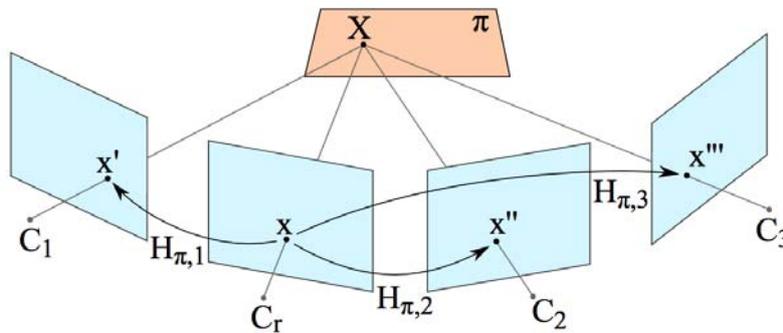


Fig. 2: Multi-view geometry of *Gipuma*: a tangent plane π in 3D scene space defines a planar homography between the reference view C_r and every other view C_i .

A small technical complication when working in scene space is to sample unbiased random normals as initialization. A fast method to pick random vectors that are uniformly distributed over a sphere can be found in (MARSAGLIA 1972). If the sampled normal points away from the camera, we reverse its sign. Moreover, it is better to sample uniformly from the range of possible disparities and convert the result to depth, rather than sampling depth values directly. This will account for the anisotropic depth resolution of stereo reconstruction and provide a finer sampling of depth values in the near field, where they make a difference; and a sparser sampling in the far field, where small variations do not produce an observable difference.

2.3 Surface Normal Diffusion

The original *PatchMatch* propagation scheme described above is inherently sequential. It has been proposed to alternate between row-wise and column-wise propagation (BAILER et al. 2012; BAO et al. 2014; HEISE et al. 2013; ZHENG et al. 2014), still one does not play to the strengths of modern multi-core GPU. Instead, we employ a diffusion-like propagation scheme. The reference image is partitioned into a checkerboard pattern of “red” and “black” pixels, see Figure 3. In this way, one can update all “black” pixels at once by propagating from nearby “red” ones, and vice versa. The red-black scheme is a standard trick to parallelize message-passing and related schemes, c.f. red-black Gauss-Seidel for linear equation solving. Even with this more “local” propagation scheme we find that 6-8 iterations are sufficient for the depth map to converge.

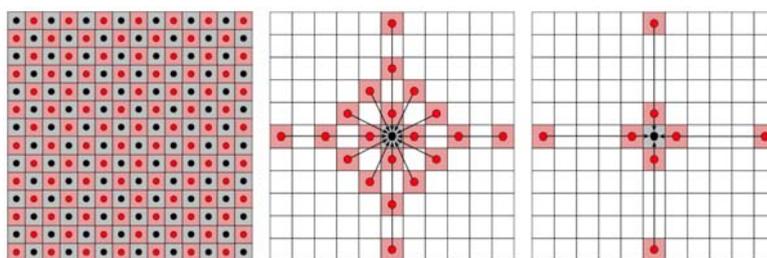


Fig. 3: Red-black scheme for massively parallel propagation. All black pixels are updated in parallel, by propagating tangent planes from the red pixels. *Center*: standard pattern. *Right*: reduced pattern optimised for speed.

3 Fusibile – 3D Point Cloud Generation

Like most multi-view reconstruction methods, we first compute a separate depth map for each view by repeatedly running *Gipuma*. These depth maps are then fused into a 3D point cloud, using the redundancy to eliminate blunders and improve accuracy. Given multiple depth maps covering the same object surface, most mismatches can be detected: if a 3D point is wrong due to a lack of texture or due to occlusion, then it is unlikely that it is consistent with the measurement in another view (even if that measurement is wrong, too). To filter out such cases, we visit each depth map in turn, convert the depth measurements to 3D points and project those points back to the other views. If the projection is not consistent with the observed disparity and normal in at least two other views, the point is removed. For those points, which are retained, the 3D positions and normals are averaged over the consistent views to further suppress noise.

The thresholds that determine whether a point’s reprojection is consistent enough to be added to the point cloud can be tuned, to achieve more accurate or more complete reconstructions. The right trade-off depends on the application, for example computer graphics typically needs complete models that “look correct”, whereas industrial metrology prefers high accuracy even at the cost of a sparser reconstruction. We test different options in the experiments. Since the fusion step always uses the same depth maps it is fast (≈ 15 seconds for 50 depth maps of size 1600×1200), so one can even interactively explore different possibilities.

4 Experiments

As a main dataset for quantitative evaluation we use the DTU large scale multi-view dataset (JENSEN et al. 2014). It consists of 80 different objects, each covered by 49-64 images of resolution 1600×1200 pixels. The scenes vary in reflectance, texture and geometry, see Figure 4. The images have been acquired with a robot arm to accurately position the cameras. Ground truth is captured by a structured light scanner. The error metrics defined by the authors of the dataset are mean and median reconstruction errors, calculated once for the 3D point cloud and once for a surface mesh derived from the points. Accuracy is defined as the distance from the estimated surface to the ground truth, and completeness as the distance from the ground truth to the surface (both measured in mm).

To find suitable stereo partners, we select images with reasonable angular baseline: for a given reference image, we chose all other views whose viewing directions differ by at least 10 and at most 30 degrees from the reference view. *Gipuma* in its high-accuracy setting achieves the best accuracy among the tested methods while at the same time delivering the second-highest completeness, see Table 1. When tuned for completeness, our method achieves the highest completeness at the second-highest accuracy. The runtime is ≈ 50 seconds per depth map with 9 stereo partners. We also test an extreme setting tuned for high speed (window size 15 instead of 25, 6 iterations instead of 8, at most 9 stereo partners per view). With these settings it takes ≈ 2.7 seconds to compute a depth map, respectively 7 minutes for a complete reconstruction, (including disk I/O). Even this version is competitive with the state of the art, see Table 1.



Fig. 4: Example objects from DTU dataset. *Left*: ground truth point cloud. *Center*: reconstructed point cloud. *Right*: reconstructed surface mesh.

We further evaluate *Gipuma* on the popular Middlebury benchmark ([//vision.middlebury.edu/](http://vision.middlebury.edu/)), although we note that it is by now rather saturated, and somewhat outdated (only 2 objects, image size 640×480 pixels). Our results rank second for the “full” version (363 views) of “Dino” and fifth for the “full” version (312 views) of the “Temple”. For the smaller “ring”

(48/47 views) and “sparse” (16 views) versions of both objects we are also always in the top 8 out of ca. 50 submitted results. To generate a depth map from 10 views our method needs 2.5 seconds. Finally, we have also run *Gipuma* on oblique aerial images and on KITTI (cameras mounted in a vehicle), see qualitative examples in Figure 5.

Table 1: Quantitative comparisons on DTU data. Mean and median errors in mm (lower is better).

points	accuracy		completeness	
	mean	median	mean	median
<i>Gipuma</i> acc.	0.273	0.196	0.687	0.260
<i>Gipuma</i> comp.	0.379	0.234	0.400	0.188
<i>Gipuma</i> fast	0.289	0.207	0.841	0.285
TOLA et al. 2010	0.307	0.198	1.097	0.456
FURUKAWA & PONCE 2010	0.605	0.321	0.842	0.431
CAMPBELL et al. 2008	0.753	0.480	0.540	0.179

surfaces	accuracy		completeness	
	mean	median	mean	median
<i>Gipuma</i> acc.	0.363	0.215	0.766	0.329
<i>Gipuma</i> comp.	0.631	0.262	0.519	0.309
<i>Gipuma</i> fast	0.358	0.221	0.939	0.350
TOLA et al. 2010	0.488	0.244	0.974	0.382
FURUKAWA & PONCE 2010	1.299	0.534	0.702	0.405
CAMPBELL et al. 2008	1.411	0.579	0.562	0.322

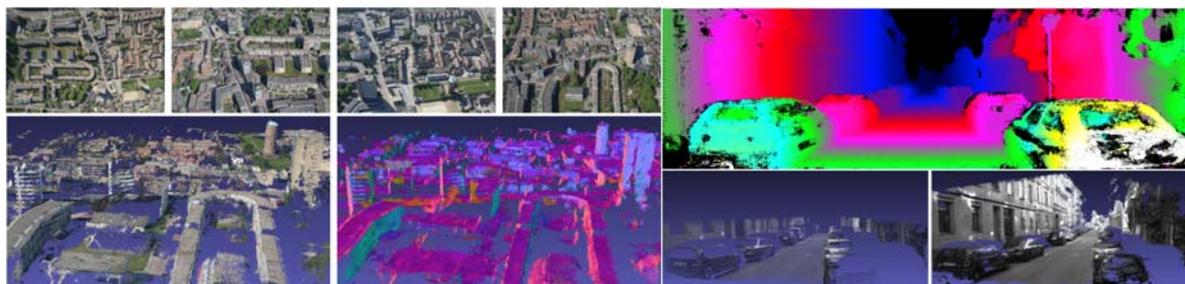


Fig. 5: Further *Gipuma/Fusibile* results. *Left*: Oblique aerial images. *Right*: Vehicle-mounted cameras from KITTI.

5 Conclusion

The *Gipuma/Fusibile* multi-view stereo suite has been designed to provide high-quality point clouds from multi-view imagery. It is optimised for parallel computation on off-the-shelf GPUs. Technically, the system is a multi-view extension of *PatchMatch Stereo* tailored to modern GPUs with thousands of parallel threads. Like the original *PatchMatch* it uses slanted support windows and does not suffer from fronto-parallel bias. It is thus particularly well-suited for locally smooth scenes with a large extent in depth. Quantitative results on the DTU and Middlebury benchmarks confirm that the software is at the same time more accurate than state-of-the-art methods such as PMVS and much faster. *Gipuma* and *Fusibile* are available as open-source software at <https://github.com/kysucix/gipuma>, respectively <https://github.com/kysucix/fusibile>.

6 Bibliography

- BAILER, C., FINCKH, M. & LENSCH, H.P., 2012: Scale robust multi view stereo. European Conference on Computer Vision (ECCV), Springer Berlin Heidelberg, 398-411.
- BAO, L., YANG, Q. & JIN, H., 2014: Fast edge-preserving Patchmatch for large displacement optical flow. IEEE Conference on Computer Vision and Pattern Recognition, 3534-3541.
- BARNES, C., SHECHTMAN, E., FINKELSTEIN, A. & GOLDMAN, D.B., 2009: PatchMatch: A randomized correspondence algorithm for structural image editing. ACM Transactions on Graphics **28** (3), 24.
- BLEYER, M., RHEMANN, C. & ROTHER, C., 2011: PatchMatch Stereo - stereo matching with slanted support windows. British Machine Vision Conference (BMVC) **11**, 1-11.
- BURT, P., WIXSON, L. & SALGIAN, G., 1995: Electronically directed "focal" stereo. 5th International Conference on Computer Vision (ICCV), 94-101.
- CAMPBELL, N.D., VOGIATZIS, G., HERNÁNDEZ, C. & CIPOLLA, R., 2008: Using multiple hypotheses to improve depth-maps for multi-view stereo. European Conference on Computer Vision (ECCV), 766-779.
- COLLINS, R.T., 1996: A space-sweep approach to true multi-image matching. Computer Vision and Pattern Recognition (CVPR), 358-363.
- DEVERNAY, F. & FAUGERAS, O., 1994: Computing differential properties of 3-d shapes from stereoscopic images without 3-d models. Computer Vision and Pattern Recognition (CVPR), 208-213.
- EINECKE, N. & EGGERT, J., 2013: Stereo image warping for improved depth estimation of road surfaces. Intelligent Vehicle Symposium (IVS), 189-194.
- FELZENSZWALB, P.F. & HUTTENLOCHER, D. P., 2006: Efficient belief propagation for early vision. International Journal of Computer Vision, **70**(1), 41-54.
- FURUKAWA, Y. & PONCE, J., 2010: Accurate, dense, and robust multiview stereopsis. IEEE Transactions on Pattern Analysis and Machine Intelligence, **32**(8), 1362-1376.
- FUSIELLO, A., ROBERTO, V. & TRUCCO, E., 1997: Efficient stereo with multiple windowing. Computer Vision and Pattern Recognition (CVPR), 858.
- GALLUP, D., FRAHM, J.-M., MORDOHAJ, P., YANG, Q. & POLLEFEYS, M., 2007: Real-time plane-sweeping stereo with multiple sweeping directions. Computer Vision and Pattern Recognition (CVPR), 1-8.
- GEIGER, A., LENZ, P. & URTASUN, R., 2012: Are we ready for autonomous driving? The KITTI vision benchmark suite. Computer Vision and Pattern Recognition (CVPR), 3354-3361.
- HEISE, P., KLOSE, S., JENSEN, B. & KNOLL, A., 2013: PM-Huber: PatchMatch with Huber regularization for stereo matching. International Conference on Computer Vision (ICCV), 2360-2367.
- HIRSCHMÜLLER, H., 2008: Stereo processing by semi-global matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence **30**(2), 328-341.
- HU, X. & MORDOHAJ, P., 2012: Least commitment, viewpoint-based, multi-view stereo. 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 531-538.

- INTILLE, S.S. & BOBICK, A. F., 1994: Disparity-space images and large occlusion stereo. European Conference on Computer Vision (ECCV), Springer Berlin Heidelberg, 179-186.
- JENSEN, R., DAHL, A., VOGIATZIS, G., TOLA, E. & AANÆS, H., 2014: Large scale multi-view stereopsis evaluation. Computer Vision and Pattern Recognition (CVPR), 406-413.
- KANADE, T. & OKUTOMI, M., 1994: A stereo matching algorithm with an adaptive window: Theory and experiment. IEEE Transactions on Pattern Analysis and Machine Intelligence **16** (9), 920-932.
- KAZHDAN, M. & HOPPE, H., 2013: Screened Poisson surface reconstruction. ACM Transactions on Graphics **32** (3), 29.
- KOLMOGOROV, V. & ZABIH, R., 2001: Computing visual correspondence with occlusions using graph cuts. International Conference on Computer Vision (ICCV), 508-515.
- MARSAGLIA, G., 1972: Choosing a point from the surface of a sphere. Annals of Mathematical Statistics **43** (2), 645-646.
- OKUTOMI, M. & KANADE, T., 1993: A multiple-baseline stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence **15** (4), 353-363.
- RANFTL, R., GEHRIG, S., POCK, T. & BISCHOF, H., 2012: Pushing the limits of stereo using variational stereo estimation. Intelligent Vehicles Symposium (IVS), 401-407.
- RHEMANN, C., HOSNI, A., BLEYER, M., ROTHER, C. & GELAUTZ, M., 2011: Fast cost-volume filtering for visual correspondence and beyond. Computer Vision and Pattern Recognition (CVPR), 3017-3024.
- SCHARSTEIN, D. & SZELISKI, R., 2002: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision **47** (1-3), 7-42.
- SEITZ, S.M., CURLESS, B., DIEBEL, J., SCHARSTEIN, D. & SZELISKI, R., 2006: A comparison and evaluation of multi-view stereo reconstruction algorithms. Computer Vision and Pattern Recognition (CVPR), 519-528.
- TOLA, E., LEPETIT, V. & FUA, P., 2010: Daisy: An efficient dense descriptor applied to wide-baseline stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence **32** (5), 815-830.
- YOON, K.-J. & KWEON, I.S., 2006: Adaptive support-weight approach for correspondence search. IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (4), 650-656.
- ZABIH, R. & WOODFILL, J., 1994: Non-parametric local transforms for computing visual correspondence. European Conference on Computer Vision (ECCV), Springer Berlin Heidelberg, 151-158.
- ZHENG, E., DUNN, E., JOJIC, V. & FRAHM, J.-M., 2014: PatchMatch based joint view selection and depthmap estimation. Computer Vision and Pattern Recognition (CVPR), 1510-1517.
- ZINNER, C., HUMENBERGER, M., AMBROSCH, K. & KUBINGER, W., 2008: An optimized software-based implementation of a census-based stereo matching algorithm. Advances in Visual Computing, 216-227.

CityGML Semantik unter der Lupe – welche Konsequenzen ergeben sich für die Validierung von 3D Stadtmodellen?

DETLEV WAGNER¹, EGBERT CASPER² & VOLKER COORS²

Zusammenfassung: CityGML ist als semantisch-geometrisches Datenmodell für 3D Stadtmodelle bekannt. Im Zuge der Verbesserung der Datenqualität von 3D Stadtmodellen ist die Entwicklung von Prüfroutinen für valide Geometrie inzwischen weit fortgeschritten. Semantische Anteile der Modelle bleiben in der Regel jedoch unbeachtet, obwohl das semantische Modell neben der Geometrie eigenständiger und wesentlicher Bestandteil eines Datensatzes ist. Die Kohärenz der beiden nebeneinander bestehenden Modelle ist dabei wesentlicher Bestandteil von CityGML.

Ein Konzept zur unabhängigen Validierung des semantischen Modells auf Grundlage von eindeutigen Regeln wird im Paper vorgestellt. Es kann um die Validierung geometrisch-semantischer Kohärenz erweitert werden, worauf an Hand konkreter Beispiele Bezug genommen wird.

Das Paper geht von der Semantik von Objekten der realen Welt aus und konkretisiert diese an Hand eines thematischen Beispiels, wie etwa Gebäude, und stellt Bedingungen für ein valides Gebäude auf. Diese Bedingungen umfassen die Existenz, Anzahl sowie geometrische Ausprägung und Anordnung von Bestandteilen des Gebäudes, wie etwa den Wänden, dem Dach und der Grundfläche. Dies dient als Grundlage zur Ableitung von eindeutigen Regeln, welche dem semantischen Konzept zugrunde liegen und die im semantischen Konzept von CityGML ihre Entsprechung finden sollten. Kann keine eindeutige Regel abgeleitet werden, so ist die Bedingung in der vorliegenden Form nicht eindeutig und kann bei der Validierung keine Verwendung finden.

Mit Hilfe des semantischen Modells von CityGML kann dann eine eindeutige Formalisierung der Regeln abgeleitet werden, die eine automatisierte Validierung unterstützt. Die Eindeutigkeit der Formalisierung ermöglicht somit erst unabhängige und nachvollziehbare Validierungsergebnisse. Beispielhafte Regeln, auch mit Bezug zur geometrisch-semantischen Kohärenz, werden vorgestellt und erforderlichenfalls Parameter und Toleranzen eingeführt sowie die Umsetzung in Prüfroutinen dargestellt. Die Grenzen derart eindeutiger und klar ableitbarer semantischer Restriktionen werden abschließend diskutiert.

1 Einleitung

3D-Stadtmodelle existieren inzwischen in einer Vielzahl von Varianten und Datenformaten, wobei sich CityGML als gemeinsamer Standard u.a. zum Datenaustausch etabliert hat. CityGML stellt ein ausgereiftes geometrisches und semantisches Datenmodell bereit. Die in den Modellen enthaltene semantische Information ist in vielen Anwendungsbereichen entscheidend, auch im Hinblick auf deren kohärentem Zusammenspiel mit der Geometrie. Die Sicherstellung der Kohärenz der beiden Bereiche liegt jedoch allein in der Verantwortung des Datenerstellers und wird nicht implizit vom CityGML-Datenmodell gefordert. Vielmehr verweist das

¹ Universität Teheran, Department of Cartography, E-Mail: dwagner4@gmx.at

² Hochschule für Technik Stuttgart, Fakultät Vermessung, Informatik und Mathematik, Schellingstr. 24, 70174 Stuttgart, E-Mail: [egbert.casper, volker.coors]@hft-stuttgart.de

Standarddokument in Kapitel 6.3 auf die Bedeutung dieses Umstandes als einem der wichtigsten Gestaltungsleitsätze von CityGML (GRÖGER et al. 2012: 12). Daher ist die Validierung nicht nur hinsichtlich des geometrischen Modells erforderlich, sondern darauf aufbauend die Validierung der Semantik sowie der geometrisch-semantischen Kohärenz.

Ausgehend von der Semantik als Bedeutungslehre (Abschnitt 3) und sehr allgemein gehaltenen Gebäudedefinitionen (Abschnitt 4) werden Aussagen identifiziert, die teilweise durch Interpretation (Abschnitt 5) zur Grundlage von präzisen Anforderungen werden. In Abschnitt 6 werden diese mit den vom Building-Modul von CityGML bereitgestellten Features abgeglichen. Darauf aufbauend werden in Abschnitt 7 Regeln formuliert, die im Abschnitt 8 in geeigneter Weise formalisiert werden.

2 Problemstellung

Bei Datenmodellen wie X3D und glTF, die in der Computergraphik entwickelt wurden, stehen Geometrie und visuelle Erscheinung (Photometrie) im Vordergrund. Die Bedeutung der Geometrie wird i.d.R. allenfalls durch eine Assoziation von Semantik und Geometrie definiert. Im Gegensatz dazu steht in CityGML ein semantisches Modell im Vordergrund, räumliche Ausdehnung und Photometrie sind dem semantischen Modell als Attribute untergeordnet. Allerdings ist auch in CityGML keine formale Definition der Bedeutung von verwendeten Begriffen wie „Building“ gegeben. Für die automatisierte Validierung der semantisch-geometrischen Kohärenz ist eine formale Spezifikation dieser Begriffe aber unabdingbar. Bisherige Ansätze zur Validierung von CityGML Modellen prüfen beispielsweise, ob eine Gebäudegeometrie geschlossen ist. Auch das geometrische Modell eines Schokoladenhasen würde bei dieser Prüfung als valide Gebäudegeometrie akzeptiert werden.

Im vorliegenden Paper wird ein neuartiger Ansatz vorgestellt, der ausgehend von einer formalen Spezifikation von Begriffen Prüfregeln ableitet, mit der semantisch-geometrische Kohärenz validiert werden kann. Dies wird anhand des Beispiels *Gebäude* nachvollzogen und exemplarisch dargestellt.

Die Vorstellungen darüber, was ein Gebäude ist, wie es aussieht und strukturiert sein sollte, wo es sich befindet und wie seine Umgebung beschaffen sein kann, können individuell sehr weit auseinander gehen. Um diese Vorstellungen auf eine verbindliche Ebene zu bringen, ist es notwendig, einen Katalog von grundlegenden Anforderungen zu definieren, der die meisten Fälle abdecken kann und zugleich nicht so allgemein formuliert ist, dass die Variationsmöglichkeiten unüberschaubar werden. Auch im Hinblick auf standardisierte Qualitätsanforderungen und Interoperabilität ist dies erforderlich.

Es gilt herauszuarbeiten, wie weit diese Anforderungen gefasst werden können, um noch eine sinnvolle Modellierung zu gewährleisten, und ein Abdriften in die Beliebigkeit zu vermeiden.

3 Semantik und semantische Modelle

Semantik im weiteren Sinne beschäftigt sich mit der Bedeutung von Zeichen. In der Linguistik sind diese Zeichen Phoneme bzw. Grapheme, in der Informationstechnologie sind damit

weiterführende domänenspezifische Konzepte gemeint. Semantische Modelle dienen somit vornehmlich der konzeptuellen Modellierung.

Im Bereich der 3D Stadtmodelle spielen sowohl semantische als auch geometrische Modelle eine Rolle. Beides sind konzeptuelle Modelle, die einen bestimmten fachspezifischen Aspekt beschreiben. Während das geometrische Modell ein Konzept zur Modellierung der räumlichen Eigenschaften von Objekten definiert, umfasst das zugehörige semantische Modell die Definition der jeweiligen Bedeutung dieser Objekte und ihrer Beziehungen zueinander. Beide Modelle können isoliert voneinander existieren, werden aber in der Praxis meist kombiniert. Da sie den selben Gegenstand zum Thema haben, ist eine Grundforderung für einen validen Datensatz die Kohärenz von geometrischer und semantischer Modellierung (STADLER & KOLBE 2007).

4 Semantik von Gebäuden

Jedes Modell stellt im Idealfall eine für die fachspezifische Anwendung geeignetes Konzept zur Abstraktion eines Objektes der Realwelt dar. Im Folgenden wird am Beispiel des Gebäudes nachvollzogen, wie dieser Abstraktionsprozess verlaufen kann, und welche Schlüsse sich daraus für das Modell ziehen lassen. Zunächst wird dazu eine Definition für das Gebäude benötigt, die z.B. so aussehen kann:

Definition 1:

„Gebäude sind selbstständig benutzbare, überdeckte bauliche Anlagen, die von Menschen betreten werden können und geeignet oder bestimmt sind, dem Schutz von Menschen, Tieren oder Sachen zu dienen“ (BAUMINISTERKONFERENZ 2002).

Es gibt weitere Definitionen, die jedoch ähnlich allgemein gehalten sind. Die Ableitung von allgemeinen Eigenschaften gestaltet sich angesichts der Vielzahl vorkommender Gebäudearten i.d.R schwierig. Diese Vielzahl umfasst alle Varianten zwischen einem simplen quaderförmigen Einraumgebäude und hochkomplexen Bauten wie etwa Kathedralen oder repräsentativen öffentlichen Gebäuden. Hinzu kommen noch die Variationen in der Gebäudegröße, die sich zwischen wenigen Metern und mehreren Hundert Metern Seitenlänge und Höhe bewegen kann. Dennoch werden all diese Varianten unter ein und demselben Konzept Gebäude zusammengefasst. Es könnten also weitere Gemeinsamkeiten existieren, die eine Subsummierung unter einem Dach rechtfertigen und über den allgemeinen Ansatz der obigen Definition hinausgehen.

Die obigen Anmerkungen zum Variantenreichtum bei Gebäuden legen den Fokus unbewusst auf deren geometrische Ausprägung. Man wird wohl große Schwierigkeiten haben, auf diesem Wege Gemeinsamkeiten herauszuarbeiten. Anders sieht es aus, wenn man die Semantik hinzunimmt.

Dann lassen sich unmittelbar weitere Anforderungen aus der Definition ableiten:

1. Gebäude sind *selbstständig benutzbar*
Ein Gebäude hat eigenständigen Charakter, es erfüllt seinen Zweck unabhängig von seiner Umgebung. Dazu sollte es ein in sich abgeschlossenes Objekt sein.
2. Gebäude sind *überdeckt*
Das Gebäude besitzt ein Dach. Wände oder Boden sind demnach nicht zwingend.
3. Gebäude sind *bauliche Anlagen*
Im Unterschied zu natürlich vorkommenden Objekten ist ein Gebäude ein künstlich errichtetes Objekt.

4. Gebäude *können von Menschen betreten werden*
Eine Mindestgröße muss gegeben sein, um menschlichen Zugang ungehindert zu gewährleisten.

5. Gebäude sollen *Schutz bieten*
Gebäude dienen einem bestimmten Zweck. Hier wird auf die Funktion des Gebäudes verwiesen, die Schutz im weitesten Sinne für Menschen, Tiere und Sachen bedeutet.

Dass die Definition immer noch sehr weitreichend ist, zeigt sich an Gebäuden wie überdachten Haltestellen oder Wachhäuschen, die ebenso darunter fallen wie etwa große offene Lagerstellen, z.B. für Baumaterialien, solange diese eine Überdachung besitzen.

Ein anderer Ansatz wird von der Definition bei ALKIS verfolgt:

Definition 2:

“Gebäude' ist ein dauerhaft errichtetes Bauwerk, dessen Nachweis wegen seiner Bedeutung als Liegenschaft erforderlich ist sowie dem Zweck der Basisinformation des Liegenschaftskatasters dient” (ALKIS).

Hier liegt der Fokus auf der Funktionsweise des zugrundeliegenden Systems und weniger beim Gebäude selbst. Das Gebäude ist in diesem Fall lediglich ein wegen seiner Bedeutung als Liegenschaft zu verwaltes Objekt. Das Gebäude besitzt somit zur Verwaltung erforderliche Fachinformationen.

Dennoch lassen sich weitere Anforderungen ableiten:

6. Das Gebäude wurde *dauerhaft* errichtet,
7. und ist *ortsfest* (da es sich um eine Liegenschaft handelt).
8. Zudem muss es eine nicht näher spezifizierte *Bedeutung* haben, um als Liegenschaft qualifiziert zu sein. Letztlich ist diese Bedeutung länderabhängig und ergibt sich aus der Anwendung der rechtlichen Rahmenbedingungen für das Katasterwesen.

Dies sind nur zwei von vielen möglichen Definitionen zur konkretisierten Darstellung des Konzeptes Gebäude. Man kann deutlich erkennen, dass entsprechend dem Anwendungsbereich der Fokus der jeweiligen Definition unterschiedlich ausfallen kann. Entsprechend verschiebt sich der Bedeutungsgehalt der Definition und somit des Definiendums Gebäude. Unterschiedliche Definitionen desselben Gegenstandes bedingen demnach semantische Abweichungen. Diese können sich ergänzen, in manchen Fällen jedoch auch widersprüchlich sein. Letztlich ist der Anwendungsbereich maßgeblich, um zu entscheiden, ob die Definition eines semantischen Konzepts für diesen geeignet ist.

Unser Anwendungsbereich ist die 3D-Modellierung einer Stadt. Bestehende Konzepte basieren in der Regel auf formalisierten Modellbeschreibungen, wie etwa KML, CityGML oder Openstreetmap3D. Diese enthalten implizite oder explizite Definitionen der zu modellierenden Objekte, in jedem Fall aber exakte Definitionen der sie repräsentierenden Modelle. Am Beispiel von CityGML wird dies im folgenden Abschnitt erläutert.

5 Regeln für valide Semantik

Die Ableitung von Regeln aus den in Abschnitt 4 genannten Bedingungen und Einschränkungen muss in einer Art und Weise geschehen, die nachvollziehbar von den eindeutig abgeleiteten Aussagen zu eindeutig definierten und überprüfaren Regeln führt. Im Rahmen des CityGML

Quality Interoperability Experiments des OGC (COORS & WAGNER 2015) wurde ein Prozessmodell nach der Business Process Model Notation (BPMN) Spezifikation entwickelt, welches dies für CityGML Conformance Requirements sicherstellt. Dieses Modell kann in seiner allgemeinen Form dazu dienen, um aus den in Abschnitt 4 genannten Bedingungen (Test Requirements, TR) konkrete Regeln für valide Semantik zu formulieren. (Abb. 1).

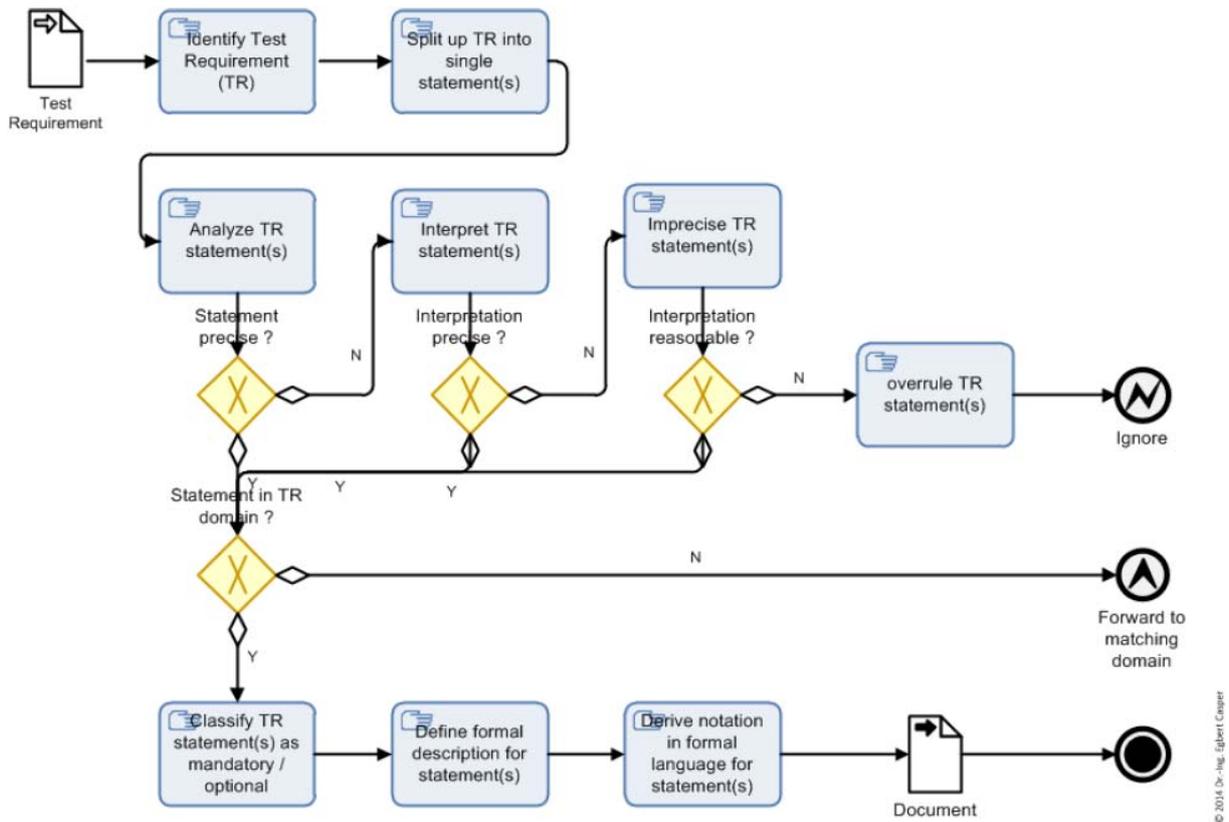


Abb. 1: BPMN-Prozessmodell zur Analyse von Testbedingungen (© E.Casper).

Der erste Schritt ist die Identifizierung von Bedingungen, die ein Gebäude definieren. Diese Bedingungen wurden am Beispiel zweier Definitionen in Abschnitt 4 identifiziert und in einzelne jeweils mit einer Nummer versehene Aussagen aufgeteilt. Jede einzelne Aussage wird im Folgenden auf ihre Eignung als eindeutig überprüfbares Kriterium untersucht (siehe Tab. 1).

Tab. 1: Analyse der Gebäudedefinitionen.

Nr.	präzise?	interpretierbar?	vernünftig?	Bereich
1	Nein	Ja	Ja	Geometrie/Semantik
2	Ja	-	-	Geometrie/Semantik
3	Nein	Ja	Ja	Semantik
4	Nein	Ja	Ja	Geometrie/Semantik
5	Nein	Ja	Ja	Semantik
6	Ja	-	-	Semantik
7	Ja	-	-	Geometrie
8	Nein	Ja	Ja	Semantik

Die meisten Kriterien bedürfen einer weiteren Interpretation, da sie ansonsten zu wenig präzise sind. Die Interpretation kann in allen Fällen für den Bereich der 3D Stadtmodelle so klar gefasst werden, dass die Ableitung von formalen Regeln möglich sein sollte. Die Unterscheidung der Regeln in verpflichtende oder optionale Regeln ist hier nicht erforderlich, da die Zugehörigkeit zum Bereich des Konzeptes *Gebäude* unstrittig ist. Wichtig ist vielmehr, ob die jeweilige Aussage einen rein semantischen Bezug hat oder auch die Geometrie betrifft. Bevor eine weitere Ableitung formaler Regeln erfolgt, wird noch der Zusammenhang zwischen Realwelt-Gebäuden und dem Building-Modell von CityGML untersucht.

6 Semantik von *Building*-Features in CityGML

Die Semantik erstreckt sich im Rahmen von CityGML auf die Ausprägungen der jeweils modellierten Objekte (etwa *Building*) sowie deren Teilen (z.B. *BuildingInstallation*). Die semantische Ausprägung wird hier also zum einen als Label genutzt, um ein bestimmtes Objekt mit einem Typ aus dem vorhandenen Katalog der CityGML-Features zu klassifizieren, zum anderen wird auf diesem Weg die Aggregationshierarchie beschrieben (eine bestimmte *BuildingInstallation* ist Teil eines bestimmten *Building*-Features). Die semantische Beschreibung von Objekten umfasst weiterhin Attribute (z.B. *function*) und Relationen.

Eine explizite Definition des zugrundeliegenden Realwelt-Objekts *Gebäude* existiert in diesem Zusammenhang jedoch nicht. Ebenso wird der Anwendungsbereich des Konzepts implizit durch Verwendung von umgangssprachlichen Begriffen wie *building* oder *room* festgelegt, die teilweise nur in bestimmten Level of Detail (LOD) verwendet werden dürfen. So sind z.B. *BoundarySurface* nur für die LOD 2-4 zulässig. Das bedeutet zugleich einen semantischen Informationsmangel für Modelle in LOD 0 und 1, der Hand in Hand mit deren geometrischer Vereinfachung geht.

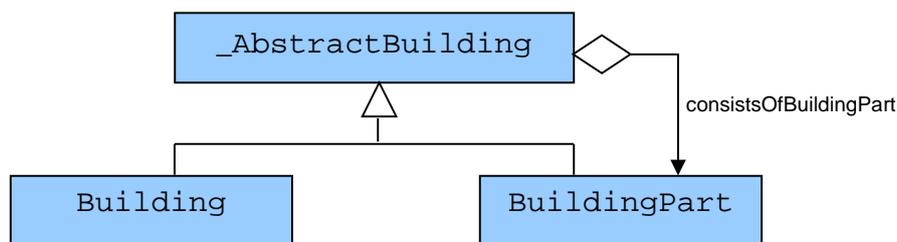


Abb. 2: Die Klasse *_AbstractBuilding* des *Building*-Moduls von CityGML.

Die Klasse *Building* ist neben der Klasse *BuildingPart* eine Spezialisierung der abstrakten Klasse *_AbstractBuilding* (Abb. 2). Daran zeigt sich, dass die beiden Klassen Gemeinsamkeiten besitzen, die sie aus der Basisklasse erben. Sie unterscheiden sich lediglich durch die Aggregationsmöglichkeit des *BuildingPart*. Das bedeutet, dass ein aus mehreren Teilen bestehendes Gebäude durch ein *Building* (repräsentiert den Hauptteil) sowie einem oder mehreren *BuildingParts* (repräsentieren untergeordnete bauliche Elemente) modelliert werden soll. Die sich daraus ergebende Problematik, wie denn die unterschiedlichen Teile (so vorhanden) voneinander abgegrenzt werden sollen, kann mit Hilfe der im Standard genannten Merkmale wie Dachform oder Geschosshöhe gelöst werden. Für die Unterscheidung zwischen

Hauptteil und untergeordneten Gebäudeteilen werden jedoch keine Angaben gemacht. Die Praxis zeigt, dass dies sehr unterschiedlich gehandhabt wird. Ein Ausweg wäre die Modellierung in der Weise, dass das *Building* lediglich als Container ohne eigene Geometrie, aber mit gemeinsamen Attributen, modelliert wird, der dann ein oder mehrere *BuildingParts* mit jeweils eigener Geometrie enthält. Diese Lösung wird von CityGML 2.0 jedoch nicht empfohlen. Gemäß dem ersten Conformance Requirement des Building Moduls soll die Geometrie des Gebäudehauptteils vom aggregierenden *Building*-Element repräsentiert werden (GRÖGER et al. 2012: 78).

Das *Building* Modell von CityGML enthält weiterhin die Features *BuildingInstallation* zur Repräsentation von An- und Aufbauten, sowie eine Reihe von *BoundarySurface*-Typen (Abb. 3), die die einzelnen flächigen Bauteile der Gebäudehülle darstellen. Diese können wiederum die Öffnungen *Door* und *Window* enthalten.

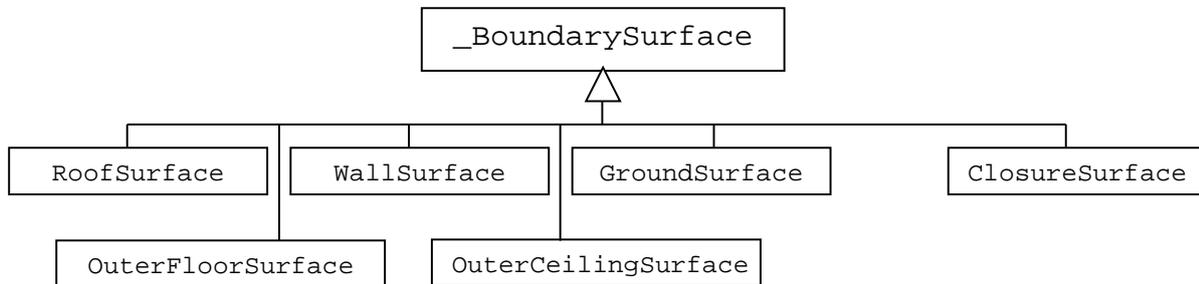


Abb. 3: Ausgewählte *BoundarySurface*-Feature Types in CityGML.

Auf die Features *IntBuildingInstallation*, *BuildingFurniture* und *Room* zur detaillierten Modellierung von Innenräumen wird in diesem Rahmen nicht weiter eingegangen.

Im Vergleich mit den in Abschnitt 4 herausgearbeiteten Gebäudeeigenschaften ergeben sich die in Tab. 2 gestellten Übereinstimmungen zwischen den Modellierungsmöglichkeiten im Rahmen von CityGML und den Anforderungen an Gebäude der Realwelt.

Tab. 2: Eigenschaften von Realwelt-Gebäuden und deren CityGML Repräsentation.

Nr.	reales Gebäude	CityGML Repräsentation
1	selbständig benutzbar	Building::usage
2	bedeckt	BoundarySurface::RoofSurface
3	bauliche Anlage	Building
4	können von Menschen betreten werden	Building::measuredHeight
5	bietet Schutz	Building::function
6	dauerhaft	Building::yearOfConstruction
7	ortsfest	Geometry::srsName, plus Koordinaten
8	Bedeutung	Building::class

Umgekehrt haben die unterschiedlichen *BoundarySurface*-Typen natürlich auch Entsprechungen in der Realwelt, ebenso wie die verschiedenen *BuildingInstallation*-Typen. Da diese Entsprechungen intuitiv sind und um die Aussagen des CityGML-Standard Dokuments nicht zu wiederholen, werden sie hier nicht näher erläutert. Interessant sind jedoch die dort gemachten

Einschränkungen, die sich etwa auf die Ausrichtung der Flächen im Raum beziehen. So wird gefordert, dass die Flächennormale von *GroundSurface* und *OuterCeilingSurface* nach unten, diejenige von *OuterFloorSurfaces* hingegen nach oben, weist. Zusätzliche Einschränkungen gibt es lediglich hinsichtlich der Verwendung in bestimmten Levels of Detail.

7 Regeln für Semantik

In der bisherigen Abhandlung steht das semantische Modell von Gebäuden im Vordergrund, die Beziehungen (sofern existent) zum geometrischen Modell werden nun im Zuge der Regelformulierungen berücksichtigt.

Bis auf die Hinweise im Standard zur Normalenrichtung einiger *BoundarySurface*-Typen sind explizite Aussagen zur Geometrie von semantischen Features nur selten anzutreffen. Dies belegt die Notwendigkeit für Regeln, die die Beziehung der beiden Modelle zum Gegenstand haben, um somit die vom Standard in Kapitel 6.3 geforderte geometrisch-semantische Kohärenz eines Stadtmodells möglichst sicherzustellen (GRÖGER et al. 2012).

Aussage 1 (vgl. Abschnitt 4) besagt, dass ein Gebäude selbständig benutzt werden kann. Geometrische Implikationen lassen sich insofern ableiten, als dass das Gebäude abgeschlossen sein sollte. Die Folgerung daraus wäre eine geschlossene Geometrie der Gebäudehülle, im Idealfall repräsentiert durch ein *Solid*. Die Regel für CityGML dazu lautet:

(A) Ein *Building* muss eine Solid-Geometrie besitzen.

Dies steht zunächst im Widerspruch zu offenen Lagerhallen oder Schuppen (Abb. 4). CityGML bietet in diesem Möglichkeit an, nicht vorhandene Wände mittels *ClosureSurface* zu modellieren und damit eine geschlossene Gebäudehülle mit virtuellen Anteilen herzustellen (Abb. 5).



Abb. 4: Schuppen mit einer offenen Seite.

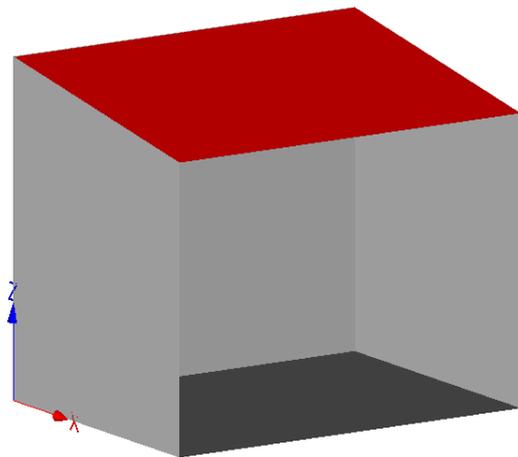


Abb. 5: Modell des Schuppens, vorne eine nicht sichtbare *ClosureSurface* zum Abschluss der Gebäudehülle.

Die zweite Aussage fordert eine Bedeckung des Gebäudes. Es muss somit eine oder mehrere Flächen vorhanden sein, die das Gebäude nach oben hin abschließen. Das deckt sich auch mit der von CityGML geforderten Normalenrichtung von *RoofSurface*-Features.

(B) Ein *Building* muss ein *RoofSurface* besitzen.

Weiterhin kann folgende Definition bei der Validierung des semantischen Modells nützlich sein, womit auch die Sonderfälle wie Zwiebelturmdach bei Kirchen etc. abgedeckt sein sollten:

Eine Wand ist eine vertikale Struktur. Sie wird nach oben von einem Dach und nach unten von einer Grundfläche begrenzt. Ein Gebäude umschließt ein Volumen („offene“ Volumen können mit *ClosureSurface* abgeschlossen werden, vgl. Abb. 5), Teile des Gebäudes (Dachflächen) können über die Wandflächen hinausragen.

Daraus folgt:

(C) Ein Strahl oberhalb des Gebäudes in negativer z-Richtung (0,0,-1) muss als erstes ein *RoofSurface* schneiden (manchmal kann dies auch ein *OuterFloorSurface* sein).

Anmerkung: Analog zu (C) schneidet ein Strahl ausgehend von einem beliebigen Punkt innerhalb des Volumens in Richtung (0,0,-1) als letztes eine Grundrissgeometrie (*GroundSurface*, ggf. *OuterFloorSurface*). Nur Flächen, die so geschnitten werden, sind Grundrissflächen. (Das gilt bei Dachflächen nicht, siehe Zwiebeltürme).

Als zusätzliche Konsequenz ergibt sich, dass auch in LOD1 die Dachflächen zwingend erforderlich sind. Das semantische Modell ergibt sonst keinen Sinn bzw. ist nicht überprüfbar. Durch die Einschränkung von *BoundarySurface*-Features auf die LOD >2 ist das semantische Modell in LOD 1 nur sehr rudimentär anwendbar.

In der dritten Aussage wird eine bauliche Anlage gefordert, um das Gebäude von naturräumlichen Objekten mit ähnlicher Gestalt oder Funktion abzugrenzen. Die Geometrie spielt in diesem Fall keine Rolle.

(D) Es muss ein *Building*-Objekt existieren

Aussage 4 verlangt eine Mindestgröße, so dass Gebäude von Menschen betreten werden können. Diese Mindestgröße lässt sich einerseits aus der Geometrie ablesen (z.B. Grundfläche), andererseits gibt das Attribut *measuredHeight* Auskunft über die Gesamthöhe.

(E) Die Gebäudegeometrie muss eine bestimmte Mindestfläche und –höhe aufweisen.

(F) Das Attribut *measuredHeight* (falls existent) muss einen Mindestwert überschreiten.

Aussage 5 bezieht sich auf die Schutzfunktion des Gebäudes und korreliert mit dem Attribut *function*. Die Geometrie ist dabei unerheblich.

(G) Das Attribut *function* (falls existent) muss einem bestimmten Wertebereich entsprechen

Im sechsten Fall geht es um die Dauerhaftigkeit des Gebäudes. Diese kann anhand des Temporärattributes *yearOfConstruction* (falls existent) belegt werden.

(H) Das Attribut *yearOfConstruction* hat einen Wert, der auf ein Datum in der Vergangenheit verweist.

Aussage 7 verlangt ein ortsfestes Gebäude. Dies wird durch fixe Koordinatenpunkte innerhalb eines Referenzsystems sichergestellt.

(I) Ein *Building* muss eine Geometrie besitzen.

Die achte Aussage hebt auf die Bedeutung eines Gebäudes ab, ohne nähere Angaben dazu zu machen. Letztlich kann sie ihren Niederschlag im Attribut *class* finden.

(J) Das Attribut *class* (falls existent) liegt in einem bestimmten Wertebereich.

Weitere Regeln können direkt aus Vorgaben des Standards abgeleitet werden, wie etwa Richtungen der Normalenvektoren von *BoundarySurfaces* (bei Dach- und Bodenflächen nach oben, Wandflächen zur Seite und Deckenflächen nach unten) sowie Beziehungen zwischen Gebäudeteilen (Ein *BuildingPart* bzw. eine *BuildingInstallation* benötigt ein *Building* als Elternelement). Für Details dazu wird auf (WAGNER et al. 2013) verwiesen.

8 Formalisierung mit RDF und GeoSPARQL

Die in Abschnitt 7 angeführten Regeln sollen im nächsten Schritt Eingang in eine automatisierte Prüfung finden. Die dazu nötige Formalisierung wird im Rahmen eines Resource Description Frameworks (RDF) verwirklicht. Die Bedingungen können dann mit Hilfe der Abfragesprache GeoSPARQL (<http://www.opengeospatial.org/standards/geosparql>) formuliert werden. Es ist eine Erweiterung von SPARQL zur Abfrage von Geodatenätzen, die die Möglichkeit bietet, Semantik *und* Geometrie unter einem Dach gemeinsam zu verarbeiten. Bisher sind geometrische Beispiele nur aus dem 2D-Bereich vorhanden, wobei 3D Geometrie grundsätzlich kein Problem darstellen sollte. Die Untersuchungen sind hier erst im Anfangsstadium, da als Zwischenschritt zunächst die Übertragung der CityGML-Features nach RDF erfolgen muss, z.B. mit einer an CityGML angelehnten und in als RDF definierten Ontologie. Ein Beispiel für die Features *Building* und *BuildingPart* wird hier vorgestellt.

Zunächst ist dazu das *Building/BuildingPart*-Konzept aus CityGML in Form einer RDF-Ontologie zu definieren (Listing1). Dazu wird die XML-Notation RDF/XML verwendet. Im Sinne der Semantik des *BuildingPart* als untergeordnetem Teil des Gesamtgebäudes wird es hier als Subclass des *Building* definiert.

Listing 1: Ontologiebeispiel in RDF/XML

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:cd="http://www.citydoctor.eu/CityGML-rdf/">

  <!-- OWL Header Example -->
  <owl:Ontology rdf:about="http://www.opengis.net/citygml/2.0">
    <dc:title>CityGML Ontology for Building module</dc:title>
    <dc:description>Ontology example for testing semantic validation
of Building features</dc:description>
  </owl:Ontology>

  <!-- OWL Class Definition - Building -->
  <owl:Class rdf:about="http://www.opengis.net/citygml/2.0#Building">

    <rdfs:label>Building</rdfs:label>
    <rdfs:comment>The CityGML Building feature.</rdfs:comment>
  </owl:Class>
  <!-- OWL Subclass Definition - BuildingPart -->
  <owl:Class rdf:about="http://www.opengis.net/citygml/2.0#BuildingPart">

    <!-- BuildingPart is a subclassification of Building-->
    <rdfs:subClassOf
rdf:resource="http://www.opengis.net/citygml/2.0#Building"/>
```

```
    <rdfs:label>BuildingPart </rdfs:label>
    <rdfs:comment>The CityGML BuildingPart feature.</rdfs:comment>
  </owl:Class>
</rdf:RDF>
```

Nun existieren *Building* und *BuildingPart* in einer Form, die mit GeoSPARQL abgefragt werden kann. Die Abfrage, ob eine Solid-Geometrie vorliegt, könnte prototypisch so aussehen wie in Listing 2 dargestellt. Zunächst werden alle *Building*- und *BuildingPart*-Elemente selektiert, die sich berühren (`geo:sfTouches`). Dann wird die kombinierte Geometrie aus allen Teilen gebildet (`geof:union`) und letztlich diese nach der Eigenschaft `geo:isSolid` abgefragt. Letzteres ist eine in GeoSPARQL nicht vorhandene Eigenschaft, deren Implementierung auf anderem Weg sichergestellt werden muss. Das Resultat sind alle Gebäudekomplexe mit Solid-Geometrie. Ist die Ergebnismenge leer, so liegt keine Solid-Geometrie vor und somit die Daten nicht valide im Sinne der Solideigenschaft für Gebäude.

Listing 2: Prototypische GeoSparQL Abfrage

```
PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX geof: <http://www.opengis.net/def/geosparql/function/>
PREFIX cd: <http://www.citydoctor.eu/CityGML-rdf/>
SELECT ?b ?bp
WHERE {
  ?b a ex:Building ;
  geo:hasGeometry ?bgeo .
  ?bp a ex:BuildingPart ;
  geo:hasGeometry ?bpgeo .
  ?bgeo geo:sfTouches ?bpgeo .
  geo:hasGeometry ?compgeo geof:union ?bgeo ?bpgeo ;
  ?compgeo geo:isSolid
}
```

Die Implementierung erfolgt über das JENA Framework (<http://jena.apache.org/>), das alle benötigten Teile als JAVA-API bereitstellt. Insbesondere sind dies die eigentliche RDF-API sowie eine API für SparQL. Daneben gibt es noch die Möglichkeit, einen sogenannten Triple-Store einzurichten, der die Daten persistent in der RDF-typischen Triplestruktur vorhalten kann und sich direkt mit SPARQL abfragen lässt.

9 Fazit und Ausblick

Regeln für die Semantik von Gebäuden im Allgemeinen und für *Building*-Features im Falle von CityGML können mit Hilfe eines Prozessmodells aus einfachen Definitionen abgeleitet werden. Diese Regeln bilden den Ausgangspunkt für die semantische Validierung von 3D-Stadtmodellen. Sie können unter Berücksichtigung des Anwendungsfalls weiter verfeinert werden, wobei darauf zu achten ist, dass die geforderten Einschränkungen eindeutig und präzise ausgedrückt werden können. Im Einzelfall ist eingehende Interpretation der Anforderungen gefragt, die ebenso vom Anwendungsfall beeinflusst werden kann.

Im Unterschied zur rein geometrischen Validierung von 3D-Stadtmodellen, die mit einem klaren Ergebnis abschließt, verbleibt bei der semantischen Prüfung zumeist ein Graubereich, der vom Anwender beurteilt werden muss und sich in der ‚weichen‘ Formulierung der Regeln niederschlägt.

Die Ausstattung der Regeln mit Parametern wie Wertebereichen, Toleranzen etc. erfordert vom Anwender genaue Kenntnis über den Einsatzbereich der zu prüfenden Daten, um eine sinnvolle Prüfkongfiguration zu erstellen. Den Grad der Abweichung legt der Nutzer des Modells fest, da nur er entscheiden kann, inwiefern die Abweichung im jeweiligen Einzelfall wünschenswert ist bzw. noch toleriert werden kann. Dies impliziert, dass absolute Aussagen über die Validität der Semantik nicht zielführend sind, da im Einzelfall immer wieder zulässige Ausnahmen auftreten können. Letztlich ergibt die Prüfung, ob die vorliegende Konfiguration von Geometrie und Semantik plausibel modelliert ist.

Weiterhin sind bestimmte Vorgaben des CityGML Standards schwierig zu interpretieren oder in einer Weise formuliert, dass deren Umsetzung kaum möglich ist. Ein Beispiel dafür ist die Unterteilung von Gebäuden in Hauptteil (*Building*) und baulich untergeordneten Teilen (*BuildingPart*). Abgesehen von der Abgrenzung der Teile untereinander ist es in vielen Fällen kaum möglich zu entscheiden, welcher der übergeordnete Teil sein soll. Eine Gleichwertigkeit aller Teile ist hingegen vom Standard nicht vorgesehen, da er zwingend die Existenz eines *Building*-Features verlangt.

Die Umsetzung der Regeln zur Semantikprüfung in eine formale Sprache wird in einem Prototyp mit RDF und GeoSparQL mit Hilfe des JENA-Frameworks erprobt. Die ersten Ergebnisse sind vielversprechend und werden weiter verfolgt. Damit soll eine Möglichkeit geschaffen werden, diese und weitere Regeln einfach zu implementieren und zu testen.

10 Literaturverzeichnis

- BAUMINISTERKONFERENZ, 2002: Musterbauordnung. www.is-argebau.de/lbo/VTMB100.pdf
- COORS, V. & WAGNER, D., 2015: CityGML Quality Interoperability Experiment des OGC. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 24, T. Kersten (Hrsg.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 16.-18. März 2015 an der Universität zu Köln, Tagungsband auf CD-ROM, 288-295.
- GRÖGER, G., KOLBE, T.H., NAGEL, C. & HÄFELE, K.-H., 2012: OGC City Geography Markup Language (CityGML) Encoding Standard V2.0. OGC Doc 12-019.
- STADLER, A. & KOLBE, T.H., 2007: Spatio-semantic Coherence in the Integration of 3D City Models. Proceedings of the 5th International Symposium on Spatial Data Quality. Enschede.
- WAGNER, D., BOGDAHN, J., WEWETZER, M., ALAM, M.N., PRIES, M. & COORS, V., 2013: Geometric-semantical consistency validation of CityGML models. Progress and New Trends in 3D Geoinformation Sciences, J. Pouliot, S. Daniel, F. Hubert, & A. Zamyadi (Hrsg.), Springer Berlin Heidelberg, 171-192.

Extending Semantic 3D City Models by Supply and Disposal Networks for Analysing the Urban Supply Situation

TATJANA KUTZNER¹ & THOMAS H. KOLBE¹

Abstract: For applications of 3D city models such as simulating the impact of failures in supply networks, information on the city objects as well as on the supply networks are required. The CityGML extension UtilityNetworkADE allows for modelling different types of networks, the focus currently being on the representation of topographical and topological aspects. However, for representing supply and disposal tasks, also functional aspects are of relevance, in particular, when no detailed modelling of supply networks is available, but the impact of a failure on a certain region is to be analysed nevertheless. This paper presents three methods complementing the CityGML UtilityNetworkADE by functional aspects regarding the modelling of supply areas, the characterisation of city objects and network features according to functional roles and the representation of the potential and current supply of commodities to city objects.

1 Introduction

Semantic 3D city models represent city objects such as buildings, bridges, tunnels, roads and vegetation. These city objects mainly constitute the visible part of a city, i.e. that part striking ones eye immediately when looking around a city. However, cities also exhibit a large number of city objects which are not apparent at first sight, but which are crucial to the functioning of the city as a system. These (often) hidden city objects – due to being below ground – contribute to the infrastructure of a city in the form of networks for water, electricity, sewage, telecommunication, and other public utilities.

In general, a network consists of a set of nodes which are connected by links (in graph theory usually referred to as vertices and edges). Networks can be classified according to various criteria. NEWMAN 2003, for instance, differentiates between social networks (i.e. networks dealing with social interactions between people), information networks (i.e. networks linking together knowledge), technological networks (i.e. networks distributing commodities or resources) and biological networks (i.e. networks representing biological systems).

In the context of our work, the focus is on technological networks, in particular on utility networks. Utility networks can also be considered as spatial networks, “i.e. networks whose nodes occupy a precise position in two or three-dimensional Euclidean space, and whose edges are real physical connections” (BOCCALETTI et al. 2006).

One well-known standard for representing 3D city models is the international OGC standard CityGML (GRÖGER et al. 2012). By means of so-called Application Domain Extensions (ADEs) the core model of CityGML can be extended systematically by application-specific attributes and object types. Specifically for applications dealing with supply and disposal networks, the *CityGML UtilityNetworkADE* (BECKER et al. 2011; BECKER et al. 2012) provides concepts which

¹ Technische Universität München, Lehrstuhl für Geoinformatik, Arcisstraße 21, D-80333 München, E-Mail: [kutzner, thomas.kolbe]@tum.de

allow for modelling different types of networks, such as electricity, freshwater, wastewater, gas or telecommunication networks.

So far, the CityGML UtilityNetworkADE focuses mainly on *topographical and structural* aspects. They allow for representing the position, the shape and the extent of networks including all individual network features such as switches, valves, pumps and pipes and for deriving spatial and computer-graphic-related answers therefrom.

However, when modelling supply and disposal networks, often *functional* aspects are of relevance as well. They are required for representing supply and disposal tasks, in particular when the available data do not reflect a detailed modelling of supply lines, but the impact of a network failure on a certain region and on the concrete city objects located in that region are to be analysed nevertheless. To be able to perform such analyses, it needs to be determined whether these city objects, e.g. buildings including their inhabitants, are supplied with the commodities in question. Furthermore, these analyses require information on the potential and actual supply of commodities to city objects and also the concrete flows and flow quantities need to be taken into account.

In the CityGML UtilityNetworkADE functional aspects can currently be represented through the topological connection of the network features which allows for simulating the propagation of failures across multi-network structures. This modelling, however, is not covering everything we need; the impact of failures on city objects and the population, for instance, cannot be simulated up to now. This paper presents three methods for how to complement the CityGML UtilityNetworkADE with these aspects. This includes the modelling of supply areas, a further characterisation of network features and city objects by the roles ‘source’ and ‘sink’ as well as complementing city objects with information on the supply of commodities to them. The UML diagrams presented in this paper are based on the ISO-compliant CityGML UML model (CITYGMLUMLMODEL 2016) which currently also serves as starting point for the further development of CityGML 3.0 (LÖWNER et al. 2014).

Section 2 provides a brief analysis of the CityGML UtilityNetworkADE and related network models regarding characteristics relevant to network modelling. Section 3 presents a refined definition of the concept of networks which includes aspects relevant to functional modelling as well. Section 4 introduces the new approaches for functional modelling to the CityGML UtilityNetworkADE and section 5 concludes this paper.

2 A brief review of related network models

Besides the CityGML UtilityNetworkADE, several other data models and formats for representing utility networks exist. The most relevant ones in the geospatial domain are the INSPIRE Utility Networks model (JRC 2013a) which is based on the INSPIRE Generic Network Model (JRC 2013b), the ISO standard Industry Foundation Classes (IFC) (ISO 16739:2013) which is predominantly used in Building Information Modeling and the ESRI Geometric Network model (ESRI 2016) based on which also distinct data models for gas, water and electricity for use with the ArcGIS software exist. For a detailed discussion of these data models please refer to BECKER et al. (2011) and BECKER et al. (2012).

Another ISO standard which allows for representing utility networks is SEDRIS (Synthetic Environment Data Representation and Interchange Specification) (SEDRIS 2016). 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.

Under the umbrella of the OGC, PipelineML, a GML-based data interchange standard for the exchange of pipeline data focusing on the oil and gas industry, is currently under development (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 exist in scientific literature. HALFAWY (2010), for instance, presents data models for water and wastewater networks, taking hereby also into account life-cycle aspects of network components such as maintenance operations or performance assessment.

The CityGML UtilityNetworkADE proposed by BECKER et al. (2011) and BECKER et al. (2012) 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

Tab. 1: Existence of characteristics relevant to network modelling in various data models

	INSPIRE Utility Networks	IFC	ArcGIS Utility Networks	SEDRIS	Pipeline ML	CityGML UtilityNetwork ADE
Representation of heterogeneous networks	+	•	•	+	•	++
Dual representation	+	++	+	++	–	++
Topographic/graphic aspects	++	++	++	++	++	++
3D geometries	–	++	–	+	–	+
Functional aspects	–	–	–	–	–	•
Hierarchical modelling						
• networks/ subnetworks	++	–	–	++	–	++
• components/ subcomponents	++	++	•	–	•	++
Interdependencies between						
• network features and city objects	–	•	–	•	–	++
• network features of different network types	–	++	–	–	–	++
– = no support, • = basic support, + = sophisticated support, ++ = comprehensive support						

(including 3D) as well as of functional aspects. Furthermore, the data model should 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 as well as between network features of different types of networks.

For the data models from INSPIRE, IFC and ArcGIS a detailed analysis regarding these characteristics is provided in BECKER et al. (2011) and BECKER et al. (2012). In the following, an overview based on this detailed analysis is provided in Table 1 which also takes into account the data models from SEDRIS and PipelineML and compares them to the CityGML UtilityNetworkADE.

The overview in table 1 and the detailed analysis show that the CityGML UtilityNetworkADE meets best the requirements for modelling utility networks regarding the characteristics in question. SEDRIS exhibits a similar good support; disadvantages, however, are the representational ambiguity of the format at runtime and the limited software support. Furthermore, CityGML can be extended more easily by the required functional aspects. Third-best support is offered by the INSPIRE Utility Networks data model, however, only 2D geometries are supported and its scope is limited to the EU. A disadvantage of IFC is its limitation to the building level. ArcGIS, in turn, is a proprietary software product, furthermore, no 3D solids are supported and its adaptation to other types of utility networks is difficult. PipelineML is currently still under development, the evaluation provided in the table is, thus, to be considered preliminary only.

3 The CityGML UtilityNetworkADE network concept

Central concept of the CityGML UtilityNetworkADE is the *network*, which can be represented in two ways: topographically as an aggregation of network features, i.e. of the individual components the network is constructed from, and topologically by means of a network graph, which is composed of the feature graphs of the individual network features. Regarding the topographical representation, the UtilityNetworkADE specifies in addition that a network can be decomposed into subnetworks and that it can be provided with information on what kind of commodity is transported by the network.

In reality, however, networks with more varied characteristics can occur. For this reason, we propose a more precise definition of the concept of a network together with a refined Core UML model of the CityGML UtilityNetworkADE. These refinements are also of importance to the further functional aspects to be introduced to the CityGML UtilityNetworkADE in section 4.

In its most general form, a network is simply a set of interconnected network features. From a semantic point of view, however, several other characteristics of what exactly constitutes a network need to be considered:

1. A network can represent a *subnetwork* of a more extensive network. Using the example of power supply, a transmission network operated by a certain transmission system operator represents one network. In the same way, a transmission network consisting of several interconnected individual transmission networks represents one network as well, the individual transmission networks being subnetworks of the interconnected network. An example are wide-area synchronous grids such as the European electricity transmission

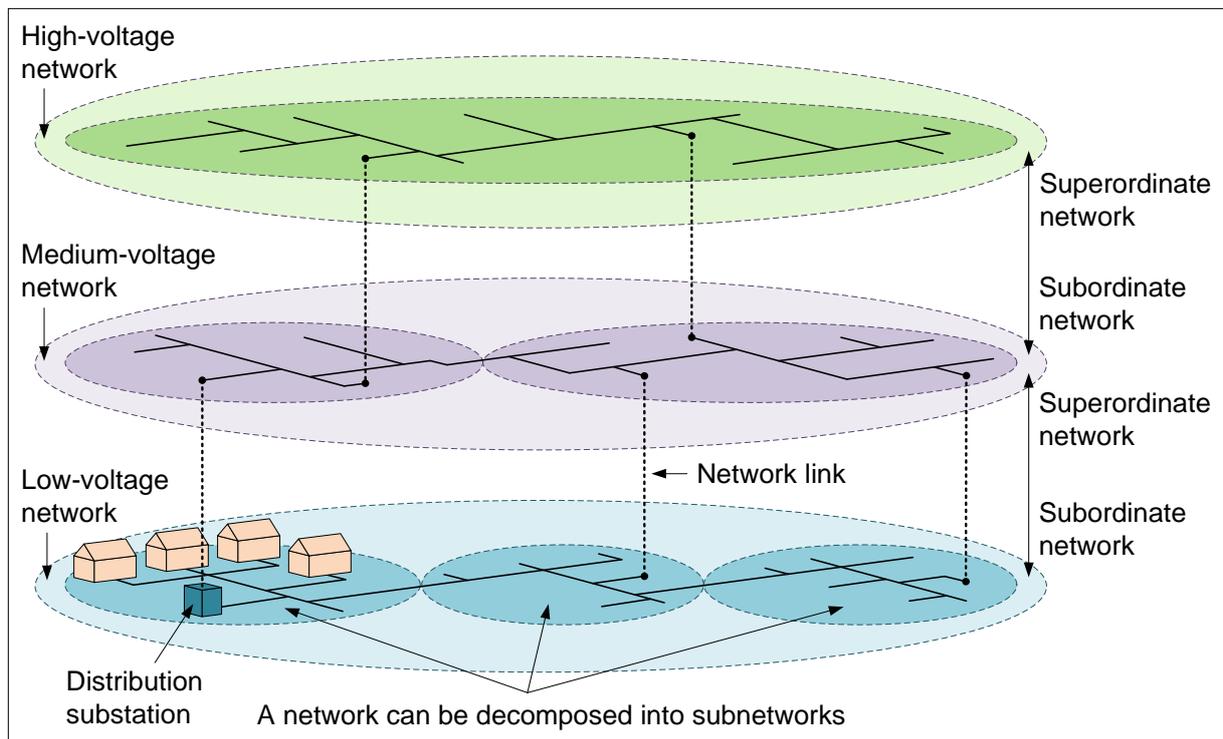


Fig. 1: Decomposition and hierarchical structuring of networks in the context of power supply

network (www.entsoe.eu) which interconnects the transmission networks of 42 electricity transmission system operators. Similarly, a distribution network operated by a certain distribution system operator represents one network. This distribution network could also be split up into several subnetworks, each subnetwork distributing power from a certain distribution substation to the connected end users (cf. Figure 1).

2. A network can represent a *subordinate network* to a *superordinate network* and vice versa. Using the example of power supply again, power is transmitted and distributed based on different voltage levels. High-voltage networks are used to transmit power across large distances, e.g. from power stations to supply regions. For distribution within a supply region to the end users, high voltage is first transformed to medium voltage and distributed via medium-voltage networks. Afterwards, medium voltage is transformed to low voltage by means of distribution substations and distributed via low-voltage networks to the end users (cf. Figure 1). From the point of view of the low-voltage network, the low-voltage network represents a subordinate network and the medium-voltage network its superordinate network. Similarly, the medium-voltage network is subordinate to the high-voltage network.
3. Although a network might initially have been constructed for transporting a certain type of commodity, it can in practice also be used for transporting other types of commodity, either simultaneously, alternately or replacing its intended usage. Power lines, for instance, cannot only distribute electricity, but they can also be used for transferring data simultaneously, which is referred to as powerline communication.

Based on these characteristics, the Core UML model of the CityGML UtilityNetworkADE is refined as follows (cf. Figure 2):

- Currently, the UML class *Network* supports the definition of subnetworks only in the sense of subordinate and superordinate networks. Since the concepts of subordinate and superordinate network are different from the concept of subnetwork, a new bidirectional association with the role names *subOrdinateNetwork* and *superOrdinateNetwork* is added to *Network* to be able to explicitly distinguish subordinate and superordinate networks from subnetworks which partition a larger network on the same network level.
- Following the attribution of feature classes in the CityGML specification, the attributes *class*, *function* and *usage* are added to *Network* as well. The attribute *class* allows for denoting the type of a network; possible values include *high-voltage network*, *medium-voltage network* and *low-voltage network* for power supply networks, as well as *high-pressure network*, *medium-pressure network* and *low-pressure network* for gas networks. The attribute *function* can be used to denote the intended usage of a network, whereas the attribute *usage* should be used to denote the actual usage of a network; possible values for these two attributes are *supply*, *disposal* and *communication*. The actual values for these attributes are, in compliance with the CityGML specification, to be provided as external code lists. More attributes as regards the specific types of commodity transported by a network are not required here; this information is already provided by the existing association *transportedMedium* between *Network* and the UML class *AbstractCommodityType* (cf. BECKER et al. 2012; CITYGMLWIKI 2016a). Since according to the third characteristic more than one commodity can be transported by the same network, only the multiplicity of the association needs to be changed from 0..1 to 0..* (not shown in Figure 2).
- Another refinement applied to *Network* is to extend it from the UML class *AbstractCityObject* instead of from *AbstractFeature*, since not only the individual network components (defined as UML subclasses of the UML class *AbstractNetworkFeature*; cf. BECKER et al. 2012 and CITYGMLWIKI 2016b) represent city objects, but also the networks themselves.
- Currently, a composition between *Network* and *NetworkLink* exists allowing for explicitly expressing links between networks of different network types, such as between a medium-voltage and a low-voltage network (cf. Figure 1). Since *NetworkLink* represents a topological concept, *Network*, however, a topographical concept, the composition should rather link the network topologies of the corresponding networks and, thus, be modelled between *NetworkGraph* and *NetworkLink*.

4 Integration of functional aspects into the CityGML UtilityNetworkADE

4.1 Representation of supply areas

The CityGML UtilityNetworkADE allows for representing in detail the structure of a network together with all the discrete network components the network is composed of and, thus, for following the route of a commodity from source to sink and for determining the whole area the commodity is supplied to by the network, i.e. the supply area. When a failure in one of the network components occurs, it is – due to the detailed modelling – easy to analyse which parts of the network itself and, in particular, which city objects within the supply area are affected by this

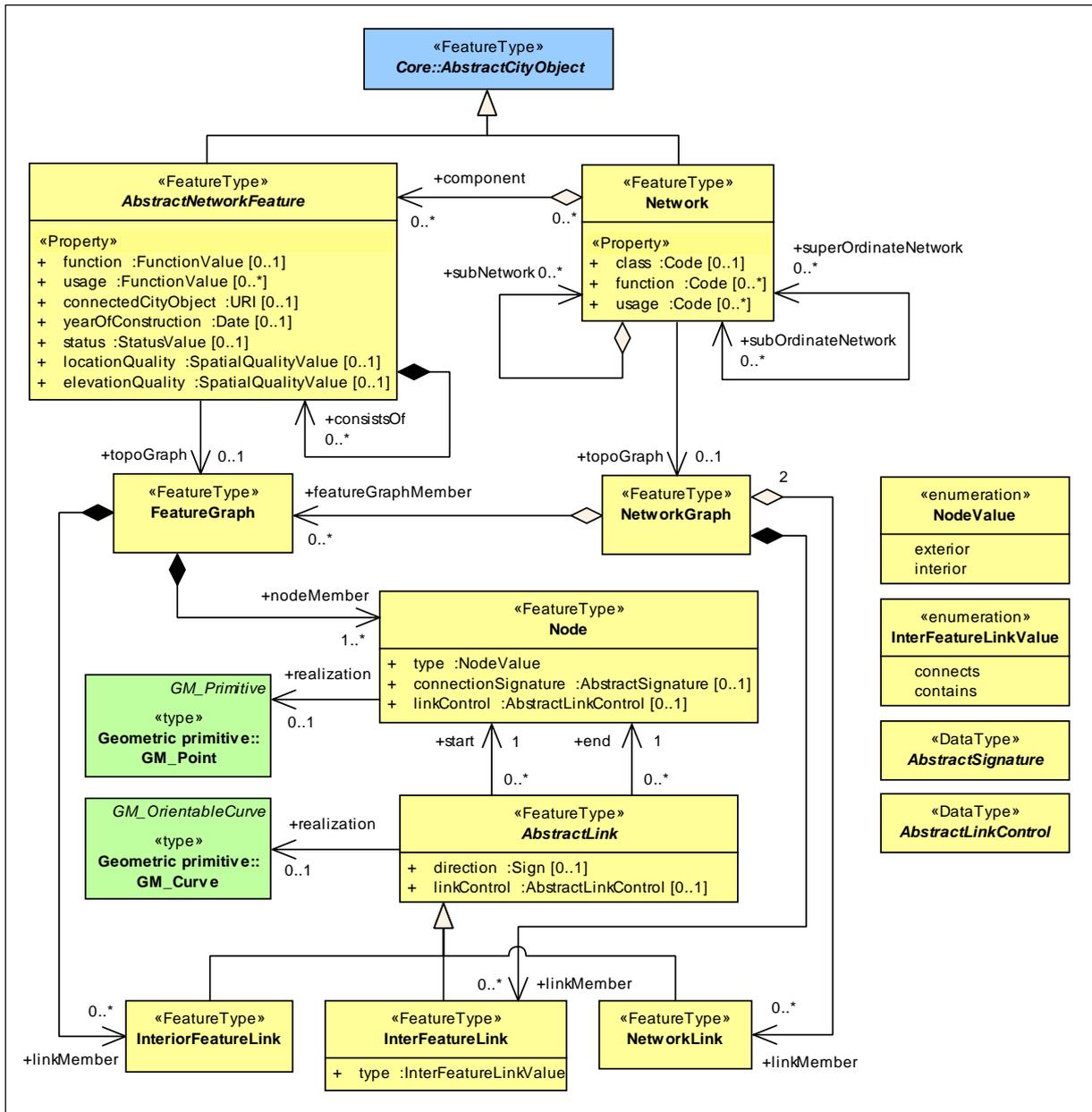


Fig. 2: CityGML UtilityNetworkADE – Refined Core UML model

failure. For certain reasons (e.g. data privacy issues), however, no information reflecting the detailed modelling of a network supplying a certain area with a commodity might be available; in that case it should still be possible to analyse which impact a failure of a source has on the corresponding supply area and on the concrete city objects located in that area.

To meet this demand, the concept of the supply area is to be introduced into the CityGML UtilityNetworkADE. We define a supply area as that geographic region a specific commodity is supplied to by a network. One could also say the supply area substitutes the network supplying that area, which, in turn, implies that the spatial extent of the supply area conforms to the spatial extent of the network it substitutes.

Furthermore, each supply area needs to be related to one or more sources supplying the commodity to that area. In the power supply context, for instance, a distribution substation transforming medium voltage to low voltage for distribution to city objects within a certain area could be regarded as such a source (cf. Figure 3). Above that, corresponding sinks need to be defined, in particular when the exact number of city objects a commodity is actually supplied to is of relevance to a certain application; otherwise it is assumed that the commodity is supplied to all city objects located within that region.

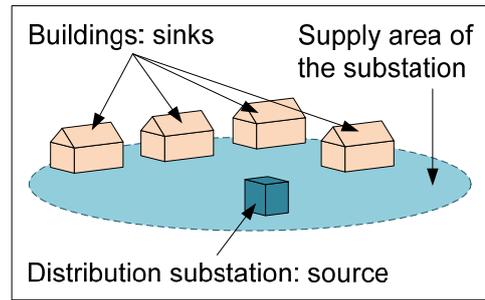


Fig. 3: The relationship between supply area, source and sink

In our new version of the CityGML UtilityNetworkADE the supply area is represented by the UML class *SupplyArea* (cf. Figure 4). By defining *SupplyArea* as a subclass of the UML class *CityObjectGroup*, each supply area can be provided with a geometry defining the spatial extent of the supply area (in particular the ISO 19107 geometry types *GM_Polygon* and *GM_MultiPolygon* are of relevance here). *CityObjectGroup* is associated with the UML class *AbstractCityObject* through the intermediate UML class *Role* which provides the attribute *role*. This association can be used to relate a supply area to the city objects located in that area. An OCL constraint is introduced which specifies that, when making use of the association, the attribute *role* needs to be set to the value *supplies*. The city objects located within a supply area can then be determined in two ways: Either explicitly through the above mentioned association between *SupplyArea* and *AbstractCityObject*, or implicitly by intersecting the geometry of the supply area with the geometries of the city objects.

Since the supply area is strongly related to the corresponding network supplying that area, a 1:1 association is defined between the UML classes *SupplyArea* and *Network*. The association is

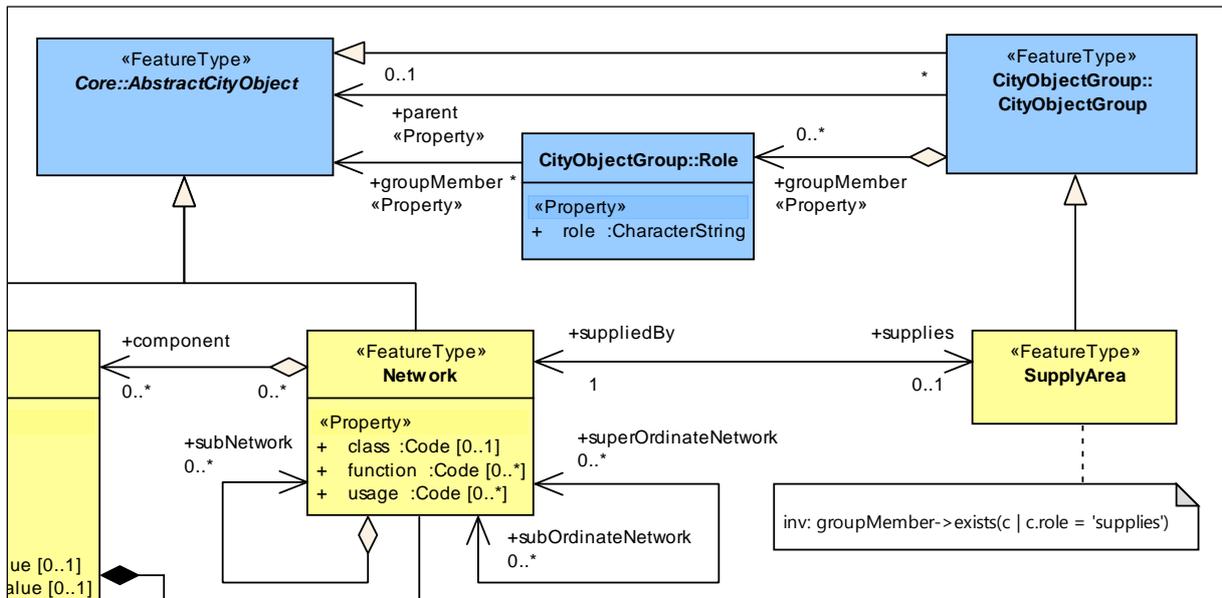


Fig. 4: Representation of supply areas in the CityGML UtilityNetworkADE. The blue classes are CityGML classes and the yellow ones are classes of the UtilityNetworkADE.

bidirectional to allow for navigating from the supply area to the corresponding network and vice versa. Thus, also the type of commodity supplied to an area can be specified by making use of the association *transportedMedium* between the UML classes *Network* and *AbstractCommodityType* (cf. BECKER et al. 2012; CITYGMLWIKI 2016b). The association between *SupplyArea* and *Network*, furthermore, allows for making use of the network characterisations provided in section 3 – subnetwork, subordinate and superordinate network. A network can now, for instance, not only be decomposed into subnetworks, but also a supply area into corresponding ‘sub’-supply areas.

The UML modelling for relating a supply area with sources and sinks will be presented as part of the next section.

4.2 Characterisation of infrastructure objects according to their functional role

Another important functional aspect is the possibility of representing the flow of commodities in a network by denoting network features and city objects with one of the roles source, sink, distribution, or storage. In particular in the context of supply areas this information is useful for relating a supply area with corresponding sources and sinks (cf. section 4.1).

The modelling of this aspect is already partially available in the existing UtilityNetworkADE (cf. Figure 5). The UML class *AbstractNetworkFeature* provides the attribute *function* of the type *FunctionValue* which contains, amongst others, the required roles (the values *feeding* and *draining* are representing the roles source and sink, respectively). This modelling states the intended role of a network feature; the role, however, might change when a network feature is used differently from its intended usage. Thus, the attribute *usage* is added to allow also for expressing the actual role of a network feature.

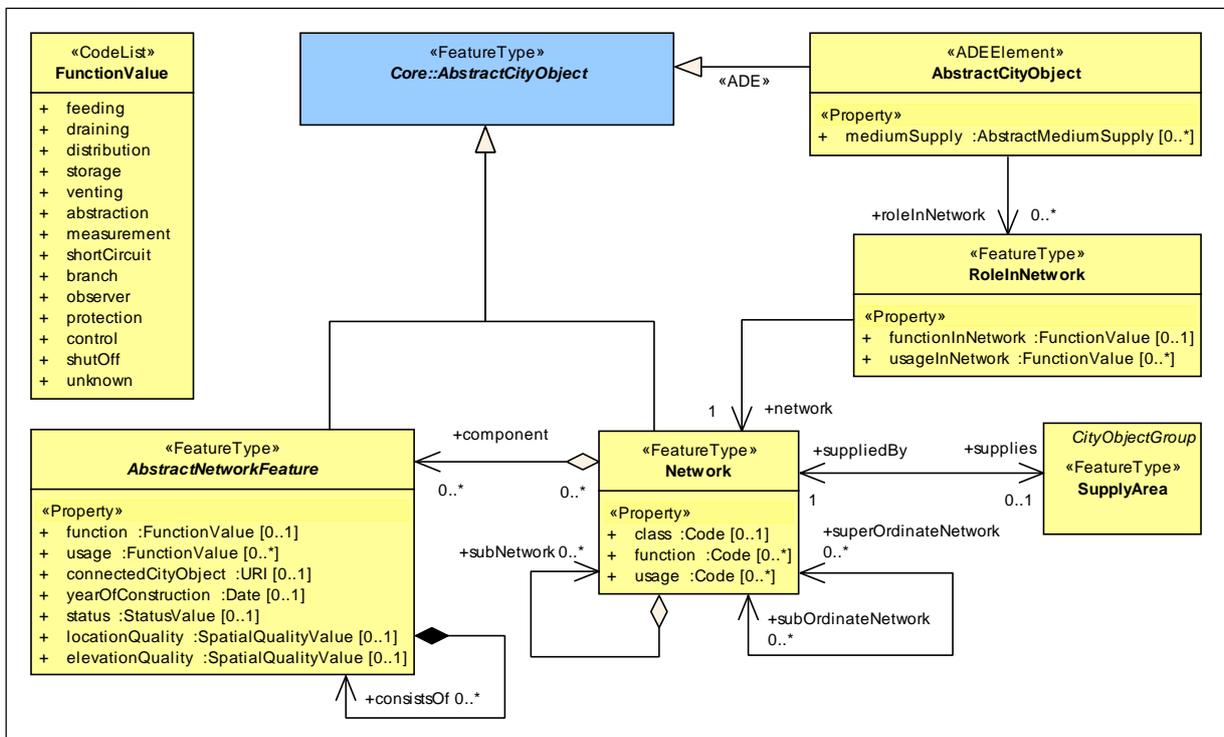


Fig. 5: Representation of functional roles in the CityGML UtilityNetworkADE

In addition, to be able to add these roles to city objects as well, a relationship between *AbstractCityObject* and *Network* is defined by means of the intermediate UML class *RoleInNetwork*. The class defines the attributes *functionInNetwork* and *usageInNetwork* which allow for denoting the intended and actual roles of a city object within a specific network. Each city object can be provided simultaneously with role information specific to different networks, i.e. a building can, for instance, at the same time be denoted as sink in a power network and as source in a wastewater network.

The relationship between *AbstractCityObject* and *Network* fulfils at the same time another important requirement; together with the relationship between *Network* and *SupplyArea* introduced in section 4.1 it yields a transitive relationship between *AbstractCityObject* and *SupplyArea* which allows for relating supply areas with the corresponding roles of city objects located in these areas. Similarly, a transitive relationship between *AbstractNetworkFeature* and *SupplyArea* via *Network* exists allowing for relating supply areas with the corresponding roles of network features located in these areas.

4.3 Representation of the potential and current supply of commodities to city objects

In particular in the context of simulations, and also in general when no information at all is available on networks, it can be useful to provide the city objects themselves with information on the potential and current supply of commodities to them. We introduce the following concepts to express this information:

- *Suppliability*: The suppliability of a city object defines the *potential supply* of a commodity to a city object, the reliability of the supply and whether storage is available allowing for autonomous supply with a commodity in case the supply via a network is interrupted, for instance due to a natural disaster.
- *Suppliedness*: The suppliedness of a city object defines the *actual supply* of a commodity to a city object at a specific point in time as well as the actual reliability of the supply and the actual state of the storage at that point in time.

In the new CityGML UtilityNetworkADE these concepts are represented by the UML class *AbstractMediumSupply* (cf. Figure 6). The class is specialised into the subclasses *ElectricalMediumSupply*, *GaseousMediumSupply*, *LiquidMediumSupply*, *OpticalMediumSupply* and *SolidMediumSupply* to be able to classify the suppliable commodities according to their physical condition. This classification conforms to the already existing classification of commodities transportable by networks used in the association *transportedMedium* between the UML classes *Network* and *AbstractCommodityType* (cf. BECKER et al. 2012; CITYGMLWIKI 2016a). The code lists indicating values for the specific commodity types can also be reused from there.

AbstractMediumSupply provides three attributes: *potentialSupply* and *currentSupply* for expressing the suppliability and suppliedness, respectively, as well as *storage* for providing information on possibly available storage. The attributes *potentialSupply* and *currentSupply* are of the type *Supply*, which, in turn, allows for explicitly stating the flow rate of a commodity (*flowRate*) as well as the supply reliability of that commodity (*status*), i.e. whether the supply is uninterrupted (*inUse*), intermittent (*tempOutOfService*), or unsupplied (*outOfService*), or

whether the supply line might even be *destroyed*. The attribute *storage* of the type *Storage* allows for providing detailed information on the type of a storage (*type*), on its potential and actual capacity (*maxCapacity*, *fillLevel*) as well as on the rate the commodity is flowing in and out of the storage (*inflowRate*, *outflowRate*). The enumeration *StatusValue* and the code list *StorageDeviceValue* are reused from the existing CityGML UtilityNetworkADE.

In order to be able to add all this information to a city object, the UML class *AbstractCityObject* is extended by the attribute *mediumSupply*, the multiplicity *0..** allowing for adding supply information of different commodities to each city object.

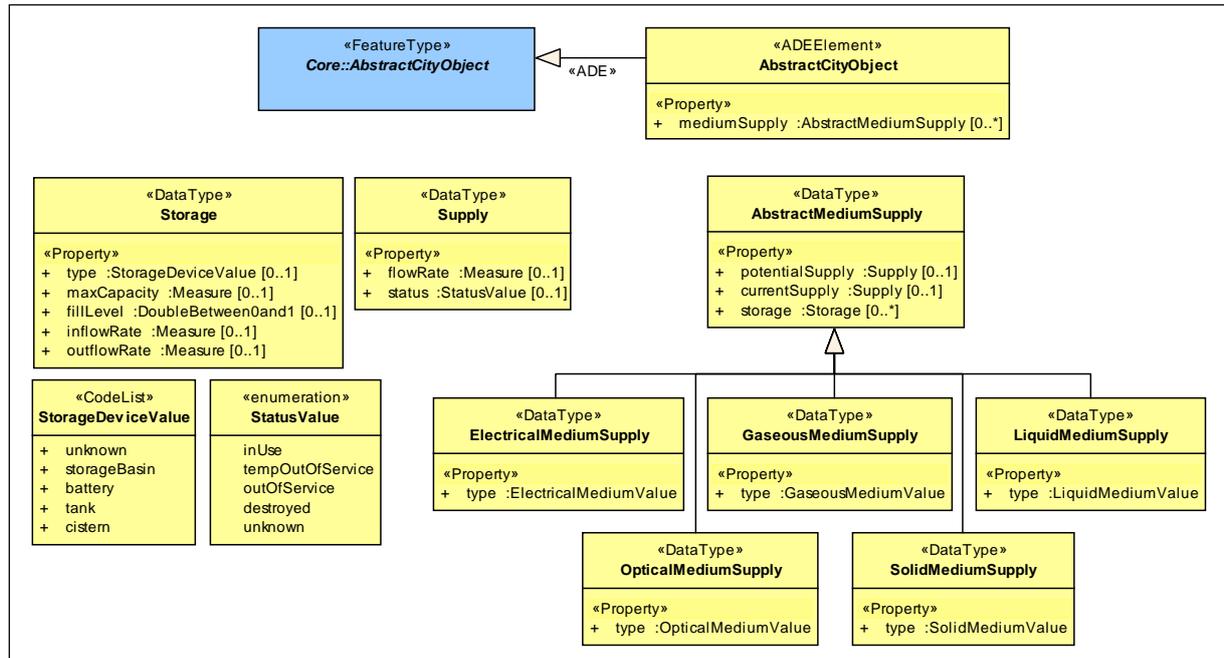


Fig. 6: Representation of the suppliability and suppliedness of city objects in the CityGML UtilityNetworkADE

5 Conclusion

The CityGML UtilityNetworkADE represents a suitable data model for modelling heterogeneous networks in the context of 3D city models. This paper introduces new concepts to the UtilityNetworkADE which add different functional aspects regarding the modelling of supply areas, the characterisation of city objects and network features according to functional roles and the representation of the potential and current supply of commodities to city objects. These functional aspects allow for representing supply and disposal tasks in cases when no detailed modelling of supply networks is available.

The CityGML UtilityNetworkADE as well as the newly introduced aspects were modelled using the software Enterprise Architect. By means of the software ShapeChange, corresponding XML schema files were successfully derived from the UML model and used in FME for testing the transformation of network information from Shape and dxf files to CityGML.

It is planned to include the CityGML UtilityNetworkADE together with the new functional aspects into the next version of CityGML (version 3.0). Another concept to be included in CityGML 3.0 is the so-called ‘Dynamizer’ concept which allows for modelling dynamic properties (CHATURVEDI & KOLBE 2015). This concept could then be used for making dynamic the functional properties introduced in this paper, i.e. the actual supply of commodities to city objects (including current flow rates, voltages, etc.) and the actual roles of city objects and network features.

6 References

- 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.
- 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.
- CITYGMLUMLMODEL, 2016: ISO-compliant definition of the CityGML 2.0 UML model using Enterprise Architect. URL: <ftp://ftp.lrz.de/transfer/DidVfCGML/> (last access on 26/04/2016).
- CITYGMLWIKI, 2016a: NetworkComponents UML model. <http://www.citygmlwiki.org/> (last access on 26/04/2016).
- CITYGMLWIKI, 2016b: NetworkProperties UML model. <http://www.citygmlwiki.org/> (last access 26/04/2016).
- ESRI, 2016: What are geometric networks? <http://desktop.arcgis.com/en/arcmap/latest/manage-data/geometric-networks/what-are-geometric-networks-.htm> (visited on 26/04/2016).
- GRÖGER, G., KOLBE, T.H., NAGEL, C. & HÄFELE, K.-H., 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.
- 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.

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., Volume **23**, Seyfert, E., Gülch, E., Heipke, C., Schiewe, J. & Sester, M. (eds.), 34. Wissenschaftlich-Technische Jahrestagung der DGPF, March 26-28 in Hamburg.

NEWMAN, M.E.J., 2003: The structure and function of complex networks. *SIAM Review* **45** (2), 167-256.

OGC, 2016: PipelineML. URL: <http://www.pipelineml.org/> (visited on 26/04/2016)

SEDRIS, 2016: SEDRIS standards. URL: <http://standards.sedris.org/> (visited on 26/04/2016)

Innovative, quantitative 3D-Planungswerkzeuge für sichere Städte der Zukunft

ANDREAS STEIN¹ & MATTHIAS BACHMANN²

Zusammenfassung: Eine voranschreitende Urbanisierung, eine steigende Komplexität kritischer Infrastruktur und die Entstehung neuer Bedrohungen, beispielsweise durch Naturkatastrophen oder Terrorismus, erhöhen die Verwundbarkeit städtischer Gebiete (UNITED NATIONS 2011; LIN et al. 2007). Risiko- und Resilienzbewertungen im Zuge von Planungs-vorgängen werden aber häufig noch qualitativ auf der Basis von Fragebögen oder in kostenintensiven Einzelberatungen durch spezialisierte Experten durchgeführt. Insbesondere konzentriert sich die Analyse dabei auf einzelne besonders gefährdete Objekte ohne ausreichende Berücksichtigung des urbanen Gebietes. Im Rahmen verschiedener Forschungsprojekte und mit Stadtplanern der öffentlichen Hand und der Privatwirtschaft wird seit 2011 die VITRUV-Software als ein ziviles Risikoanalysewerkzeug entwickelt. Die sowohl auf empirischen Daten als auch auf physikalischen Modellen basierende Risikoanalyse hilft Stadtplanern und ihren Partnern wie z.B. Architekturbüros eine wissenschaftlich fundierte Sicherheitsbewertung mit vertretbarem Aufwand durchzuführen. Jedoch gehören solche Sicherheitsbewertungen nicht unbedingt zum Tagesgeschäft von Architekten oder Ingenieurbüros und wären daher sinnvoller bei Bedarf als externer Service anzubieten – insbesondere da eine wachsende Zahl von urbanen Geoinformationen und 3D-Stadtmodellen heutzutage bereits über Online-Portale verfügbar sind.

Daher arbeiten das Fraunhofer-Institut für Graphische Datenverarbeitung (IGD) und das Fraunhofer-Institut für Kurzzeitdynamik, Ernst-Mach-Institut (EMI) in dem Projekt UrbanRiskPortal an dem gemeinsamen Ziel, den Zugang zu einer urbanen Risikobewertung für Stadtplaner und Architekten zu erleichtern. Für diesen Zweck wird die VITRUV-Software in einer vom Fraunhofer IGD entwickelten cloudbasierten service-orientierten Plattform zur Verfügung gestellt. Die Umsetzung in einer Cloud-Lösung ermöglicht bei Bedarf eine erhöhte Rechenleistung durch einen serverseitigen Rechencluster und genügt außerdem hohen Anforderungen an Verfügbarkeit und Ausfallsicherheit. Ein wichtiger Bestandteil der Umsetzung ist zudem die Gewährleistung einer hohen Sicherheit für die sicherheitsrelevanten Algorithmen, Daten und Ergebnisse. Eine grafische Benutzeroberfläche auf der Basis von 3D Stadtmodellen mit einer ansprechenden Visualisierung ermöglicht auch Planern ohne Expertenwissen eine intuitive Risikobewertung.

In den folgenden Abschnitten wird das Konzept einer Cloud-Lösung zur urbanen Risikoanalyse im Detail vorgestellt. Dabei ist die Anwendung so konzipiert, dass Expertise im Bereich von 3D-Stadtmodellen erforderlich ist, nicht aber im Bereich der Risikobewertung. In Abschnitt 1 werden die erforderlichen Daten betrachtet, wobei sowohl die Geometrie als auch die Attributierung Ansprüche erfüllen muss, um die Qualität der Analyse zu steigern. Darauf aufbauend wird in den nächsten beiden Abschnitten die eingesetzte Software vorgestellt: In Abschnitt 2 folgt das Risikoanalysewerkzeug VITRUV und in Abschnitt 3 die Software für die Aufbereitung und Verarbeitung der Geodaten. In Abschnitt 4 wird deren Zusammenspiel im UrbanRiskPortal erläutert, bevor das Paper mit einem Fazit abgeschlossen wird.

¹ Fraunhofer-Institut für Graphische Datenverarbeitung IGD, Fraunhoferstraße 5, 64283 Darmstadt, E-Mail: andreas.stein@igd.fraunhofer.de

² Fraunhofer-Institut für Kurzzeitdynamik, Ernst-Mach-Institut (EMI), Am Klingelberg 1, 79588 Efringen-Kirchen, E-Mail: bachmann@emi.fhg.de

1 Das 3D-Stadtmodell in der Cloud

Die Heterogenität und Qualität von Geodaten schwankt trotz zunehmender Standards sehr stark. Während sich CityGML als Datenformat in Deutschland durchzusetzen scheint gibt es weiterhin Städte und Kommunen, die 3D-Stadtmodelle in verschiedenen Formaten fortführen. Die daraus resultierenden Unterschiede in den Daten müssen vor einer weiteren Verarbeitung bereinigt werden. Die Hochschule für Technik in Stuttgart hat für die frühzeitige Erkennung von geometrischen Fehlern den CityDoctor entwickelt. Dabei handelt es sich um Methoden und Metriken zum Qualitätsmanagement von 3D-Stadtmodellen. Dieses System wird im UrbanRiskPortal eingesetzt um kritische geometrische Fehler vor der Analyse zu erkennen. Weiterhin ist das UrbanRiskPortal so konzipiert, dass es mit einer sehr geringen Detailierung der Daten arbeiten kann. So ist es möglich eine aussagekräftige Analyse auf Basis von Level of Detail (LoD) 1 durchzuführen. Weiterhin steigert zwar ein digitales Geländemodell die Qualität des Ergebnisses und wertet die Visualisierung auf, ist aber für die Durchführung nicht erforderlich.

Neben den geometrischen Anforderungen an die Daten ist eine korrekte Attributierung notwendig, um das Risiko einzelner Gebäude korrekt zu berechnen. Aus dem Standard CityGML kommen die drei Attribute class, function und usage. Das Attribut class beschreibt die Klassifikation der Objekte, beispielsweise Straße, Straßenbahn, Bahn, oder Platz. Das Attribut function enthält den Zweck des Objekts, wie Bundesstraße oder Kreisstraße, während das Attribut usage definiert, wie ein Objekt genutzt wird. Für diese Attribute existieren Codelisten, die zum Beispiel die verschiedenen Nutzungsarten mit einem dazugehörigen Code verknüpfen.

Weitere Eigenschaften wie Qualität der Fenster oder des Mauerwerks werden ebenfalls bei der Analyse berücksichtigt. Diese Eigenschaften können als Attribute im CityGML-Format abgelegt werden. Das System gibt bei nicht vorhandener Attributierung die Möglichkeit die Werte manuell auszuwählen oder nimmt einen Standardwert an. Durch eine so korrekt vorgenommene Attributierung und qualitätsgesicherte Geometrien kann eine aussagekräftige Risikobewertung für 3D-Stadtmodelle vorgenommen werden.

Eine alternative Lösung wäre die Entwicklung einer CityGML Application Domain Extension (ADE). In diesem Projekt wurde sich gegen die Entwicklung einer ADE entschieden, da die nötigen Eigenschaften und deren Werte auch als generisches CityGML-Attribut abgelegt werden können, was wiederum die Benutzbarkeit der Anwendung steigert. Durch die steigende Nutzung von 3D-Stadtmodellen in der Anwendungsentwicklung werden immer mehr Stadtmodelle als Open Data zur Verfügung gestellt. Während dadurch eine deutliche Steigerung von verfügbaren Anwendungen erreicht wird, birgt es für die Gesellschaft auch Gefahren. Menschen mit böswilligen Absichten könnten versuchen eine Risikoanalyse auf Basis von Open Data durchzuführen, um dadurch einen besonders verwundbaren Punkt im 3D-Stadtmodell zu finden.

Um das zu verhindern, wird die im Projekt entwickelte Software als Webservice betrieben. Die zugrundeliegenden Algorithmen der VITRUV-Software werden nicht nach außen gegeben, sondern laufen im Backend auf den Servern des Fraunhofer EMI.

2 Das Risikoanalyse-Tool für den Einsatz als Webservice

Die zivile Risikoanalyse-Software integriert Sicherheitsaspekte direkt in den städtischen Planungsprozess im Sinne einer resilienten und nachhaltigen Stadtplanung. Sie unterstützt bei der Identifikation und Beurteilung potenzieller Schwachstellen und bietet zudem gezielte Gegenmaßnahmen inklusive Kostenabschätzung.

VITRUV erfordert keinerlei Expertenwissen und unterstützt Planer und Entscheider auf zwei verschiedenen Ebenen. Zunächst hilft die empirische Risikoanalyse bei der Erkennung potenziell gefährdeter Punkte in der Stadt und bietet so eine schnelle wissenschaftlich fundierte Einschätzung der Gefährdungslage. Zweitens ermöglicht die physikalische Risikoanalyse basierend auf kurzzeitdynamischen Ingenieurmodellen und Strukturversagensbewertungen eine detaillierte Prognose möglicher Schäden. Mit Gegenmaßnahmen wie Gebäudeverstärkung, Sicherheitsglas oder Pollern ist es den Planern dann möglich die Sicherheit ihrer städtischen Gebiete bereits in der Planungsphase deutlich zu erhöhen. Schließlich schätzt VITRUV die Kosten und ermöglicht damit eine erste Kosten-Nutzen-Analyse.

2.1 Risikoanalysen

Eine städtische Umgebung besteht aus verschiedenen Objekten wie Gebäuden, Verkehrswegen, Verkehrsknotenpunkten und öffentlichen Plätzen. In dieser städtischen Umgebung gibt es für terroristische Bedrohungen verschiedene Szenarien wie beispielsweise Explosionsereignisse oder biochemische Kontamination. Dieser Artikel konzentriert sich auf Explosionsereignisse, die in verschiedene Bedrohungsklassen eingeteilt werden können - in der Größe von Rucksäcken über Koffer bis hin zu kleinen und großen Fahrzeugen. Die Frequenz $F(T_i, b_k)$ für ein Ereignis am Ort des Objektes b_k mit der Bedrohungsklasse T_i hat die Einheit Ereignisse / Jahr. Das Risiko R eines städtischen Objektes b_k für ein solches Ereignis der Bedrohungsklasse T_i ist als das Produkt der Ereignisfrequenz F mal der Konsequenz C des Ereignisses definiert:

$$R(T_i, \vec{b}_k) = F(T_i, \vec{b}_k) \cdot C(T_i, \vec{b}_k) \left[\frac{\text{consequences}}{\text{year}} \right] \quad (1)$$

Zur Schätzung der Konsequenzen gibt es zwei verschiedene Ansätze – einmal die empirischen Konsequenzen aufgrund historischer Daten und dann die physikalischen Konsequenzmodelle.

2.1.1 Empirische Risikoanalyse

Der Zweck der empirischen Risikoanalyse ist es besonders bedrohte Punkte (Hotspots) und Punkte hoher Verwundbarkeit (Schwachstellen) zu identifizieren. Die Häufigkeit des Auftretens für eine bestimmte Art von Bedrohung bezüglich einer Gebäudekategorie wird aufgrund empirischer Daten aus vergangenen Ereignissen ermittelt. Historische Daten für die Bedrohung durch intendierte Explosionsereignisse werden aus der hauseigenen Terror-Ereignis-Datenbank (TED) mit 80.000 Ereignissen weltweit entnommen (FISCHER et al. 2013). Teil des Datensatzes sind dabei auch die empirischen Konsequenzen wie die Anzahl betroffener Personen und die Anzahl der Personenschäden.

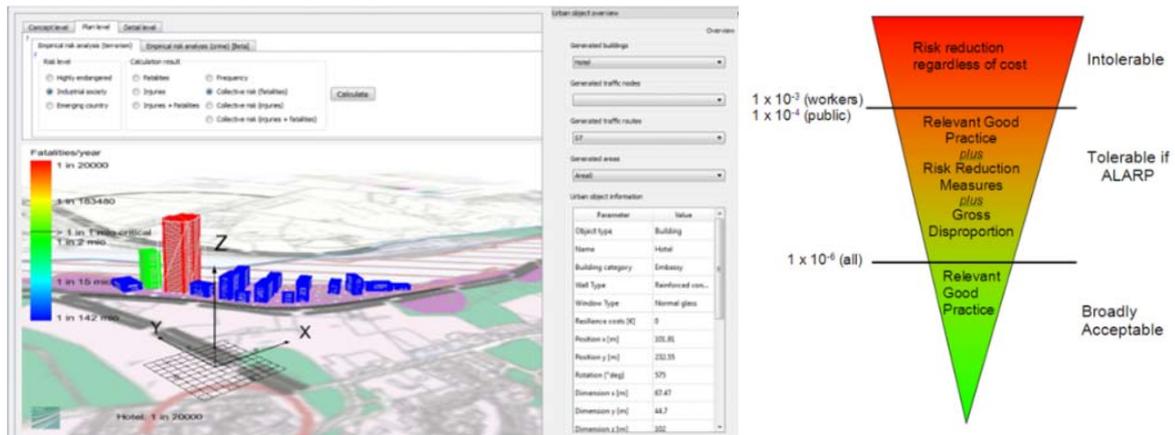


Abb. 1: Empirisches Risiko für Personenschäden in Gebäuden (links), Individuelle Risikolevel definiert durch den "UK Health & Safety Executive Report" von 2001 (rechts)

Abschließend berechnet VITRUV das kollektive Risiko nach (1) und unterstützt den Planer bei der Einschätzung, ob das Risiko kritisch oder akzeptabel ist mit Akzeptanzgrenzen. Der Vergleich zwischen dem kollektiven empirischen Risiko für Personenschäden in Abb. 1 (links) und den Akzeptanzgrenzen der britischen Health and Safety Exekutive (2001) (HSE 2014) in Abb. 1 (rechts), hilft bereits die Gefährdungslage für das betrachtete städtische Umfeld zu bewerten.

So identifiziert die empirische Risikoanalyse die kritischen Punkte im Stadtgebiet. Das ist allerdings nur der Ausgangspunkt für eine detaillierte physikalische Risikoanalyse und anschließende Optimierung des städtischen Gebietes. Die empirischen Konsequenzen betrachten nur einzelne Objekte. Expositionen durch benachbarte Objekte können mit diesem Ansatz nicht berücksichtigt werden.

2.1.2 Detaillierte physikalische Risikoanalyse

Um das Risiko und die Konsequenzen von Bedrohungsereignissen genau beurteilen zu können, ist eine detaillierte Repräsentation des Stadtgebiets nötig. Die empirische Analyse liefert eine sofortige empirische Bestimmung der am meisten gefährdeten Punkte (Hotspots) im Stadtgebiet, die jedoch noch keine Schlüsse über mögliche Konsequenzen zulässt. Die detaillierte Analyse ermöglicht dagegen eine genauere Prognose der möglichen Konsequenzen mit physikalischen Modellen (FISCHER et al. 2016), die zwischen mehreren Minuten und Stunden für die Berechnung benötigt. Die physikalischen Modelle führen somit in der Konsequenz zu besseren Risikovorhersagen. Außerdem werden neben der Nachbarschaft auch Schutz- und Verstärkungsmaßnahmen mit einbezogen. Der allgemeine Ansatz ist für eine große Anzahl von Konsequenzmodellen anwendbar. Zunächst wird dazu ein Raster von möglichen Bedrohungspositionen im Stadtgebiet erstellt. Die einzelnen Positionen werden dabei mit Hilfe einer Wahrscheinlichkeitsdichtefunktion in Abhängigkeit von ihrem Abstand zu verschiedenen Objekten gewichtet. Die Gesamtsumme der Ereignisfrequenzen F über alle Bedrohungspositionen r wird dann durch (Gleichung 2) bestimmt.

$$F(\vec{b}_k, T_i^{threat}, A_j^{event}) = \iint_{A_j^{event}} f(\vec{b}_k, T_i^{threat}, \vec{r}) dr_x dr_y \quad (2)$$

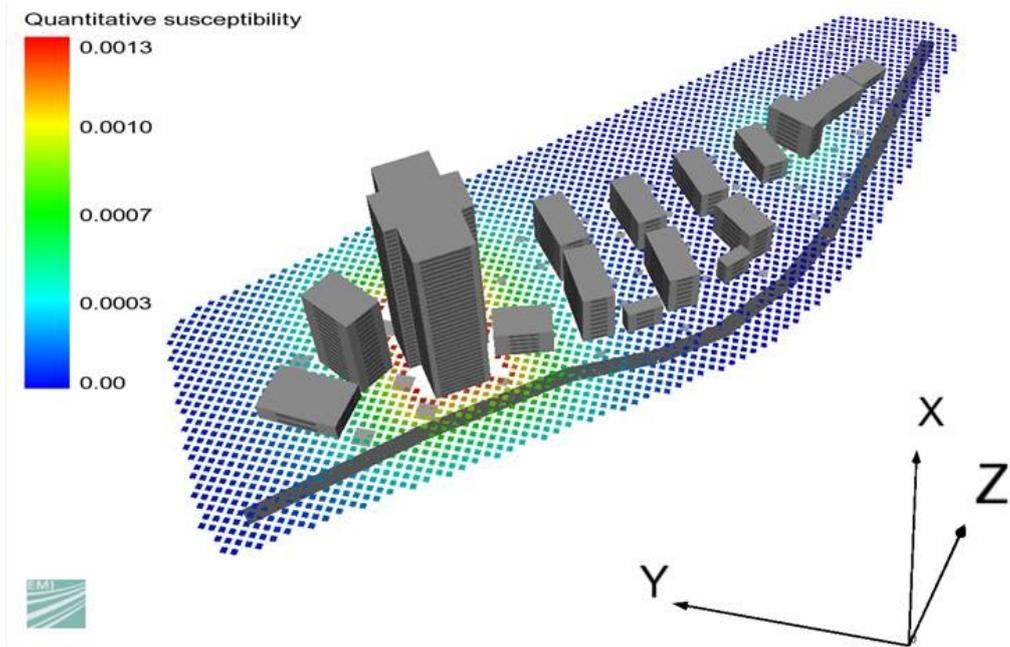


Abb. 2: Die Bedrohungspunkte erhalten Empfindlichkeitswerte in Abhängigkeit von ihrem Abstand zu umgebenden Objekten.

Die Verletzlichkeit V des Stadtgebiets ergibt sich aus der durchschnittlichen Summe aller Konsequenzen an den verschiedenen Ereignispositionen r_j nach (Gleichung 3).

$$V(H(T_i^{threat}; P), \vec{r}_o, T_n^{consequence}) = \frac{1}{n_{location}} \sum_{j=1}^{n_{location}} C(H(T_i^{threat}, r_j^{event}; P), \vec{r}_o, T_n^{consequence}) \quad (3)$$

Die konkreten Konsequenzen C des Explosionsereignisses werden dabei mithilfe verschiedener physikalischer Modelle für Strukturschäden, direkte und indirekte finanzielle Schäden sowie Personenschäden berechnet.

Die Berechnung einer Druckstoßbelastung erfolgt, indem Druck-Zeit-Ladefunktionen mit empirischen Gleichungen, beispielsweise nach Kingery-Bulmash (KINGERY & BULMASH 1984) **Fehler! Verweisquelle konnte nicht gefunden werden.** für die Massen-Abstands-Beziehungen verwendet werden. Der Schaden für Personen und Gebäude kann mit Hilfe von Probit-Funktionen (BOWEN et al. 1968) und vereinfachten mechanischen Modellen, wie zum Beispiel dem Einmassenschwinger (EMS) prognostiziert werden (RIEDEL et al. 2010) **Fehler! Verweisquelle konnte nicht gefunden werden.** Dabei werden die typischen Personendichten und Wiederaufbaukosten für alle Gebäudetypen und Nutzungsarten, sowie das Kollapsverhalten von erfahrenen Bauingenieuren bereitgestellt (MÜLLERS et al. 2015), wodurch quantitative Schadensaussagen zur Personengefährdung und zum Gebäudeschaden getroffen werden können.

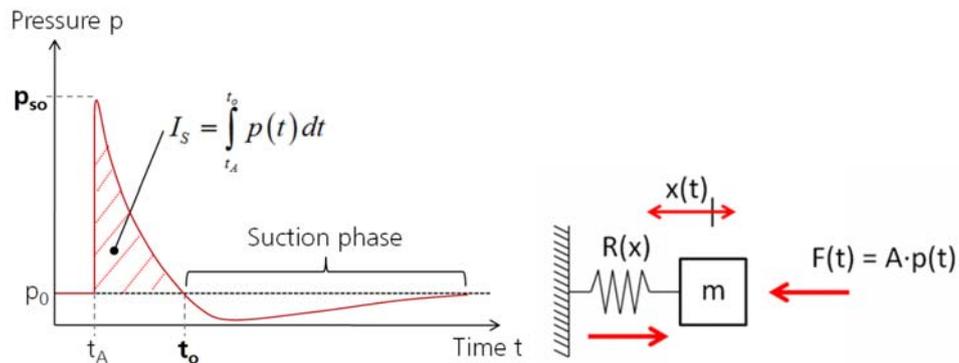


Abb. 3: Generalisierter Druck-Zeit-Verlauf während einer Explosion (links). Einmassenschwinger Modell zur vereinfachten Beschreibung des Strukturverhaltens unter Blastbelastung (rechts).

Das berechnete Ergebnis in Abb. , links unten, prognostiziert in 18% aller relevanten Ereignisse eine Zerstörung der (herkömmlichen) Verglasung des Hochhauses bis zur 8. Etage. Dieses lokale Einzelrisiko kann direkt mit den Einzelrisiken in Abb. 1 verglichen werden, wenn sie mit der Gesamt ereignisfrequenz für das Stadtgebiet aus (2) multipliziert werden.

Das lokale individuelle Risiko für Personenschäden erreicht neben dem Hochhaus im Freien 3% (Abb. links oben). Das Nachbarhaus auf der rechten Seite hat in seiner Umgebung immer noch ein vergleichbar hohes lokales Risiko, trotz seiner Verwendung als Wohngebäude.

Bezüglich des progressiven Kollaps ist das Hochhaus (Abb. rechts oben) voraussichtlich in 1 von 200 Ereignissen betroffen. Es verfügt über ein solides Tragwerk aufgrund seiner Größe und ist daher viel weniger Risiko ausgesetzt als das rot gefärbte Gebäude weiter rechts, für das fast jedes fünfte Ereignis zu einem Kollaps führen wird - ein Risiko, das es wert ist zu berücksichtigen.

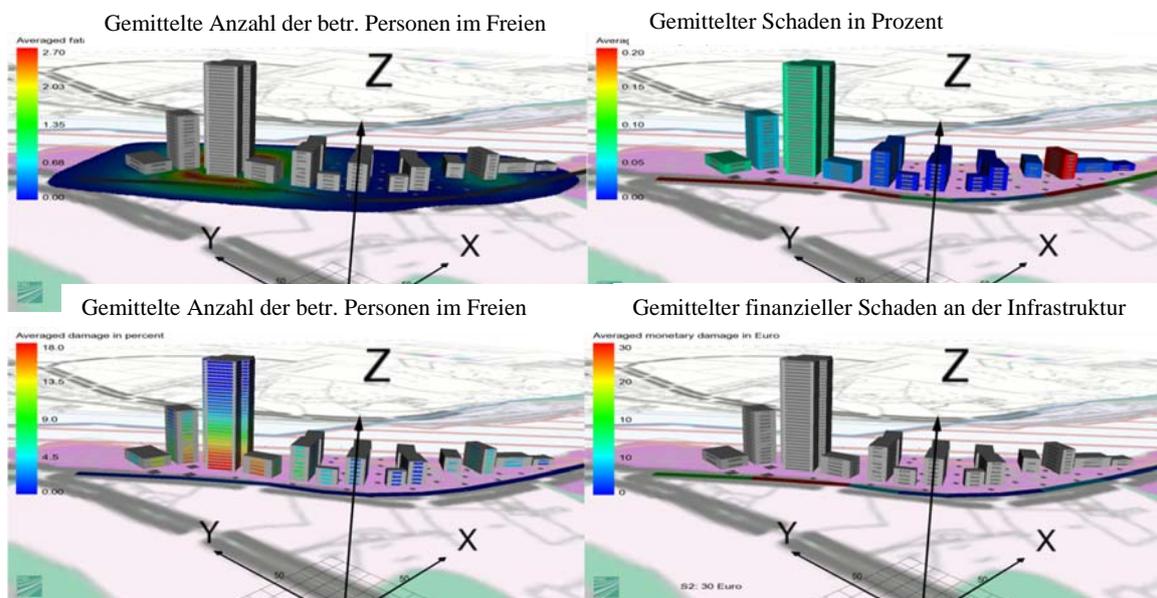


Abb. 4: Physikalische Konsequenzmodelle für die quantitative Risikoanalyse. Oben links: Personenschäden in Gebäuden nach einem angenommenen Anschlagsszenario. Oben rechts: Progressiver Einsturz von Gebäuden. Unten links: Fensterschäden. Unten rechts: Gemittelte finanzielle Schäden an der Infrastruktur.

Eine mögliche Verbesserungsmaßnahme in einer frühen Phase der Projektplanung ist eine Umgestaltung des urbanen Gebietes – also z.B. eine Verlegung von Gebäuden mit hoher Gefährdung in entsprechende zum Beispiel mit Pollern geschützte Areale. Des Weiteren kann die Konstruktion eines Gebäudes verstärkt werden. Dafür stehen verschiedene Optionen mit Kostenschätzungen zur Verfügung, wie z.B. duktiler Beton, Beton mit erhöhtem Bewehrungsgrad oder verstärktes Mauerwerk (STOLZ et al. 2014). Einige Verstärkungsmaßnahmen lassen sich auch auf bestehende Gebäude anwenden, wie zum Beispiel der Einsatz von Sicherheitsglas. Die kurzzeitdynamischen Berechnungsmodelle basieren auf Stoßrohr- und Sprengversuchen wie in Abb. dargestellt. In Abb. (rechts) ist die experimentelle Erprobung einer nachträglich installierten Schutzmaßnahme für Mauerwerk zu sehen. Das Produkt verhindert das Eindringen von Fragmenten ins Rauminnere und bietet damit ausgezeichneten Explosionsschutz.



Abb. 5: Experimentelle Stoßwellentests an normalem Glas (links) und Sicherheitsglas (Mitte). Experimentelle Erprobung einer Schutzmaßnahme für Mauerwerk (rechts).

Durch das UrbanRiskPortal wird das Expertenwissen von VITRUV für Stadtplaner und ihre Partner als Webservice zur Verfügung gestellt.

3 GeoToolbox

Bei der GeoToolbox³ handelt es sich um eine Sammlung von Werkzeugen zur Verarbeitung von Geodaten. Dabei sind die einzelnen Werkzeuge unabhängig voneinander oder in Kombination einsetzbar. Die Funktionalität geht dabei von einfachen Prozessen wie Informationen über einen Geodatensatz erfahren bis hin zu der automatischen Erstellung einer Plug-In freien Webvisualisierung. Alle Werkzeuge der GeoToolbox werden als Kommandozeilen-Tools bereitgestellt sowie als Web-Service zur Verfügung gestellt. Daher können sie ohne weitere Problem in das UrbanRiskPortal integriert werden. Dort finden sie zum einen Anwendung in der Aufbereitung der Geodaten und in der Erstellung einer Plug-In-freien Webvisualisierung.

³ www.igd.fraunhofer.de/geo

4 UrbanRiskPortal

Das UrbanRiskPortal vereint die Analyse- und Risikobewertungsfunktion der VITRUV-Software mit den Prozessierungsfunktionalitäten der GeoToolbox in einem Webportal. Der Anwender startet die Bedienung mit dem Upload des 3D-Stadtmodells. Dazu wird ein CityGML-Datensatz benötigt. Sobald der Datensatz hochgeladen wurde steht die Möglichkeit der 3D-Visualisierung des Datensatz zur Verfügung. Bevor die Analyse ausgewählt wird kann der Anwender Informationen über die Attributierung des Datensatzes beim sogenannten XML-Mapping vornehmen. Sollten Attribute in dem Datensatzes mit anderen Namen vorliegen, können in diesem Schritt die geforderten Attribute angegeben werden. Im letzten Schritt kann der Anwender die Analyse konfigurieren. Dabei hat er die Möglichkeit monetäre Schäden und Personenschäden zu betrachten. Nachdem die Analyse erfolgreich durchgeführt wurde kann der Nutzer erneut das 3D-Stadtmodell betrachten, wobei Gebäude entsprechend ihrer Gefährdungstufe eingefärbt werden.

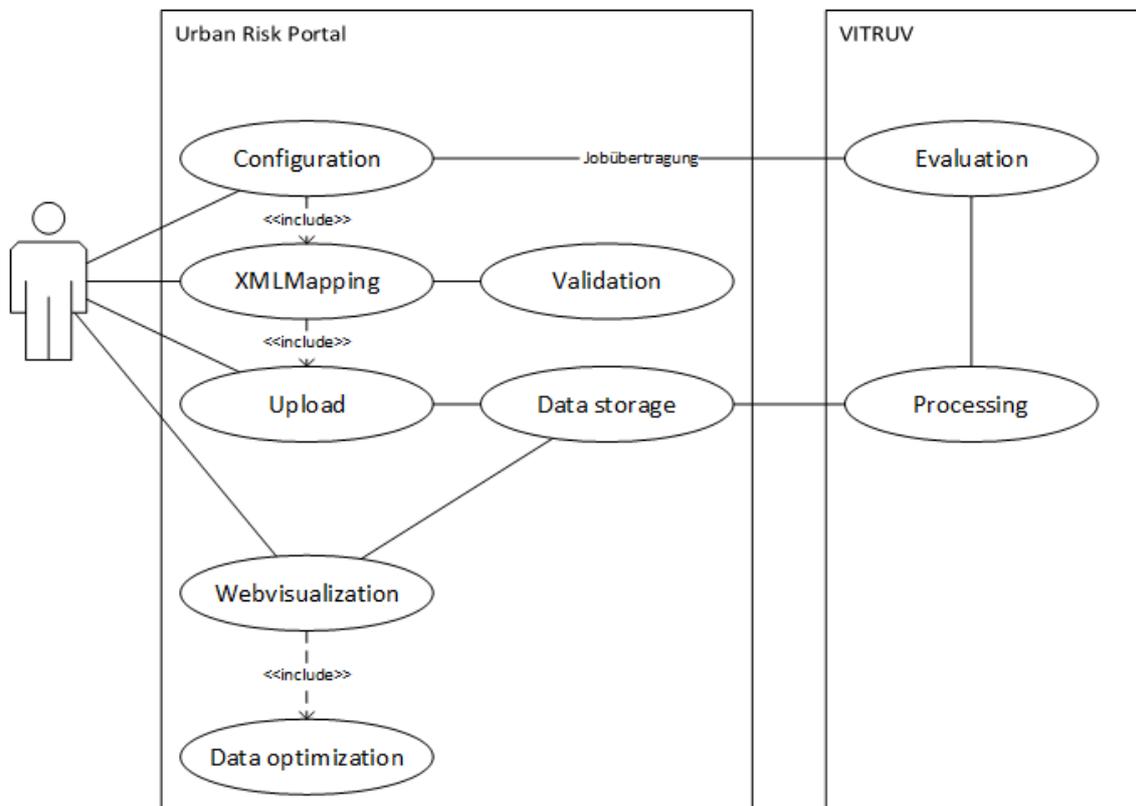


Abb. 6: Bedienung des UrbanRiskPortals aus Anwendersicht

4.2 3D Portrayal Service

Da die gesamte Benutzeroberfläche des UrbanRiskPortals webbasiert ist, ist eine entsprechende Ergebnisvisualisierung nötig. Da 3D-Stadtmodelle visualisiert werden sollen fiel die Entscheidung auf den 3D Portrayal Service (3DPS).

Mit Standards wie dem Web Map Service (WMS) ist es schon seit längerem Möglich 2D-Karten bereitzustellen, beziehungsweise in andere Applikationen einzubinden. Der 3D Portrayal Service

ist das Äquivalent des WMS für 3D-Geodaten. Er wurde entworfen um die Interoperabilität zu ermöglichen. Ein Entwurf von 3DPS wurde im Oktober 2015 veröffentlicht. Die Veröffentlichung der endgültigen Standards wird im Mai 2016⁴ erwartet. Der Standard unterstützt sowohl client- und serverseitiges Rendering und spezifiziert eine Schnittstelle um Daten von einem 3D-Geodaten-Server abzufragen und an einen Client zu übertragen.

Im Rahmen des Projekts wird der 3D Portrayal Service eingesetzt um eine ansprechende Visualisierung des 3D-Stadtmodells zu ermöglichen. Der Nutzer hat dabei die Möglichkeit das ursprüngliche Stadtmodell zu betrachten und direkt mit dem analysierten Stadtmodell zu vergleichen.

Das Ergebnis der Risikoanalysen wird farblich im 3D Stadtmodell dargestellt. Gebäude und Orte werden je nach Höhe der Gefährdungsstufe entsprechend eingefärbt.

5 Fazit

Durch die Kombination der hier beschriebenen Tools und Technologien ist es erstmals möglich eine Kombination von 3D-Stadtmodellen und quantitativer Risikoanalyse mit kurzzeitdynamischen Ingenieursmodellen zur Einschätzung der Gefährdungssituation im städtischen Gesamtkontext durchzuführen. Die ausfallsichere Bearbeitung von rechenintensiven Anfragen wird durch eine darunterliegende Cloud-Infrastruktur gewährleistet.

Die Risikoanalyse wird als Webdienst konzipiert, um so als „*Expertise on demand*“ ohne hohe, fixe Lizenzkosten dem Kunden zur Verfügung zu stehen. Die sicherheitsrelevanten und Know-How-intensiven Analyseverfahren bleiben in der sicheren Hand von Fraunhofer. Es wurde Wert auf einen intuitiven und einfachen Zugang über ein Web-Portal gelegt, damit Risikobewertungen auch von Anwendern durchgeführt werden können, die keine ausgewiesene Expertise in der Risikoanalyse haben.

Die konzipierte Systemarchitektur ermöglicht schnelle und effiziente Anpassungen an die jeweiligen Anforderungen zur Analyse und Visualisierung unterschiedlicher Szenarien.

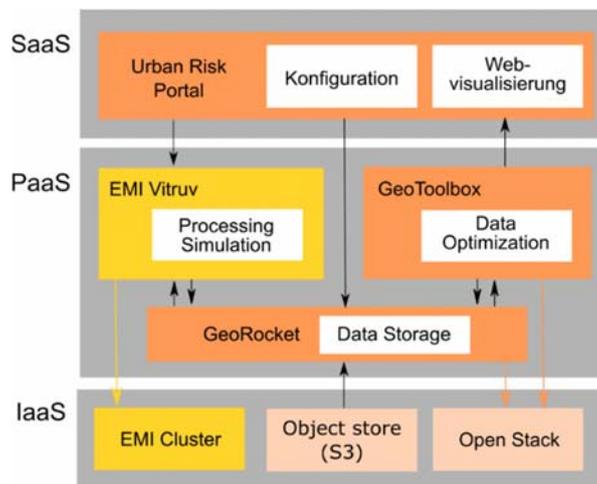


Abb. 7: 1 Die Komponenten des UrbanRiskPortals im Zusammenspiel

⁴ <http://www.opengeospatial.org/pressroom/pressreleases/2311/>

5 Literaturverzeichnis

- BOWEN, I, FLETCHER, E. & RICHMOND, D.R., 1968: Estimate of man's tolerance to direct effects of air blast. Headquarters Defense Atomic Support Agency, 37 p.
- FISCHER, K., SIEBOLD, U., VOGELBACHER, G., HÄRING, I. & RIEDEL, W., 2013: Empirical analysis of security critical events in urban areas. *Bautechnik*, **91**(4), 262-273.
- FISCHER, K., VOGELBACHER, G., HÄRING, I., RIEDEL, W. & HIERMAIER, S., 2016: Susceptibility, vulnerability and averaged risk analysis for resilience enhancement of urban areas. *International Journal of Protective Structures*, **7** (1), 45-76.
- HSE, 2014: UK Health & Safety Executive (HSE) Report. http://www.hse.gov.uk/foi/internal-ops/hid_circs/permissioning/spc_perm_37/
- KINGERY, C.N. & BULMASH, G., 1984: Airblast parameters from TNT spherical air bursts and hemispherical surface bursts. Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, US.
- LIN, C., LIOU, D. & WU, K., 2007: Opportunities and challenges created by terrorism. *Technological Forecasting & Social Change*, **74**, 148-164.
- MÜLLERS, I., FISCHER, K. & NAWABI, A., 2015: Design against Explosions and Subsequent Progressive Collapse. *Structural Engineering International*, **25** (3), 319-325.
- RIEDEL, W., THOMA, K., MAYRHOFFER, C., STOLZ A., 2010: Engineering and Numerical Tools for Explosion Protection of Reinforced Concrete. *International Journal of Protective Structures*, Multi-Science Publishing, **1**, 85-101.
- ROLLER, C. & STOLZ, A., 2014: A holistic protection portfolio for urban infrastructure. *Future Security - 9th Security Research Conference 2014*.
- STOLZ, A., FISCHER, K., ROLLER, C. & HAUSER, S., 2014: Dynamic bearing capacity of ductile concrete plates under blast loading. *International Journal of Impact Engineering*, **69**, Elsevier, 25-38.
- UNITED NATIONS, 2011: World Urbanization Prospects, the 2011 Revision. Department of Economic and Social Affairs, http://esa.un.org/unup/CD-ROM/WUP2011-F02-Proportion_Urban.xls.

Parameters Influencing Forest Gap Detection Using Canopy Height Models Derived From Stereo Aerial Imagery

KATARZYNA ZIELEWSKA-BÜTTNER¹, PETRA ADLER²,
MAIKE PETERSEN² & VERONIKA BRAUNISCH^{1,3}

Zusammenfassung: Lücken sind wichtige Strukturelemente für die Waldbiodiversität. Zur automatisierten Kartierung von Lücken in Relation zur umgebenden Bestandeshöhe und –bedeckung entwickelten wir eine Methode, welche auf von Stereo-Luftbildern abgeleiteten Kronenhöhenmodellen (CHMs) und einem LiDAR- Geländemodell beruht. Zur Evaluierung der Methode und der Bestimmung der wichtigsten Fehlerquellen wurden in einem 1021 ha großen Modellgebiet im Schwarzwald (Südwestdeutschland) die Kartiererergebnisse aus drei Befliegungen (2009, 2012, 2014) verglichen. Die Befliegungen von 2009 und 2012 hatten eine Bodenauflösung von 20cm und eine Überlappung von 60 % in Flugrichtung und 30% quer. 2014 war die Bodenauflösung 10cm und Überlappung 80%, respektive 60%. Die Validierung erfolgte durch visuelle Stereointerpretation. Schattenvorkommen und die geometrischen Grenzen der Stereobildauswertung wurden als Hauptfehlerquellen erkannt.

Abstract: Gaps in the canopy are important elements for forest biodiversity. We developed a method based on Canopy Height Models (CHMs) derived from stereoscopic aerial imagery and a LiDAR-based Digital Terrain Model (LiDAR DTM) to automatically delineate forest gaps in relation to height and cover of the surrounding forest. To evaluate the factors affecting the mapping accuracy, we compared the results from three different flight campaigns (2009, 2012 and 2014) in a 1021-ha model region in the Black Forest, Southwestern Germany. The public campaigns of 2009 and 2012 were taken with an overlap of 60% within stripe and 30% between stripes and an overall resolution on ground of 20cm. Data from 2014 had a 10cm resolution and an overlap of 80% within stripe and 60% between stripes. The validation was done by visual stereo-interpretation. Shadow occurrence and geometric limitations of the stereo aerial imagery were identified as main error sources.

1 Introduction

Forest gaps are considered important structural elements in forest ecology. They play a key role in forest regeneration processes (GETZIN et al. 2014) and provide suitable habitat structures for animal species that depend on semi open habitats (SIERRO et al. 2001; MÜLLER & BRANDL 2009; ZELLWEGER et al. 2013). Canopy gaps are therefore of great interest for research in the fields of stand structure and regeneration dynamics as well as biodiversity and nature conservation. In

¹ Forest Research Institute Baden-Württemberg (FVA), Department of Forest Nature Conservation, Wonnhaldestr. 4, D-79100 Freiburg, Germany; E-Mail: Katarzyna.Zielewska-Buettner@forst.bwl.de, Veronika.Braunisch@forst.bwl.de

² Forest Research Institute Baden-Württemberg (FVA), Department of Biometry and Information Sciences, Wonnhaldestr. 4, D-79100 Freiburg, Germany; E-Mail: Petra.Adler@forst.bwl.de, petersen.maike@outlook.de

³ University of Bern, Institute of Ecology and Evolution, Conservation Biology, Baltzerstrasse 6, CH-3012 Bern, Switzerland; email: veronika.braunisch@iee.unibe.ch

addition to the widely used traditional field-data collection for identification and quantification of the canopy gaps in ecological studies, the use of remote sensing data has been recently recognized as a good source of suitable data enabling the analysis of the canopy structure at various, often broad spatial scales. The first method that is usually chosen for forest gap detection (VEPAKOMMA 2010; VEPAKOMMA et al. 2012) and habitat mapping for biodiversity and nature conservation purposes (SEIBOLD et al. 2014; BRAUNISCH et al. 2014; BÄSSLER et al. 2010) is Light Detection and Ranging (LiDAR) that is considered to deliver a more detailed picture of the horizontal and vertical forest structure than any other remote sensing system. However the recent technical advances in the field of digital photogrammetry demonstrate the great potential of the automatic image matching for the generation of Canopy Height Models (CHMs) and for deriving important forest parameters (STRAUB et al. 2013; WANG et al. 2015; KOTREMB 2014, BETTS et al. 2005). Thus, to assess the viability of gap detection based on publicly available data we focused our research on CHMs derived from the standard stereo aerial imagery and the official LiDAR based Digital Terrain Model (LiDAR DTM), which are delivered in regular time intervals by the regional mapping agency of Baden-Württemberg (LGL). We aimed for a gap mapping tool which would deliver standardized and replicable results when applied on publicly available data either in form of original aerial imagery, point clouds or a raster CHM.

Gaps were detected and delineated in relation to height and cover of the surrounding forest in three steps: (1) open and dense forest are identified, (2) dense forest is classified into low and high forest and (3) gaps are extracted in the latter two classes. The method is described in ZIELEWSKA-BÜTTNER et al. (2016). In this conference paper we present parameters influencing the method performance with regard to canopy gaps detection (1). In addition we test in more detail the benefits of using a shadow mask (2) and discuss effects associated with variance in flight conditions (3). We also consider the variance introduced by different image matching algorithms (4). Finally, the influence of spatial resolution and overlap of the stereo aerial images are presented comparing the results obtained with data of different flight campaigns (5).

2 Material and Method

2.1 Study area

The study area of 1021 ha (excluding the mountain lake surface of the Huzenbacher See) is located in the State of Baden-Württemberg, Southwestern Germany, in the northern Black Forest (8° 34' E, 48° 58' N). It is characterized by a heterogeneous topography with elevation ranging from 493 to 941 m, and a high variance in forest successional stages. Most slopes (77,5 %) are very steep (> 20°) or strongly inclined (10 - 20°) (AG BODEN 1996). Among the dominant tree species are Norway spruce (*Picea abies* L.) with admixture of Silver fir (*Abies alba* Mill.) and Scots pine (*Pinus sylvestris*). The broadleaved tree species account for less than 30 % in most (> 80 %) forest stands. The area is covered by a dense forest road network of 187 m/ha and underlying different protection regimes.

2.2 Remote sensing data

As primary input data for the method development, aerial imagery datasets from three flight campaigns (2009, 2012, 2014) were used. Data (including the absolute orientation of the images)

were provided by the state agency of spatial information and rural development of Baden-Württemberg (LGL) as pan sharpened, 4 channels (red, green, blue and near-infrared (RGB NIR)) stereo aerial images with radiometric resolution of 8 (2009) and 16 (2012 and 2014) bit. Data of 2014 originated from a special flight campaign of the Black Forest National Park. The overall spatial resolution of the imagery was 20 cm with an overlap of 60 % (end lap) and 30 % (side lap) in 2009 and 2012; and 10cm, 80% and 60% respectively, in 2014. In line with our goal of using only publicly available data, we limited the additional data used in the study to the products of the LGL (LiDAR DTM) or internal data of the forestry administration (forest road network dataset).

Tab. 1: Technical characteristics of the aerial image data used in the method development (2009, 2012) and the higher resolution and overlap data comparison (2014) (from ZIELEWSKA-BÜTTNER et al. 2016, modified)

Year	2009	2012	2014
Camera	UltraCamXp	DMC II 140 – 006	<i>UltraCamXp</i>
Panchromatic / color lens focal length	100 / 33 mm	92 mm	<i>100.5 mm</i>
Resolution	20 cm	20 cm	<i>10 cm</i>
Overlap	60 % / 30 %	60 % / 30 %	<i>80 % / 60 %</i>
Image type	Digital color infrared (RGB NIR)	Digital color infrared (RGB NIR)	<i>Pansharpened digital color infrared (RGB NIR)</i>
Angle-of-view from vertical, cross track (along track)	55° (37°)	50,7° (47,3°)	<i>55° (37°)</i>
No. of stripes in the block file	3	6	<i>4</i>
No. of pictures in the block file	23	48	<i>69</i>
Flight date	23.05.2009	01.08.2012	<i>17.07. – 19.07.2014</i>

2.3 Gap mapping method

The gap mapping method was based on Canopy Height Models (CHMs) of 1 m ground resolution including the potential vegetation points of height between -1 and 55 m vs. the LiDAR DTM. The Digital Surface Models (DSMs) serving as basis for the CHMs generation were calculated from the stereoscopic aerial imagery using two image matching algorithms: Leica Photogrammetry Suite enhanced Automatic Terrain Extraction (LPS eATE (ERDAS 2012)) and Semi Global Matching (SGM XPro (Hexagon Geospatial 2015)). As the two algorithms returned different point clouds partially complementing each other, in the initial study for the method development we decided, based on visual assessment, for a combination of three point clouds from eATE and SGM processed with the pyramid levels 0, 1 and 2 respectively to reach the best point coverage in a reasonable processing time. The detailed settings of both algorithms and the single processing steps are given in ZIELEWSKA-BÜTTNER et al. (2016). The point cloud editing was carried out with LAStools (ISENBURG 2014) whereas the LAS to a raster transformation was done in ArcGIS (“LasDataset to Raster”). For the gap detection a constantly closed surface was produced by filling the no-data areas with a including inverse distance weighting (IDW) interpolation method.

The gap detection was carried out in ArcGIS 10.3 (ESRI 2014) (raster and vector based) in three steps: (1) identification of open and dense forest, (2) classification of dense forest into height classes of low and high forest and (3) gap extraction in the latter two classes. We defined gaps as canopy openings in dense forest ($\geq 60\%$ canopy cover) of at least 10 m^2 reaching through all forest strata down to maximum 2 m vegetation height in high forest stands ($\geq 8 \text{ m}$ height) and down to maximum 1 m in low forest stands ($< 8 \text{ m}$ height). A minimum stand size of 0.3 ha is related to the size of the conventional minimum stand size in Baden-Württemberg (MATHOW 2016). Areas with canopy cover less than 60% and exceeding 0.5 ha were classified in line with AHRENS et al. (2004) as “open forest”, where the free spaces between the trees are considered as inherent stand characteristic and thus not mapped as gaps. 10 m^2 is the minimum size of a gap defined in line with MÜLLER & WAGNER (2011) and SCHLIEMANN & BOCKHEIM (2011). The maximum gap-vegetation height was set to 2 m after BROKAW (1982) and adapted by the authors to 1 m in the lower stands.

2.4 Validation

To evaluate the gap mapping performance we compared the automatic mapping results with the visual stereo-interpretation of the original aerial imagery on an independent dataset of sample plots using Stereo Analyst for ArcGIS 10.2 (GEOSYSTEMS GMBH 2014). As we expected the results to vary in relation to the terrain situation, 120 plots with a radius of 25 m (covering 2.4 % of the dense forest area) were placed according to a stratified random design into stands of three steepness classes ($0\text{-}10^\circ$, $10\text{-}20^\circ$, $>20^\circ$ (AG BODEN 1996)) and four aspect classes (N, E, S, W), resulting in 12 terrain classes represented by 10 sampling plots each. Gaps with an area of at least 10 m^2 inside the plot (168 in 2009 and 171 in 2012) were visually assessed, delineated and compared with the automatically mapped gaps located with at least 10 m^2 inside the evaluation plot. The gap-absence was evaluated on circles of 95 m^2 (mean size of the visually mapped gaps in both years) randomly placed in dense forest within the sampling plots in an amount equal to the visually verified gaps per year. At least 8 m^2 (80% of the minimum gap size) of overlap with the visually identified gaps was needed to confirm the correct classification of the automatically detected gaps or to classify a “non-gap” circle as incorrect. The agreement between visual and automatic mapping was then quantified in form of overall, producers’ and users’ accuracy as well as Cohen’s Kappa. An effect of selected parameters such as height of the surrounding forest, shadow occurrence (assessed visually), gap size, slope, aspect and gap location in relation to forest road, skidding trail or an open area (storm throw, open forest) on gap mapping results was tested using the Conditional Inference Trees (ctree) vignette of R-package “partykit” (HOTHORN et al. 2006).

To evaluate the influence of the missing original height information on the gap mapping performance, a no-data mask was generated as a raster of 1 m resolution from the final point clouds (combination of eATE and SGM pyramid 1 and 2 point clouds) of 2009 and 2012. It included only the raster cells, where no points were directly matched during the image matching process. As for the gap detection a constantly closed surface was used, by a comparison with the no-data mask the resulting improvement in accuracy was evaluated.

To quantify the effect of sun elevation on the image quality, we calculated a shadow mask for the data of 2009 and 2012. We defined as shadow an area without any textural differentiation. The classification was done by a visually defined threshold. As the data from the two study years had different radiometric resolution, two different methods had to be used to calculate the shadow fraction in the aerial images. For the images from 2009 that had been resampled to 8-bit resolution we used for this year a Ratio S calculated according to SARABANDI et al. (2004) as $S = \arctan(\text{Blue}/\max\{\text{Red}/\text{Green}\})$. For 2012 data with 16-bit resolution we used the Intensity channel of the transformed images (CONRAC CORP. 1980).

2.5 Comparison with data of higher overlap and resolution

To evaluate the potential influence of higher resolution and overlap of the aerial imagery on the method performance, PETERSEN (2015) applied the gap mapping method to a study polygon of 95 ha located in the south-western corner of the original research area (Fig. 1) using data from a special flight campaign of the National Park Black Forest in 2014 (Tab. 1). Gap mapping results based on these pansharpened RGB NIR aerial imagery of 10 cm resolution and 80% and 60% end and side lap were compared with those obtained from the lower-resolution public data of 2012, using the eATE algorithm for point cloud generation. The CHMs used for gap extraction in 2014 were calculated based on the LiDAR derived DTM of the National Park Black Forest, as obtained from their own flight campaign in 2015.

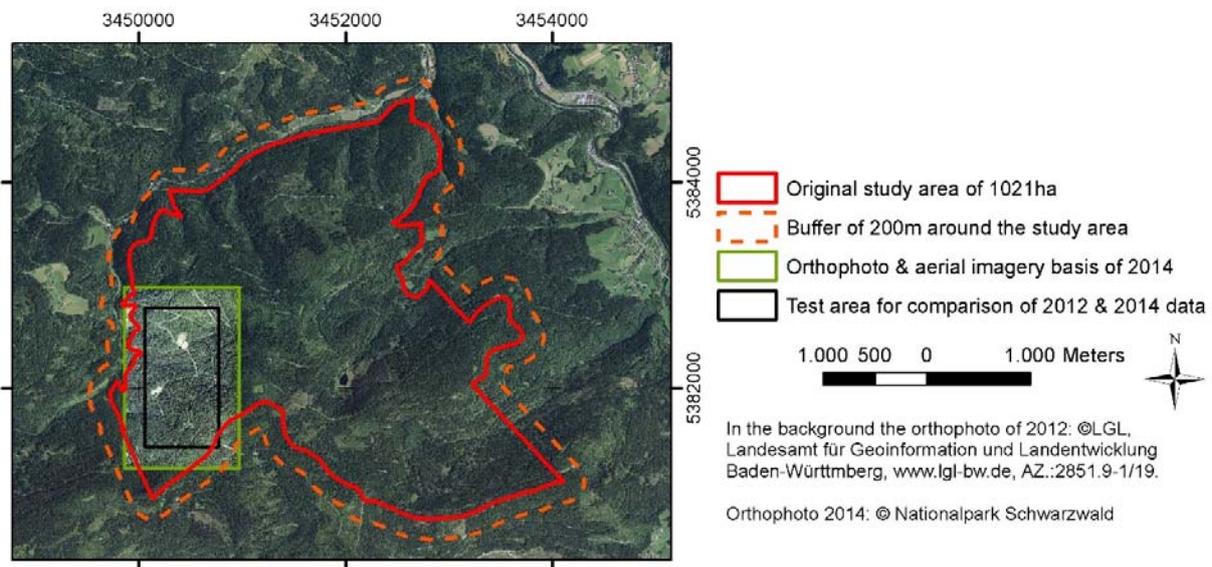


Fig. 1: Location the test area for comparison of 2012 and 2014 data within the original study area presented on the background of the available orthophotos from 2012 and 2014.

3 Results

3.1 Gap mapping results

We detected 4575 (2009) and 4667 (2012) gaps in the dense forest of the study area using the automated method, what results in a total gap density of 4.9 gaps per ha (7.2 % of the dense

forest area) in 2009 and 4.7 gaps per ha (6.3 %) in 2012. Considering the forest height classes, more gaps (13.7 and 14.6 N/ha in 2009 and 2012) covering a greater area (45 and 46 ha respectively) were mapped in the forest stands lower than 8 m compared to the higher forests with a gap density of 2.0 and 2.8 N/ha and mapped gap area of 25 and 16 ha in 2009 and 2012, respectively. The most (> 75%) of all detected gaps in both study years were very small or small (less than 100 m²) accounting for 13 % and 23% of the total gap area per year in 2009 and 2012, respectively. The visual validation resulted in an overall accuracy of 0.90 and 0.82 in 2009 and 2012 and the corresponding Kappa values of 0.80 and 0.66 (Tab. 2). Producer's accuracies greater than 0.96 confirmed almost all automatically detected gaps as correctly classified. Yet, a fraction of the visually identified gaps were not detected during the automated mapping process, which is reflected in lower user's accuracies of 0.84 in 2009 and 0.72 in 2012. However, more than 70 % of the visually but not automatically identified gaps in both study years were adjacent to the automatically mapped gaps, what suggest that gaps were correctly localized, but they were detected with a too small extent.

Tab. 2: Mapping accuracies of automatically generated gaps per year and forest high class derived from a comparison with the results of visual interpretation (accessed with 95 % confidence interval (CI)) (from ZIELEWSKA-BÜTTNER et al. 2016, modified)

		Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy	Kappa	Overall accuracy
		Gap	Gap	"Non-gap"	"Non-gap"		with 95 % CI
2009	DF	0.97	0.84	0.84	0.97	0.80	0.90
	LF	0.98	0.93	0.68	0.89	0.73	0.93
	HF	0.98	0.70	0.87	0.98	0.73	0.88
2012	DF	0.96	0.72	0.73	0.96	0.66	0.82
	LF	0.98	0.85	0.59	0.94	0.93	0.86
	HF	0.96	0.52	0.76	0.96	0.84	0.79

3.2 Shadow occurrence

Among the variables tested only the height of the surrounding forest and shadow occurrence significantly affected the gap mapping results. The occurrence of full shadow in the lower sections of the forest canopy was identified as the main cause for gap mapping omission errors in both years (ctree, $p < 0.001$). This was confirmed also by means of visual verification, as the most of the visually identified but not automatically mapped gaps (70%–87%) were identified in areas of total or partial shadow. The height of the surrounding forest stands (LF and HF) is strongly linked to shadow occurrence as it determines the depth in the canopy, to which the light can penetrate. Despite similar producer's accuracies of 0.96-0.98 and overall accuracies of more than 0.79, gaps in LF were mapped with higher user's accuracies than those in HF (0.93 vs. 0.70 in 2009 and 0.85 vs. 0.52 in 2012).

The shadow masks identifying complete shadow cells covered 29 % of the study area in 2009 and 16 % in 2012. However, the comparison with the location of automatically mapped gaps showed that only less than 5 % of the gaps were automatically detected in these areas. Shadow

occurrence was mostly linked to steep slopes and exposition as well as to heterogeneous vertical structure and stand height (Fig. 2), indicating a strong influence of the sun angle and associated time of data acquisition. The two flight campaigns of 2009 and 2012 were carried out in May and August, so the sun position at the time of data acquisition didn't correspond which produced shadow in different areas (Fig. 2).

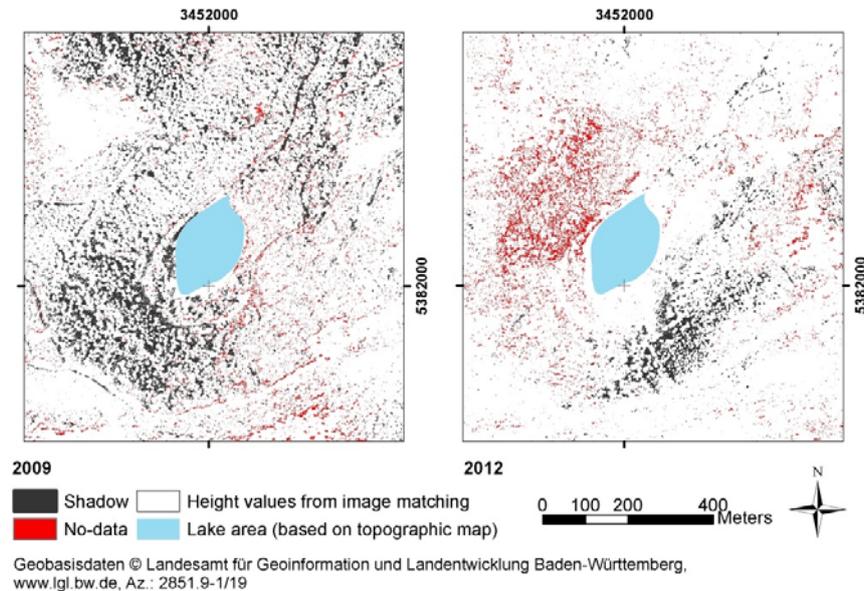


Fig. 2: Example of complete shadow and no-data cells distribution in a steep part of the study area around the mountain lake “Huzenbacher See”

3.2.1 Image matching algorithm

The amount and distribution of no-data cells was influenced by the algorithms and pyramid levels of the images used for point matching. The area with no data ranged between 38% with SGM pyramid level 2 in 2012 and 9 % with SGM pyramid level 1 in 2009. Combining different algorithms and pyramid levels led to a reduction of no-data within the study area to less than 3%. Evaluating whether missing information in some raster cells could be a reason for a fraction of the undetected gaps, we found that only about 10 % of the visually, but not automatically identified gap cells (gap area) in both years belonged initially to the no-data cells.

The distribution of no-data and shadow raster cells (Fig. 2) revealed that the points were mismatched not only in shadowy areas of the forest stands (8 % (2009) and 5 % (2012) of no-data cells intersected with the shadow mask) but also in low forest stands and on hilltops where aerial photographs should theoretically deliver good material for image matching. No-data cells were often located along flight strips (2009) or at the outer parts in the overlapping zone of the images (both years).

3.2.2 Image resolution and overlap

Comparing the results based on the original data from 2012 and the high-resolution dataset of 2014, a slightly larger amount of open forest (3 % in 2012 and 2 % in 2014) was mapped for 2012. Also the percentage of low forest was higher in 2012 (21 %) than in 2014 (14 %) (Fig. 3).

To be able to compare the results of the gap mapping of both years the study area was reduced by the area that had been classified as open forest within either of the datasets. For the remaining area of 94 ha a larger total gap area was obtained with the dataset of 2012 (4.9 ha) compared to the dataset of 2014 (2.5 ha) though a larger number of gaps was identified in the latter (2012: 240 gaps, 2014: 281 gaps). The reason for this can be found in the size of the mapped gaps. While more very small ($10\text{ m}^2 - 30\text{ m}^2$) and small ($31\text{ m}^2 - 100\text{ m}^2$) gaps were detected with the dataset of 2014 there were more large ($100\text{ m}^2 - 1000\text{ m}^2$) and very large gaps with a size of more than 1000 m^2 mapped with the data of 2012 (Fig. 4). The large and very large gaps were located mostly within the class of low forest or along forest tracks.

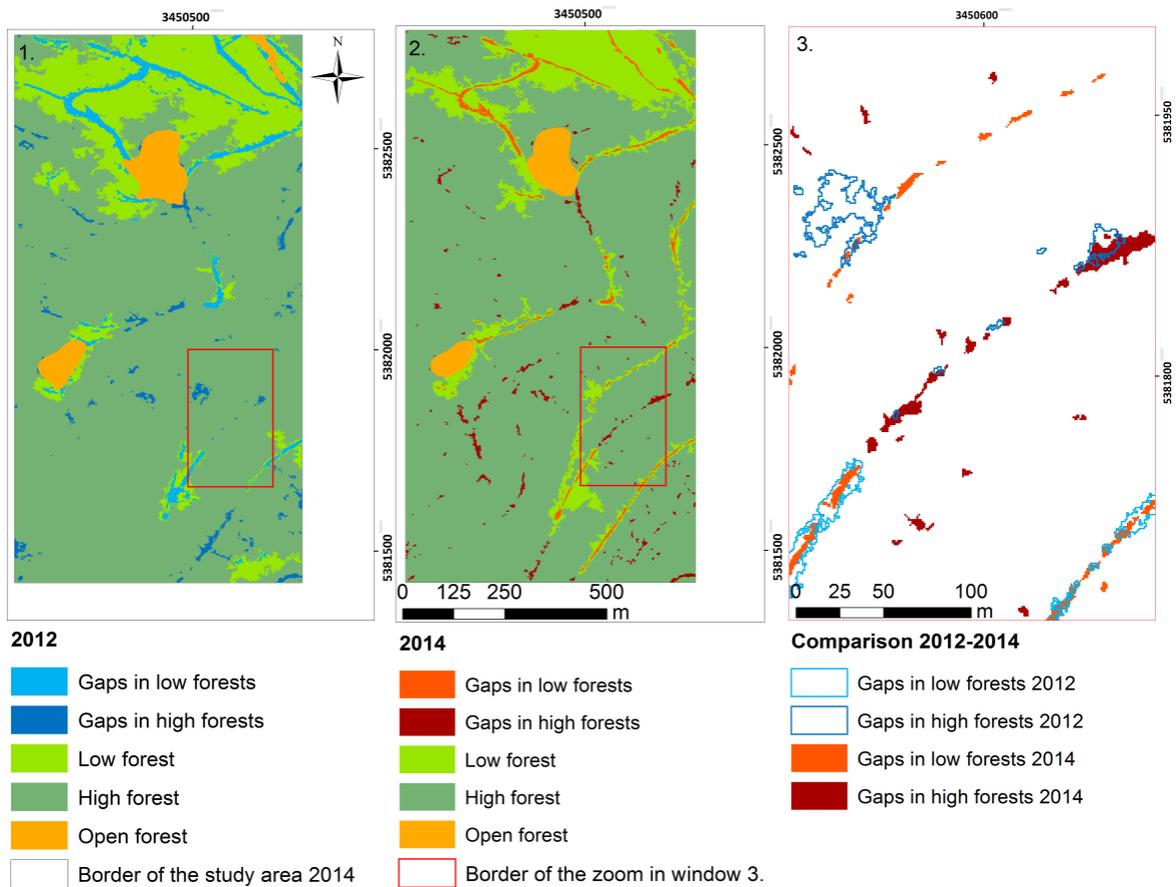


Fig. 3: Results of the automated gap mapping in the test area for the comparison of data with different resolution and overlap: 1) results from 2012 (20 cm, overlap 60% /30%), 2) results from 2014 (10 cm, 80% /60%), 3) Zoom-in window as example for a comparison of 2012 and 2014 results.

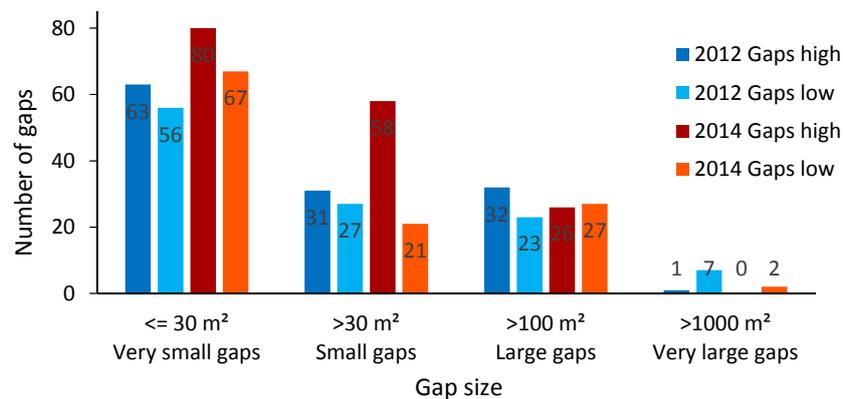


Fig. 4: Distribution of gap sizes in the dataset of 2012 (blue bars) and 2014 (red/orange bars). “High” and “low” indicate high and low forest stands.

4 Discussion and Conclusions

Gap mapping from stereo aerial imagery with 20 cm resolution and 60/30% overlap proved to deliver promising results, with a good overall method performance and even very good results in low stands. Depending on the quality of the aerial imagery and the input CHM as well as the height of the surrounding forest and the associated shadow occurrence, the results in stands higher than 8 m were moderate to insufficient, depending on the study year. The mapping results might also depend on the topography and the structure of the forest stands. GINZLER & HOBI (2015) observed better mapping accuracies in a flat terrain than in rugged mountainous topography, which also characterised our study area, whereas ADLER et al. (2014) found that even in flat terrain different DSM matching algorithms produce different results, especially in highly structured canopy situations. In addition, mountainous forests are likely more structured than intensively managed stands in the lowlands.

Shadow occurrence in aerial images is related to exposition, surface characteristics and caused by sun inclination and angle. Therefore the occurrence and distribution of shadow varies a lot between the different flight campaigns, especially in hilly areas comparable to our study area. Comparing the overall amount of shadow within the images with the shadow pixels within the classified gaps, we see that only a very small portion of gaps was detected in the shadow areas. This can be influenced by a strong fragmentation of the shadow areas with single patches not bigger than 10m² (minimum gap size). It can also be interpreted, that shadow pixels may have influenced negatively the image correlation for point matching within stand surface openings. The latter argument was confirmed by the visual interpretation of the gaps as 70-80% of visually interpreted gaps that hadn't been detected automatically were located in partial or total shadow. The superimposition of no-data areas with shadow areas showed no direct correlation. The appearance of no-data areas was more related to the geometric characteristics of the image, as they mainly occurred along image und flight strip borders.

By combining point clouds generated with two different image matching methods we expected a compensation effect and improvement of the point cloud structure in areas where no points were matched using only one of the algorithms. The results of our study underline the importance of

the image matching method. Not all algorithms perform equally well with regard to specific mapping goals e.g. detection of canopy gaps, mapping of the tree tops or calculation of forest stand parameters. Developments in technology and image matching algorithms are rapid, which makes the choice of the “best” algorithm combination very difficult, with “best” being often only valid for the used data and software combination.

The data used to analyse the influence of image resolution and overlap originated from two different years. The images with 20 cm resolution and 60/30 % overlap were taken in 2012, whereas the images with 10 cm resolution and an overlap of 80/40 % were from 2014. However, as there were no disturbances in the two years between the flights, and the forest stands wasn't in an age-class where natural mortality causes the disappearance of single trees, a decrease in detected gaps would have been expected. Nevertheless, the number of forest gaps increased from 240 to 281 with simultaneous decrease in size. This change in average gap size could be either explained by vegetation growth, at the gap edges, which reduces gap size. The increasing number of very small and small gaps could be explained ingrowth of vegetation, partially closing larger forest gaps, leaving more and smaller gaps behind. A visual examination, however, showed that the many of these very small and small gaps detected in 2014 were not mapped in 2012. Especially within high forest stands the number of detected gaps rose by almost 30 % while the amount of high forest only increased by 8 % between the years. This increase can not only be explained by ingrowth but by a better insight into the canopy structure due to a higher image overlap and resolution in the 2014 data.

The results from different flight campaigns indicate shadow occurrence and geometric limitations of the aerial imagery as serious constraints, both bearing a high potential for improvement. Flight campaigns should consider the issues arising from varying flight time and associated solar altitude. Moreover, an increase in spatial resolution and overlap of the aerial images could considerably improve the spatial accuracy of the results. Further improvements can be expected from an amelioration of the image matching algorithms. The use of shadow and no-data masks proved useful for the interpretation and evaluation of the automatically produced gap maps and we recommend them especially for change detection. Further research on these topics could help to optimize and standardize future flight campaigns, so that they can be used for reliable monitoring of gaps and other forest structure parameters.

5 Acknowledgements

We thank Joao Paulo Pereira for his inputs into gap mapping and generation of the no data masks and Martin Denter for calculation of the shadow masks. The study was carried out at the Forest Research Institute Baden-Württemberg (FVA) with the data contribution from the Black Forest National Park © Nationalpark Schwarzwald and the State Agency of Spatial Information and Rural Development of Baden-Württemberg (“Geobasisdaten ©LGL, Landesamt für Geoinformation und landentwicklung Baden-Württemberg, www.lgl-bw.de, AZ.:2851.9-1/19.)

6 Bibliography

- ADLER, P., NAAKE, T., PETERS, S., GINZLER, C., BAUERHANSL, C. & STEPPER, C., 2014: Reliability of forest canopy height extraction from digital aerial images. ForestSAT Conference, 4-7. November 2014, Riva del Garda (TN), Italy.
- AG BODEN, 1996: Bodenkundliche Kartieranleitung. 4. Auflage, Ad-hoc-Arbeitsgruppe Boden der Geologischen Landesämter und der Bundesanstalt für Geowissenschaften und Rohstoffe der Bundesrepublik Deutschland, Nachdruck, Bundesanstalt für Geowissenschaften und Rohstoffe und Geologische Landesämter: Hannover, 392.
- AHRENS, W., BROCKAMP, U. & PISOKE, T., 2004: Zur Erfassung von Waldstrukturen im Luftbild. Arbeitsanleitung für Waldschutzgebiete Baden-Württemberg. Waldschutzgebiete Baden-Württemberg **5**, 54.
- BÄSSLER, C., STADLER, J., MÜLLER, J., FÖRSTER, B., GÖTTLEIN, A. & BRANDL, R., 2010: LiDAR as a rapid tool to predict forest habitat types in Natura 2000 networks. *Biodiversity and Conservation* **20**, 456-481.
- BETTS, H.D., BROWN, L. J. & STEWART, G.H., 2005: Forest canopy gap detection and characterisation by use of high-resolution Digital Elevation Models. *New Zealand Journal of Ecology* **29**, 95-103.
- BRAUNISCH, V., COPPES, J., ARLETTAZ, R., SUCHANT, R., ZELLWEGER & F., BOLLMANN, K., 2014: Temperate mountain forest biodiversity under climate change: compensating negative effects by increasing structural complexity. *PloS one* **9**, e97718.
- BROKAW, N.V.L., 1982: The Definition of Treefall Gap and Its Effect on Measures of Forest Dynamics. *Biotropica* **14**, 158-160.
- CONRAC CORP., 1980: Raster Graphics Handbook; Conrac Corp.: Covina, California.
- ERDAS, ©INTERGRAPH CORPORATION., 2012: LPS EATE. From Image to Point Clouds: Model Your World Like Never Before. Product Sheet. Vol. 5/12 GEO-US-0033A-ENG.
- ESRI, 2014: ArcMap 10.3. Redlands, California.
- GEOSYSTEMS GMBH, ©INTERGRAPH® CORPORATION., 2014: Stereoauswertung mit Erdas extensions für ArcGIS®. Product sheet. Germering.
- GETZIN, S., NUSKE, R. S. & WIEGAND, K., 2014: Using Unmanned Aerial Vehicles (UAV) To Quantify Spatial Gap Patterns In Forests. *Remote Sensing* **6**, 6988-7004.
- GINZLER, C. & HOBI, M. L., 2015: Countrywide Stereo-Image Matching for Updating Digital Surface Models in the Framework of the Swiss National Forest Inventory. *Remote Sensing* **7**, 4343-4370.
- HEXAGON GEOSPATIAL, ©INTERGRAPH® CORPORATION, 2015: ERDAS IMAGINE Help. XPro Semi-Global Matching.
- HOTHORN, T., HORNIK, K. & ZEILEIS, A., 2006: Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* **15**, 651–674.
- HOTHORN, T.; ZEILEIS, A. PARTYKIT: A modular toolkit for recursive partytioning in R. Reference Manual. <https://cran.r-project.org/web/packages/partykit/partykit.pdf> (accessed on 3 August 2015).
- ISENBURG, M., 2014: Rapid lasso - fast tools to catch reality. 2007-2014.

- KOTREMB, C., 2014: Hochauflösende fernerkundliche Erfassung von Waldstrukturen im GIS am Beispiel der Kernzone "Quallgebiet der Wieslauter" im Pfälzerwald. *AFZ-DerWald* **9**, 12-15.
- MATHOW, T., 2015: Personal communication. ForstBW. Referat 84 Forsteinrichtung und Forstliche Geoinformation. Forstdirektion. Regierungspräsidium, Freiburg,
- MÜLLER, J. & BRANDL, R., 2009: Assessing biodiversity by remote sensing in mountainous terrain: the potential of LiDAR to predict forest beetle assemblages. *Journal of Applied Ecology* **46**, 897-905.
- MÜLLER, K.H. & WAGNER, S., 2003: Störungslücken in Fichtenbeständen des Erzgebirges: Initiale eines Waldumbaus. *Forst und Holz*, 58. Jahrgang **13/14**, 407-411.
- PETERSEN, M., 2015: Influence of Different Aerial Datasets on the Characterisation of Vertical and Horizontal Forest Stand Structure. Master Thesis, Albert-Ludwigs University, Freiburg. 65.
- SARABANDI, P., YAMAZAKI, F., MATSUOKA, M. & KIREMIDJIAN, A., 2004: Shadow Detection and Radiometric Restoration in Satellite High Resolution Images. In *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings, IEEE International: Anchorage, Alaska*, **6**, 3744-3747.
- SCHLIEMANN, S.A. & BOCKHEIM, J. G., 2011: Methods for studying treefall gaps: A review. *Forest Ecology and Management* **261**, 1143-1151.
- SEIBOLD, S.; BRANDL, R.; BUSE, J.; HOTHORN, T.; SCHMIDL, J.; THORN, S. & MÜLLER, J., Association of extinction risk of saproxylic beetles with ecological degradation of forests in Europe. *Conservation Biology* **0**, 1-9.
- SIERRO, A., ARLETTAZ, R., NAEF-DAENZER, B., STREBEL, S. & ZBINDEN, N., 2001: Habitat use and foraging ecology of the nightjar (*Caprimulgus europaeus*) in the Swiss Alps: towards a conservation scheme. *Biological Conservation* **98**, 325-331.
- STRAUB, C., STEPPER, C., SEITZ, R. & WASER, L. T., 2013: Potential of UltraCamX stereo images for estimating timber volume and basal area at the plot level in mixed European forests. *Canadian Journal of Forest Research* **43**, 731-741.
- WANG, Z., WASER, L. T. & GINZLER, C., 2015: A novel method to assess short-term forest cover changes based on digital surface models from image-based point clouds. *Forestry*, 1-12.
- VEPAKOMMA, U., 2012: Spatial contiguity and continuity of canopy gaps in mixed wood boreal forests: persistence, expansion, shrinkage and displacement. *Journal of Ecology* **100**, 1257-1268.
- VEPAKOMMA, U.; KNEESHAW, D. & FORTIN, M-J., 2010: Interactions of multiple disturbances in shaping boreal forest dynamics: a spatially explicit analysis using multi-temporal lidar data and high-resolution imagery. *Journal of Ecology* **98**, 526-539.
- ZELLWEGER, F.; BRAUNISCH, V.; BALTENSWEILER, A. & BOLLMANN, K., 2013: Remotely sensed forest structural complexity predicts multi species occurrence at the landscape scale. *Forest Ecology and Management* **307**, 303-312.
- ZIELEWSKA-BÜTTNER, K., ADLER, P.; EHMANN, M. & BRAUNISCH, V., 2016: Automated Detection of Forest Gaps in Spruce Dominated Stands Using Canopy Height Models Derived from Stereo Aerial Imagery. *Remote Sensing* **8**, 175.

Verwendung von multispektralen Sentinel-2 Daten für die Baumartenklassifikation und Vergleich mit anderen Satellitensensoren

MARKUS IMMITZER¹, FRANCESCO VUOLO¹, KATHRIN EINZMANN¹, WAI TIM NG¹,
SEBASTIAN BÖCK¹ & CLEMENT ATZBERGER¹

Zusammenfassung: In dieser Arbeit wurde die Eignung von Sentinel-2 Daten für die Baumartenunterscheidung analysiert. Dabei wurde auch ein Vergleich mit räumlich höher (WorldView-2) als auch geringer (Landsat-8) aufgelösten Satellitendaten mit ähnlicher spektraler Information durchgeführt. Die Analysen zeigen, dass die spektralen Signaturen der sieben untersuchten Baumartenklassen in allen Datensätzen sehr ähnlich sind. Deutlichere Unterschiede wurden dagegen bei den erzielten Klassifikationsgenauigkeiten festgestellt. So wurden mit den aus den Sentinel-2 Daten gewonnenen Informationen wesentlich höhere Genauigkeiten erzielt als mit Landsat-8. Die noch besseren Ergebnisse mit WorldView-2 belegen einen Zusammenhang zwischen räumlicher Auflösung und Klassifikationsgenauigkeit. Dabei ist jedoch zu berücksichtigen, dass WorldView-2 und andere kommerzielle Erdbeobachtungsdaten nur eingeschränkt verfügbar und mit hohen Kosten verbunden sind. Sentinel-2 Daten sind dagegen global kostenfrei und in hoher zeitlicher Frequenz verfügbar.

1 Einleitung

Der steigende Bedarf an detaillierten Informationen über Ökosysteme erfordert die Entwicklung von objektiven, großflächig anwendbaren Ansätzen, welche ggfls. gut auf andere Gebiete übertragbar sind. Dabei wird zunehmend auf den Einsatz von optischen Erdbeobachtungsdaten gesetzt, deren Verfügbarkeit in den letzten Jahren stark zugenommen hat. Mittlerweile stehen Satellitendaten mit häufiger Wiederholrate, sehr guter spektraler und räumlicher Auflösung zur Verfügung. Mit dem Start von Sentinel-2 (S2) wurde diesbezüglich ein weiterer wesentlicher Schritt gemacht, welcher die Verfügbarkeit an hoch aufgelösten Spektraldaten erhöht. Der erste der beiden S2 Satelliten wurde am 23. Juni 2015, im Rahmen des europäischen Copernicus Programms, erfolgreich gestartet und sendete bereits wenige Tage danach die ersten Szenen. Der Satellit trägt einen innovativen MSI Sensor (Multi-Spectral Imager) welcher Daten in 13 Spektralkanälen mit räumlichen Auflösungen von 10, 20 bzw. 60 m liefert und dabei eine Streifenbreite von 290 km abdeckt (DRUSCH et al. 2012; ESA 2015).

Nach dem Start des Zwillingsatelliten (voraussichtlicher Starttermin im Januar 2017) ergibt sich eine Wiederholrate der Datenaufnahme von maximal fünf Tagen (im Überlappungsbereich der Streifen doppelt so häufig). Durch die Kombination von hoher räumlicher Auflösung von bis zu 10 m, mit innovativen Spektralkanälen, wie drei Red-Edge sowie die beiden SWIR-Kanälen, stehen extrem nützliche Daten zur Verfügung, welche für die Charakterisierung der

¹ Universität für Bodenkultur Wien (BOKU), Institut für Vermessung, Landinformation und Fernerkundung; Peter-Jordan-Str. 82, A-1190 Wien, E-Mail: [Vorname.Nachname]@boku.ac.at

Landbedeckung und dabei vor allem von Vegetation neue Perspektiven eröffnen (MALENOVSKÝ et al. 2012; IMMITZER et al. 2016).

Der zunehmende Informationsbedarf bezüglich Waldgebieten hat in den letzten Jahren bereits zu einem verstärkten Einsatz von Fernerkundungsdaten geführt (WULDER et al. 2004). So sind detaillierte Informationen über die Baumartenzusammensetzung, sowohl für die Bewirtschaftung der Wälder als auch zur ökologischen Bewertung unumgänglich. Dies spiegelt sich auch in der steigenden Anzahl an Studien zur Klassifikation von Baumarten mit den unterschiedlichsten Erdbeobachtungssensoren wieder (NAGENDRA 2001; IMMITZER et al. 2012). Neben einer adäquaten räumlichen Unterscheidung von verschiedenen Klassen und der Beschreibung der aktuellen Situation, steigt aber auch das Interesse an der Erkennung von Veränderungen. Eine hohe zeitliche Wiederholrate, im Idealfall unter vergleichbaren Aufnahmegegebenheiten (Blickrichtung, Uhrzeit, usw.), bietet die Möglichkeit rasche Veränderungen wie z.B. Windwürfe zu detektieren. Aber auch langsamere Vitalitätsveränderungen sowie Wechsel der Baumartenzusammensetzung sind erfassbar. Bedingt durch sich ändernde Umweltbedingungen wird das Interesse an diesen Informationen in den nächsten Jahren noch deutlich zunehmen (ASNER 2013).

Ziel dieser Studie war die Untersuchung einer monotemporalen S2 Szene zur Klassifizierung von Baumarten in Mitteleuropa und ein Vergleich mit bereits etablierten Satellitensensoren, sowohl mit höherer als auch geringerer räumlicher Auflösung. Dazu wurden Referenzpolygone auf den Daten der Sensoren S2, WorldView-2 (WV2) und Landsat-8 (L8) abgegrenzt um damit Random Forest (RF) Klassifikationsmodelle zu erstellen. Dies ermöglicht einen Vergleich der Modellergebnisse in Bezug auf die verfügbare Spektralinformationen sowie die räumliche Auflösung. Da bislang erst wenige wolkenfreie S2 Szenen zur Verfügung stehen, konnte eine multitemporale Analyse im Rahmen dieser Arbeit noch nicht durchgeführt werden.

2 Datenmaterial und Methoden

2.1 Untersuchungsgebiet und Referenzdaten

Das Untersuchungsgebiet liegt im bayerischen Alpenvorland östlich von München (Abb. 1). Die Analysen konzentrieren sich auf die zwei von den Bayerischen Staatsforsten (BaySF) bewirtschaftete Waldgebiete Ebersberg und Altötting von welchen WV2 Sommeraufnahmen sowie Referenzdaten zur Verfügung standen. Die analysierten Waldgebiete sind geprägt durch heterogene Mischbestände mit nur geringen Anteilen von Reinbeständen. Das Gebiet wird durch eine Jahresdurchschnittstemperatur von 7.6°C und einem mittleren jährlichen Niederschlag zwischen 850 und 950 mm, mit einem Niederschlagsmaximum in den Sommermonaten, charakterisiert. Die potenziellen natürlichen Waldgesellschaften sind Buchen-dominierte Wälder mit beigemischter Fichte bzw. Tanne. Die aktuelle Bestockung wird von Fichte (*Picea abies*, (L) Karst.) dominiert, wobei ebenso Rot-Buche (*Fagus sylvatica*, L.) und Waldkiefer (*Pinus sylvestris*, L.) mit Anteilen von bis zu 10%, sowie Eichenarten (*Quercus* sp.) mit rund 5% Flächenanteil, eine wichtige Rolle spielen. Im Untersuchungsgebiet kommen darüber hinaus Tanne (*Abies alba*, Mill.), Europäische Lärche (*Larix decidua*, Mill.), Douglasie (*Pseudotsuga menziesii*, (Mirb.) Franco), Hainbuche (*Carpinus betulus*, L.), sowie Ahorn- (*Acer* sp.), Birken- (*Betula* sp.), Erlen- (*Alnus* sp.) und Weidenarten (*Salix* sp.) vor (BAYSF 2013).

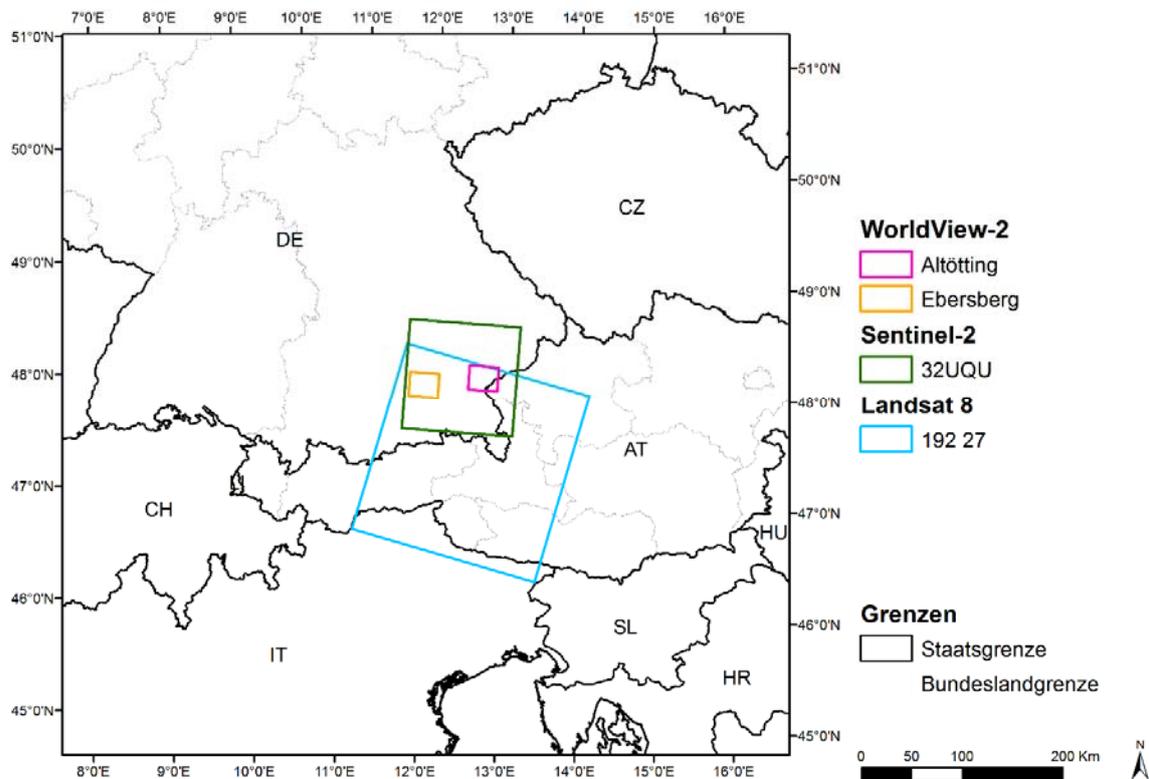


Abb. 1: Lage des Untersuchungsgebietes und Abdeckung der für die Analysen verwendeten Fernerkundungsszenen.

Die Referenzinformationen stammten aus Forstinventuren und wurden mittels visueller Interpretation von Orthophotos und WV2 Szenen ergänzt. Auf Grund der geringen Anzahl an Douglasie-Referenzdaten konnte diese Baumart in den Analysen nicht berücksichtigt werden. Ebenso mussten die wenigen Referenzdaten von Hainbuche, Ahorn, Birke, Erle und Weide zur Klasse „sonstiges Laubholz“ zusammengefügt werden. Damit wurden insgesamt sieben Klassen unterschieden (drei Laubholzklassen und vier Nadelholzklassen).

2.2 Satellitendaten und Vorverarbeitung

Der WV2 Datensatz erstreckt sich auf zwei Szenen, aufgenommen am 12. Juli 2011 (Ebersberg) bzw. am 16. August 2013 (Altötting). Die Szenen wurden mit der Prozesskette CATENA (REINARTZ 2010; KRAUB et al. 2013) des DLR vorprozessiert. CATENA beinhaltet eine parametrischen Orthorektifizierung basierend auf ORTHO (MÜLLER et al. 2005) und eine automatische Atmosphärenkorrektur basierend auf ATCOR (RICHTER & SCHLÄPFER 2012). Um einen möglichst homogenen Datensatz zu erhalten wurde zusätzlich ein Histogramm-Matching angewendet.

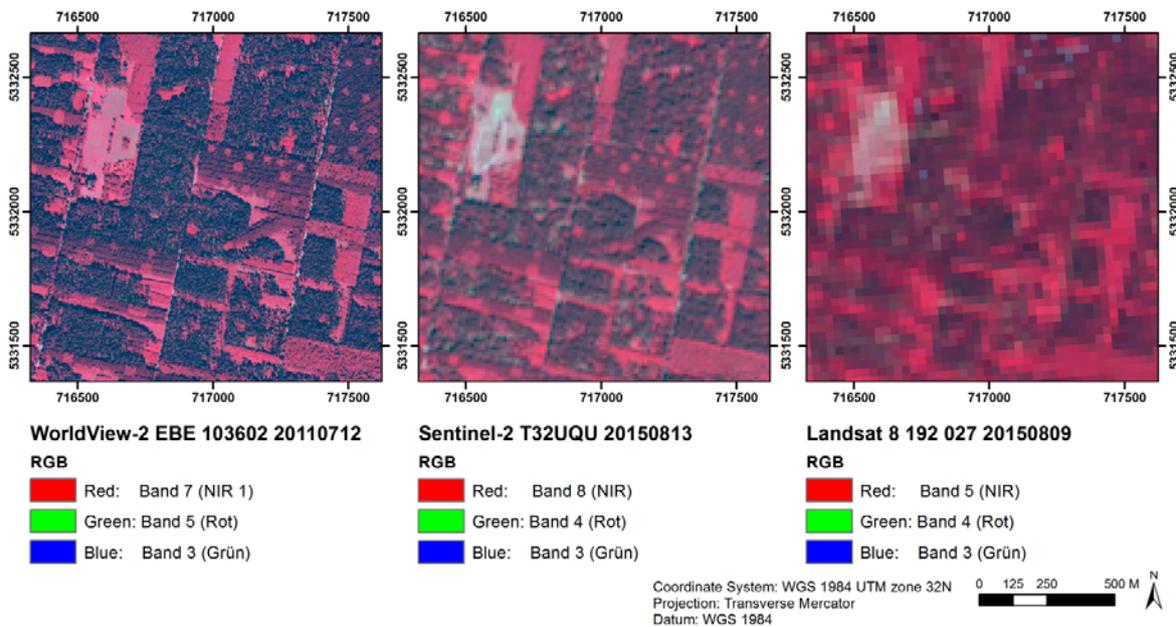


Abb. 2: Vergleich der drei verwendeten Satellitendaten WorldView-2, Sentinel-2 und Landsat-8 jeweils in CIR Darstellung

Der verwendete S2 Datensatz wurde am 13. August 2015 aufgenommen und war somit eine der ersten wolkenfreien S2 Aufnahmen. Für die Analysen wurde der am 8. März 2016 von der ESA neu prozessierte Datensatz verwendet. Die Atmosphärenkorrektur erfolgte mit Sen2Cor. Für die Analysen wurden die sechs 20 m Spektralbänder auf 10 m umgerechnet und mit den originalen 10 m-Bänder zusammengefügt. Die drei S2 Bänder in 60 m Auflösung wurden nicht verwendet. Die Aufnahme der L8 Szene erfolgte am 9. August 2015. Es konnten vom USGS bereits atmosphärenkorrigierten Daten (surface reflectance, L8SR) heruntergeladen werden. Die unterschiedlichen räumlichen Auflösungen der drei Satellitensysteme sind in Abb. 2 dargestellt. Mittels automatisierter Segmentierung wurden homogene Bestandteile bzw. Baumgruppen abgegrenzt. Die WV2 Szenen wurden mit eCognition segmentiert (BAATZ & SCHÄPE 2000) (*scale* = 25, *shape* = 0.9 und *compactness* = 0.95). Bei der Segmentierung wurden die Kanäle Rot und NIR mit 1,5 gewichtet. Anschließend wurden aneinander angrenzende ähnliche Segmente (mittels *merge* = 20) nochmals zusammengefügt. Weitere Details zu der Szenenvorverarbeitung und Segmentierung finden sich in IMMITZER et al. (2015).

Für die S2 Szene wurde eine Large Scale Mean Shift (LSMS)-Segmentierung (COMANICIU & MEER 2002) basierend auf den 10 m Bändern (Band 2, 3, 4 und 8) durchgeführt. Details zur Datenaufbereitung, zum Segmentierungsalgorithmus und zur Parameterfestlegung finden sich in IMMITZER ET AL. (2016). Umgesetzt wurde die Segmentierung mit der open source software Orfeo Toolbox Version 5.0.0, eine Beschreibung zu dieser findet sich bei MICHEL et al. (2015). Die verwendeten Parameter waren: *spatial radius* (hs) = 10, *range radius* (hr) = 10 und *minimum size* (ms) = 4.

Zur Segmentierung der L8 Daten wurde die S2 Segmentierung auf die L8 Szene übertragen, wobei für die Berechnung der Metriken, die jeweilige Deckung des Polygons mit den (gröberen) Pixeln berücksichtigt wurde.

Für alle drei Datensätze und Segmentierungen wurde für jedes Polygon der Mittelwert, die Standardabweichung sowie die 5., 10., 25., 50., 75., 90. und 95. Perzentilen berechnet. Diese Metriken wurden für die RF Klassifikation verwendet.

2.3 Klassifikation und Genauigkeitsanalyse

Basierend auf den Metriken der Referenzpolygone wurden für jeden Datensatz RF Klassifikationsmodelle erstellt (BREIMAN 2001). RF ist ein auf einer Vielzahl von Entscheidungsbäumen basierender Algorithmus. Durch das integrierte Bootstrapping basieren auch die Modellergebnisse auf unabhängige Schätzung: es werden nur Samples klassifiziert die in dem jeweiligen Durchgang nicht gezogen (out-of-bag) und somit nicht beim Training des Entscheidungsbaumes berücksichtigt wurden. Detaillierte Informationen zu RF für die Klassifikation von Fernerkundungsdaten finden sich in der Literatur (PAL 2005; GISLASON et al. 2006; IMMITZER et al. 2012). Für die beiden einzustellende Parameter *n*tree (Anzahl an Entscheidungsbäume) und *m*try (Anzahl an zufällig gewählten Metriken/Features für jeden Entscheidungsknoten) wurden die Standardeinstellungen verwendet: 500 Bäume (*n*tree) und die Wurzel der insgesamt verfügbaren Features (*m*try). Die Umsetzung erfolgte in R 3.2.3 (R CORE TEAM 2015) mit dem Package *randomforest* (LIAW & WIENER 2002).

Die Bewertung der Modelle erfolgte mittels Klassifikationsmatrizen und den daraus abgeleiteten Standardmaßen: Nutzer-Genauigkeit (NG), Produzent-Genauigkeit (PG), Gesamtgenauigkeit (GG) sowie Kappa (FOODY 2002).

3 Ergebnisse und Diskussion

3.1 Spektrale Signaturen

Die extrahierten spektralen Signaturen für die verschiedenen Baumarten zeigen eine sehr hohe Übereinstimmung zwischen den drei Satellitensystemen (Abb. 3). Während im sichtbaren Bereich die Unterschiede relativ gering sind, sind die Baumarten vor allem in Wellenlängenbereich des nahen Infrarots gut unterscheidbar. Die Laubbaumarten weisen dabei naturgemäß deutlich höhere Reflexionswerte auf als die Nadelbäume. Die höchsten Reflexionswerte zeigte die Buche, die geringsten Fichte und Kiefer. Dies deckt sich mit anderen Studien so zeigten die Analysen von Einzelkronen dieselben Tendenzen der Baumarten, jedoch etwas bessere Unterscheidbarkeit zwischen den einzelnen Baumarten (IMMITZER et al. 2012).

Durch die zusätzlichen Kanäle im kurzwelligen Infrarot weisen die beiden Sensoren S2 und L8 einen wesentlichen Vorteil gegenüber den räumlich höher aufgelösten WV2 Sensor auf. In diesen SWIR-Kanälen sind ähnlich wie im NIR große Unterschiede zwischen den Baumarten erkennbar (dies wurde beim Nachfolger, dem im Jahr 2015 gestarteten WorldView-3, berücksichtigt). WV2 und S2 haben neben den konventionellen Bändern auch einen bzw. drei Spektralkanäle im Red-Edge Bereich was ebenfalls für Vegetationsstudien von Vorteil ist. Sowohl die Wichtigkeit von SWIR als auch jene der Red-Edge Bänder wurde bereits in mehreren Studien belegt (SCHUSTER et al. 2012; RAMOELO et al. 2015; IMMITZER et al. 2016).

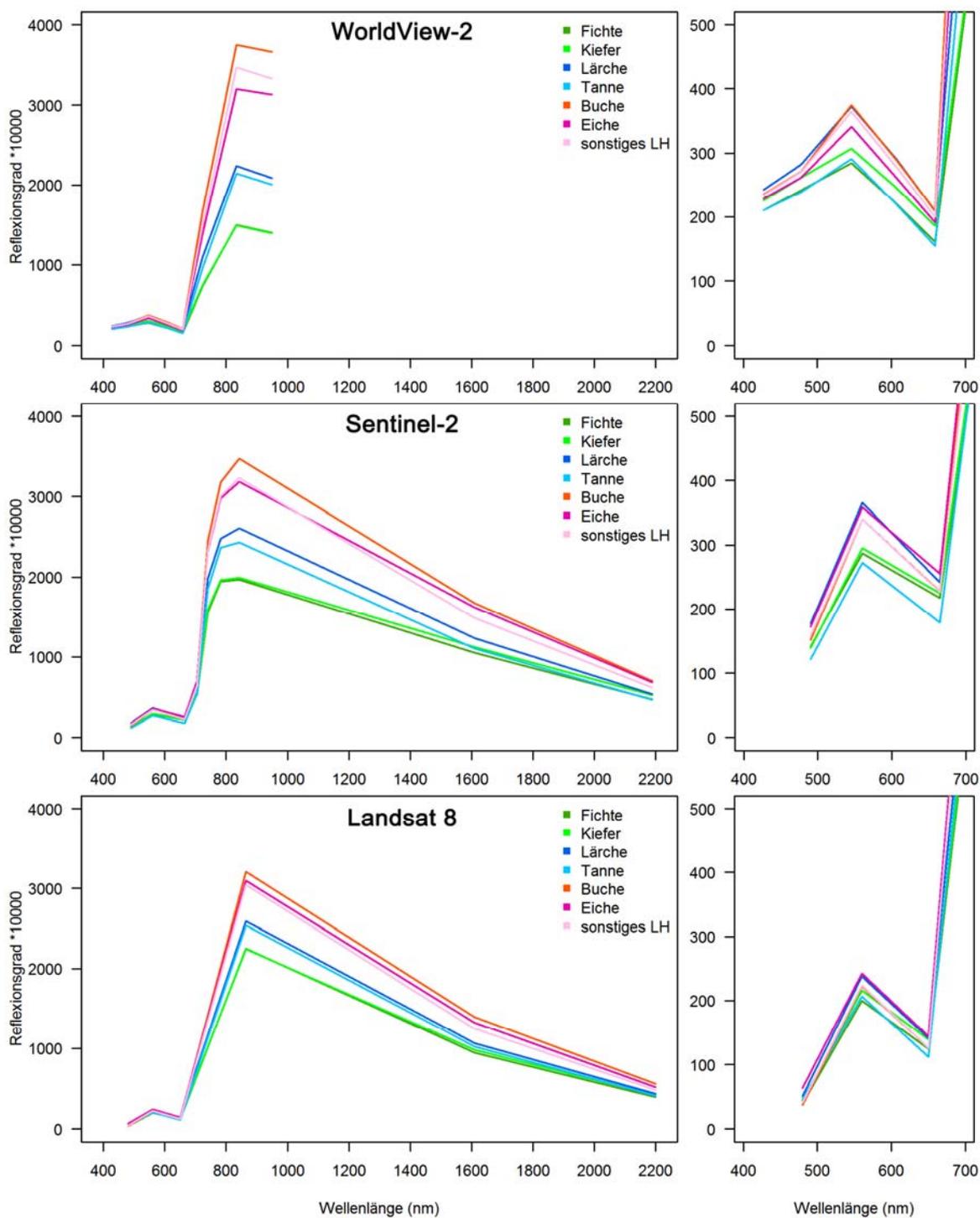


Abb. 3: Spektrale Signaturen für die sieben untersuchten Baumartenklassen und die drei analysierten Satellitensensoren. Dargestellt sind die Reflexions-Mittelwerte der jeweiligen Spektralbänder aus den Referenzobjekten

3.2 Klassifikationsergebnisse

Die Klassifikationsmodelle für sieben Baumarten erreichten mit Gesamtgenauigkeiten von 50 bis 75% (Tab. 1). Die Ergebnisse für S2 und WV2 liegen damit im Bereich vergleichbarer Studien mit anderen Datensätzen (siehe Studienübersicht in IMMITZER et al. (2012)). Aggregiert man die Klassifikationsergebnisse der einzelnen Baumarten zu den Baumartengruppen ‚Nadelhölzer‘ und ‚Laubhölzer‘ zeigt sich folgendes Bild: mit WV2 können die beiden Gruppen mit einer Genauigkeit von 97% getrennt werden, bei S2 mit 93% und bei L8 mit 83%.

Tab. 1: Klassifikationsmatrizen der RF Modelle für die untersuchten Datensätze WorldView-2, Sentinel-2 und Landsat-8 (soLH: sonstige Laubhölzer; NG: Nutzergenauigkeit, PG: Produzentengenauigkeit, GG: Gesamtgenauigkeit)

WorldView-2		Referenz									
		Fichte	Kiefer	Lärche	Tanne	Buche	Eiche	soLH	Σ	NG	
Klassifiziert als	Fichte	121	5	5	6	0	0	3	140	0.864	
	Kiefer	2	42	3	0	0	0	0	47	0.894	
	Lärche	4	3	48	1	0	1	2	59	0.814	
	Tanne	1	0	0	22	0	2	2	27	0.815	
	Buche	0	0	0	0	73	11	26	110	0.664	
	Eiche	0	0	1	3	5	33	14	56	0.589	
	soLH	0	0	0	1	15	12	33	61	0.541	
PG		128	50	57	33	93	59	80	500	0.744	
									GG		
									Kappa		0.691

Sentinel-2		Referenz									
		Fichte	Kiefer	Lärche	Tanne	Buche	Eiche	soLH	Σ	NG	
Klassifiziert als	Fichte	70	14	3	3	0	0	0	90	0.778	
	Kiefer	2	5	0	0	0	0	0	7	0.714	
	Lärche	1	1	28	1	0	3	4	38	0.737	
	Tanne	2	0	3	19	2	0	2	28	0.679	
	Buche	0	0	0	2	48	10	8	68	0.706	
	Eiche	1	1	2	0	1	9	6	20	0.45	
	soLH	1	0	2	1	15	10	37	66	0.561	
PG		77	21	38	26	66	32	57	317	0.681	
									GG		
									Kappa		0.611

Landsat-8		Referenz									
		Fichte	Kiefer	Lärche	Tanne	Buche	Eiche	soLH	Σ	NG	
Klassifiziert als	Fichte	50	13	5	9	3	4	5	89	0.562	
	Kiefer	6	4	0	1	0	0	0	11	0.364	
	Lärche	6	3	22	4	2	2	3	42	0.524	
	Tanne	3	0	1	4	1	0	2	11	0.364	
	Buche	6	1	2	4	41	12	12	78	0.526	
	Eiche	3	0	2	1	5	6	6	23	0.261	
	soLH	3	0	6	3	14	7	29	62	0.468	
PG		77	21	38	26	66	31	57	316	0.494	
									GG		
									Kappa		0.380

Die Modellergebnisse zeigen einen klaren Zusammenhang zwischen Klassifikationsgenauigkeit und räumlicher Auflösung der zugrundeliegenden Daten. Die höchste Gesamtgenauigkeit, wie auch die höchsten klassenspezifischen Genauigkeiten, wurde mit den WV2 Daten erzielt. Die etwas höhere Anzahl an Referenzdaten (kleinere Segmente als bei S2) wird dabei einen geringen positiven Einfluss haben; die Verwendung von zwei unterschiedlichen Szenen einen negativen Einfluss. Die mit S2 erzielten Ergebnisse sind deutlich höher als jene mit L8, was auf die bessere räumliche und spektrale Auflösung zurückzuführen ist.

Während bei WV2 die klassenspezifischen Ergebnisse relativ ausgeglichen sind, sind bei S2 und L8 große Unterschiede zwischen einzelnen Klassen erkennbar, wobei die Fichte in allen Datensätzen am besten klassifiziert wird. Die Kiefer weißt bei WV2 nur geringfügig schlechtere Ergebnisse auf, bei S2 und L8 ist hingegen die Produzentengenauigkeiten deutlich geringer als bei den anderen Baumarten. Ähnliches gilt auch für Eiche. Dabei wird die Tatsache, dass diese beiden Baumarten im Vergleich zu Fichte bzw. Buche in den untersuchten Gebieten weniger häufig vorkommen, vor allem auch was Reinbestände betrifft, eine Rolle spielen. Dadurch ist der Anteil an Mischpixel bzw. nicht baumartenreinen Segmenten bei den Datensätzen mit geringerer räumlicher Auflösung höher, was sicherlich zur Verschlechterung der Ergebnisse beiträgt. Neben der geringeren Mischpixelproblematik bietet die höhere räumliche Auflösung von WV2 auch den Vorteil, Texturunterschiede besser abzubilden. Dies erhöht zum Beispiel die Trennbarkeit von Fichte und Kiefer mit stark unterschiedlichen Kronenformen.

Verglichen mit den ersten (prä-operationellen) Daten die von der ESA im Herbst 2015 zu Testzwecken zur Verfügung gestellt wurden, sind die jetzt operationell verfügbaren Daten noch einmal qualitativ verbessert worden. So konnten in dieser Arbeit, im Vergleich zu einer Vorläuferstudie, die noch mit den prä-operationellen Daten gearbeitet hat (IMMITZER et al. 2016), eine Verbesserung in den spektralen Eigenschaften als auch in der erzielten Klassifikationsgenauigkeit beobachtet werden. Dies zeigt die Anstrengungen, die von Seiten der ESA unternommen werden, um den maximalen Informationsgehalt der aufgezeichneten Daten verfügbar zu machen.

Insbesondere sind die sechs S2 Bänder mit einer räumlichen Auflösung von 20 m sind detaillierte Analysen in sehr strukturreichen Wäldern zu grob, was sich mit den Erkenntnissen von STRATOULIAS et al. (2015) deckt. Bei einer Auflösung von 20 m ist bereits die Referenzdatenabgrenzung sehr schwierig. Oftmals reichen auch Aufnahmen mit 10 m Auflösung nicht aus, um einzelne Bäume zu erkennen bzw. zu analysieren (WULDER et al. 2004).

Für sehr detaillierte Auswertungen könnten Kombinationen von S2 mit höher aufgelösten Datensätzen wie Orthophotos eine potenzielle Lösung darstellen. Darüber hinaus sind vor allem S2 Zeitreihen von großem Interesse. Durch die Verwendung von nur einer einzelnen Szene wurde in dieser Studie nicht das volle Potenzial von S2 ausgeschöpft. So können Aufnahmen zu anderen Zeitpunkten, bzw. vor allem die Verwendung von kontinuierliche Zeitreihendaten, die Klassifikation noch wesentlich verbessern. (ELATAWNEH et al. 2013; TIGGES et al. 2013; LI et al. 2015). Die gute Vergleichbarkeit der Spektralkanäle ermöglicht auch eine Kombination von S2 und L8 Daten (WULDER et al. 2015). Durch die hohen Wiederholraten von 5 bzw. 16 Tagen wird sich die Anzahl an wolkenfreien Datensätzen deutlich erhöhen. Zusätzlich werden Verfahren zur Verfügung stehen, hochfrequente, wolkenfreie Komposite zu generieren (VUOLO et al. under review).

4 Fazit & Ausblick

Mit den europäischen Sentinel-2 (S2) Satelliten stehen der Fernerkundungsgemeinschaft seit Ende 2015 Daten mit bislang nicht verfügbaren Charakteristika zur Verfügung. Die hohe räumliche Auflösung (bis zu 10 m), die zehn Spektralkanäle zur Erdbeobachtung (sowie drei weitere Kanäle zur Atmosphärenkorrektur) und v.a. auch die zeitliche Auflösung von zukünftig maximal 5 Tagen bei globaler Abdeckung sind einzigartig. Die spektrale Information von S2 ähnelt dabei sehr stark dem räumlich höher aufgelösten WorldView-2 (WV2) als auch Landsat-8 (L8) mit geringerer räumlicher Auflösung. Die Analyse von monotemporalen S2 Daten für die Klassifikation von sieben Baumartenklassen zeigte gute Ergebnisse. So wurden mit S2 deutlich höhere Genauigkeiten erzielt als mit L8. Mit WV2 konnten zwar noch bessere Ergebnisse erzielt werden, dieser Sensor ist jedoch nur kleinflächig verfügbar und mit hohen Kosten verbunden. Die Analysen bestätigen damit die gute Eignung von S2 Daten für die Unterscheidung von Landbedeckungsklassen, wobei das große Potenzial von S2 – die Verfügbarkeit von multitemporalen Datensätzen – in dieser Arbeit noch nicht untersucht werden konnte.

Die Verwendung von S2 Zeitserien wird weitere Einsatzmöglichkeiten aufzeigen und die Klassifikationsgenauigkeiten weiter verbessern. Neben der Verwendung von multitemporalen Datensätzen wird auch ein großes Potenzial in der Kombination mit anderen z.B. räumlich höheraufgelösten Daten wie Orthophotos gesehen. Solche Fragestellungen werden ein wichtiger Schwerpunkt zukünftiger Forschungsarbeiten sein, wie auch die Nutzung von physikalisch basierten Strahlungstransfermodellen zur Ableitung von biophysikalischen Vegetations-Charakteristika.

Es wurde eine deutliche Qualitätsverbesserung der S2 Daten im Vergleich zu bereits analysierten prä-operationellen Daten beobachtet. Dies belegt die aktuellen Aktivitäten von Seiten der ESA, im Zuge der ramp-up Phase noch Optimierungen an den Daten vorzunehmen, um in Zukunft die bestmöglichen Daten verfügbar zu machen. Von Anwenderseite ist es jetzt ein vorrangiges Ziel, dass sich ESA und die Mitgliedsstaaten auf eine operationelle Bereitstellung von atmosphärenkorrigierten Daten unter Verwendung der temporalen Information (sog. MACCS Algorithmus) einigen.

5 Literaturverzeichnis

- ASNER, G.P., 2013: Geography of forest disturbance. *Proceedings of the National Academy of Sciences* **110** (10), 3711-3712.
- BAATZ, M. & SCHÄPE, A., 2000: Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. Wichmann Verlag, Karlsruhe, 12-23.
- BAYSF, 2013: Regionales Naturschutzkonzept für den Forstbetrieb Wasserburg am Inn. Bayerische Staatsforsten Forstbetrieb Wasserburg, http://www.baysf.de/fileadmin/user_upload/01-ueber_uns/05standorte/FB_Wasserburg_a._Inn/Naturschutzkonzept_Wasserburg.pdf [letzter Zugriff 15. Dezember 2015].
- BREIMAN, L., 2001: Random forests. *Machine learning*, **45** (1), 5-32.
- COMANICIU, D. & MEER, P., 2002: Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24** (5), 603-619.

- DRUSCH, M., DEL BELLO, U., CARLIER, S., COLIN, O., FERNANDEZ, V., GASCON, F., HOERSCH, B., ISOLA, C., LABERINTI, P., MARTIMORT, P., MEYGRET, A., SPOTO, F., SY, O., MARCHESI, F. & BARGELLINI, P., 2012: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment* **120**, 25-36.
- ELATAWNEH, A., RAPPL, A., REHUSH, N., SCHNEIDER, T. & KNOKE, T., 2013: Forest tree species identification using phenological stages and RapidEye data: a case study in the forest of Freising. 5. RESA Workshop From the Basics to the Service DLR e.V., Neustrelitz, 21-38.
- ESA, 2015: Sentinel-2 delivers first images. , Available at: http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/Sentinel-2_delivers_first_images [letzter Zugriff 7. Januar 2016].
- FOODY, G.M., 2002: Status of land cover classification accuracy assessment. *Remote Sensing of Environment* **80**, 185-201.
- GISLASON, P.O., BENEDIKTSSON, J.A. & SVEINSSON, J.R., 2006: Random Forests for land cover classification. *Pattern Recognition Letters* **27** (4), 294-300.
- IMMITZER, M., ATZBERGER, C. & KOUKAL, T., 2012: Tree species classification with Random Forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sensing* **4** (9), 2661-2693.
- IMMITZER, M., EINZMANN, K., BÖCK, J., MATTIUZZI, M., NG, W.-T., WALLNER, A., PINNEL, N., REICHMUTH, A., FROST, M., MÜLLER, A., SEITZ, R. & ATZBERGER, C., 2015: Erstellung von Fichten- und Kiefernanteilskarten auf Basis von Satellitendaten für Bayern. *Forstliche Forschungsberichte München* (214), 21-34.
- IMMITZER, M., VUOLO, F. & ATZBERGER, C., 2016: First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sensing* **8** (3), 1-27.
- KRAUB, T., D'ANGELO, P., SCHNEIDER, M. & GSTAIGER, V., 2013: The Fully Automatic Optical Processing System CATENA at DLR. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (1/W1), 177-183.
- LI, D., KE, Y., GONG, H. & LI, X., 2015: Object-Based Urban Tree Species Classification Using Bi-Temporal WorldView-2 and WorldView-3 Images. *Remote Sensing* **7** (12), 16917-16937.
- LIAW, A. & WIENER, M., 2002: Classification and regression by randomForest. *R news* **2** (3), 18-22.
- MALENOVSKÝ, Z., ROTT, H., CIHLAR, J., SCHAEPMAN, M.E., GARCÍA-SANTOS, G., FERNANDES, R. & BERGER, M., 2012: Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment* **120**, 91-101.
- MICHEL, J., YOUSSEFI, D. & GRIZONNET, M., 2015: Stable Mean-Shift Algorithm and Its Application to the Segmentation of Arbitrarily Large Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing* **53** (2), 952-964.
- MÜLLER, R., HOLZWARTH, S., HABERMAYER, M. & MÜLLER, A., 2005: Ortho Image Production within an Automatic Processing Chain for hyperspectral Airborne Scanner ARES. *EARSeL Workshop 3D-Remote Sensing*, Porto, Portugal.

- NAGENDRA, H., 2001: Using remote sensing to assess biodiversity. *International Journal of Remote Sensing* **22** (12), 2377-2400.
- PAL, M., 2005: Random forest classifier for remote sensing classification. *International Journal of Remote Sensing* **26**, 217-222.
- R CORE TEAM, 2015: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- RAMOELO, A., CHO, M., MATHIEU, R. & SKIDMORE, A.K., 2015: Potential of Sentinel-2 spectral configuration to assess rangeland quality. *Journal of Applied Remote Sensing* **9** (1).
- REINARTZ, P., 2010: The CATENA Processing Chain-Multi-Sensor Pre-processing: Orthorectification, Atmospheric Correction, Future Aspects. *Proceedings of Geoland Forum_6*, Toulouse.
- RICHTER, R. & SCHLÄPFER, D., 2012: Atmospheric/Topographic Correction for Satellite Imagery (ATCOR-2/3 User Guide). German Aerospace Center.
- SCHUSTER, C., FÖRSTER, M. & KLEINSCHMIT, B., 2012: Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. *International Journal of Remote Sensing* **33** (17), 5583-5599.
- STRATOULIAS, D., BALZTER, H., SYKIOTI, O., ZLINSZKY, A. & TÓTH, V.R., 2015: Evaluating sentinel-2 for lakeshore habitat mapping based on airborne hyperspectral data. *Sensors* **15** (9), 22956-22969.
- TIGGES, J., LAKES, T. & HOSTERT, P., 2013: Urban vegetation classification: Benefits of multitemporal RapidEye satellite data. *Remote Sensing of Environment* **136**, 66-75.
- VUOLO, F., NG, W.-T. & ATZBERGER, C., under review: Innovative approach for smoothing and gap-filling of high resolution multi-spectral time series: Example of Landsat data.
- WULDER, M.A., HALL, R.J., COOPS, N.C. & FRANKLIN, S.E., 2004: High spatial resolution remotely sensed data for ecosystem characterization. *BioScience* **54** (6), 511-521.
- WULDER, M.A., HILKER, T., WHITE, J.C., COOPS, N.C., MASEK, J.G., PFLUGMACHER, D. & CREVIER, Y., 2015: Virtual constellations for global terrestrial monitoring. *Remote Sensing of Environment* **170**, 62-76.

Mathematikunterricht aus dem All – Interdisziplinäre Lernwerkzeuge für den Einsatz von Erdbeobachtung im Schulunterricht

ANDREAS RIENOW¹, VALERIE GRAW¹, SASCHA HEINEMANN¹,
JOHANNES SCHULTZ¹, FABIAN SELG¹ & GUNTER MENZ¹

Zusammenfassung: Die Erdbeobachtung mit Bildern aus dem All ermöglicht einen umfangreichen Einblick auf natürliche und vom Mensch beeinflusste Phänomene auf der Erde. Trotz ihrer visuellen Ästhetik und ihres wissenschaftlichen Nutzens finden Erdbeobachtungsdaten aber noch selten einen reflektierten Eingang in den alltäglichen Schulunterricht. Hier knüpft das vom Deutschen Zentrum für Luft- und Raumfahrt (DLR) geförderte Projekt „Columbus Eye – Live-Bilder von der ISS im Schulunterricht“ (www.columbuseye.uni-bonn.de) der Arbeitsgruppe Fernerkundung an der Universität Bonn an. Ziel des didaktischen Projekts ist es die Erdbeobachtung aus dem All im Allgemeinen und von der Internationalen Raumstation (ISS) im Speziellen nachhaltig und umfassend in den Schulunterricht zu integrieren. Inhaltlich werden insbesondere die Fächer aus dem MINT-Bereich abgedeckt, welches neben dem Schulfach zur Geographie auch das Fach Mathematik mit einbezieht, in welchen die naturwissenschaftlichen Hintergründe der Aufnahmen nahe am Schulcurriculum eingesetzt werden können. Das Lernmodul „Mittelwertberechnung von der ISS“ als ein Beispiel für das Fach Mathematik beschäftigt sich mit der Korrektur fehlerhafter Fernerkundungsdaten. Schülerinnen und Schüler lernen hier wie Empfangsstörungen mit Hilfe mathematischer Berechnungen geglättet werden, um die in den Bildern enthaltenen Informationen brauchbar zu machen. Durch den Dreiklang aus ‚Sehen‘, ‚Informieren‘ und ‚Interagieren‘ wird ein Grundstein in die Richtung des eigenverantwortlichen Umgangs von Schülerinnen und Schülern mit Geomedien im Bereich der Luft- und Raumfahrt gelegt.

1 Erdbeobachtung von der ISS

Die Internationale Raumstation (ISS) ist das größte künstliche Objekt im Orbit und das teuerste wissenschaftliche Labor der Menschheit. Während die ISS vor allen Dingen für Experimente zur Mikrogravitation bekannt ist, tritt ihr Nutzen als Erdbeobachtungsplattform noch weitgehend in den Hintergrund (STEFANOV & EVANS 2014). Ins Bewusstsein einer breiten Öffentlichkeit gelangt die Erdbeobachtung von der ISS vordergründig durch die Präsenz von Astronauten-Aufnahmen unserer Erde in den sozialen Medien. Hier teilen die Astronauten ihren Blick von der ISS auf unsere Erde mit der Öffentlichkeit. Das Projekt „Columbus Eye – Live-Bilder von der ISS im Schulunterricht“, gefördert vom Deutschen Zentrum für Luft- und Raumfahrt (DLR), partizipiert als einziger europäischer Partner am NASA-Experiment „High Definition Earth Viewing“ (HDEV) (NASA 2014). Das Erdbeobachtungsexperiment umfasst vier commercial off-the shelf (COTS) Kameras, welche im April 2014 am Columbus Labor der ISS montiert und in Betrieb genommen worden sind. Die COTS-Kameras decken im stündlichen Turnus drei

¹ Geographisches Institut der Universität Bonn, Meckenheimer Allee 166, D-53115 Bonn,
E-Mail: [a.rienow@geographie.uni-bonn.de]

Perspektiven ab (Abb. 1). Im Rahmen von HDEV untersucht die NASA, ob die COTS-Kameras für kommende Weltraummissionen, wie beispielsweise zum Mond oder zum Mars, geeignet sein könnten (RIENOW et al. 2015). Die Kameras filmen die Erde 24 Stunden am Tag und sieben Tage die Woche aus unterschiedlichen Perspektiven.

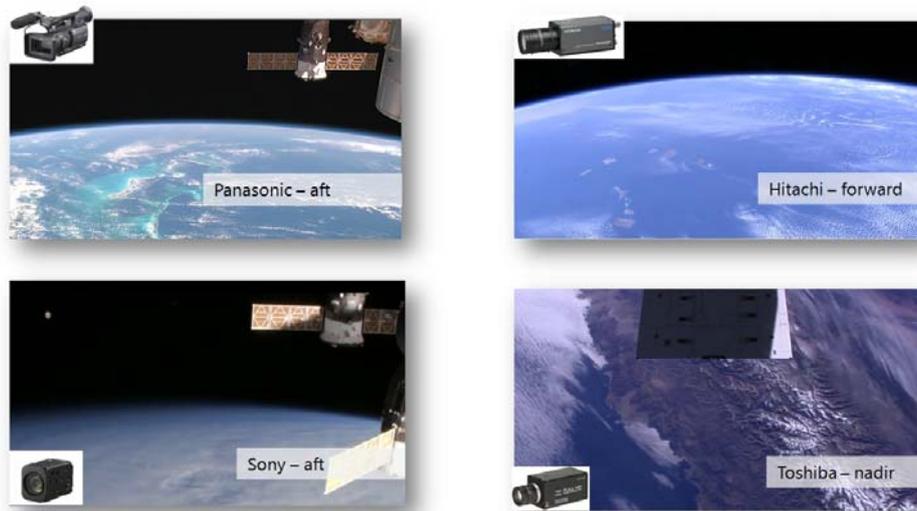


Abb. 1: Die vier am Columbus Labor der ISS montierten Kameras und ihre Blickwinkel (Quelle: NASA).

Das wesentliche Ziel von Columbus Eye wiederum ist die nachhaltige Integration der ISS-Bilddaten in den naturwissenschaftlichen Schulunterricht. Als Partner von HDEV ist das Projekt somit auch am Empfang und an der Veröffentlichung des Videomaterials beteiligt. Hierzu wird das elektronische Signal der HDEV-Payload von der ISS über das Tracking and Data Relay Satellite System nach White Sands (New Mexico, USA) gesendet und von dort aus zum Payload Operations Integration Center (POIC) am Marshall Space Flight Centre in Huntsville (Alabama, USA) weitergeleitet. Die weitere Übertragung und finale Speicherung in Bonn (Deutschland) erfolgt mit Hilfe des Telescience Resource Kit (TReK) (SCHNEIDER & LAPENTA 2001). Innerhalb von Columbus Eye werden die Aufnahmen der vier HD-Kameras archiviert und radiometrisch korrigiert (Abb. 4). Die Archivierung der Aufnahmen erfolgt ausschließlich in Bonn. Seit Februar 2016 können die Kameras zudem auch von der Universität Bonn über das POIC angesteuert werden, um die aktuelle Kameraperspektive zu verändern. Dies bietet die Möglichkeit, die für erdbeobachtungszwecke so wichtige Nadir-Sicht für eine längere Zeit beizubehalten. Mit einer Bildauflösung von 2,1 Megapixel in Kombination mit der Flughöhe der ISS, welche bei ca. 400 km liegt, ergibt sich im Nadir eine räumliche Bodenauflösung von ca. 500 m.

Die Verbreitung der HDEV-Aufnahmen und der darauf basierenden Unterrichtsmaterialien geschieht über das projekteigene Internetportal: www.columbuseye.uni-bonn.de. Neben einem Livestream von der ISS, werden hier interessante Sequenzen aus den Überflügen als ‚Highlights‘ veröffentlicht. Hierzu gehören u. a. natürlicher Phänomene wie Vulkanausbrüche, Vorgänge rund um die ISS wie das Andocken einer neuen Versorgungskapsel, sowie faszinierende Muster der Erdoberfläche wie Meteoritenkrater. Die wichtigste Komponente im Portal stellt jedoch die Unterrichtsmaterialien dar. Fernerkundungsdaten können durch ihre hohe Anschaulichkeit einen

wichtigen Beitrag zur medienbasierten Wissensvermittlung von Themen des Schulcurriculums leisten (GOETZKE et al. 2014). Hierbei kann nicht nur der Geographieunterricht durch die Erdbeobachtung innovativ gestaltet werden. Auch andere MINT-Fächer – der Begriff leitet sich aus den Fachbereichen Mathematik, Information, Naturwissenschaft und Technik ab – können von Fernerkundungsdaten und den einhergehenden Perspektivenwechsel der Schülerinnen und Schüler (SuS) profitieren.

Das in diesem Artikel vorgestellte Beispiel beschäftigt sich exemplarisch mit dem Einsatz von Bildern von der ISS im Mathematikunterricht. Mit Hilfe des digitalen Lernmoduls „Mittelwertberechnung von der ISS“ werden theoretische Konzepte aus der Stochastik praktisch umgesetzt und für die Glättung und Filterung von Bildern von der ISS verwendet. Auf diese Weise wird nicht nur Wissen aus dem Schulcurriculum vermittelt, sondern auch das eigenständige Arbeiten mit wissenschaftlichen Methoden gefördert.

2 HDEV im Schulunterricht

Die auf Basis der HDEV-Daten entwickelten Unterrichtsmaterialien können anhand von drei Interaktionsstufen unterschieden werden: einfache Arbeitsblätter, interaktive Lerntools und umfassende digitalen Lernumgebungen (Abb.2).

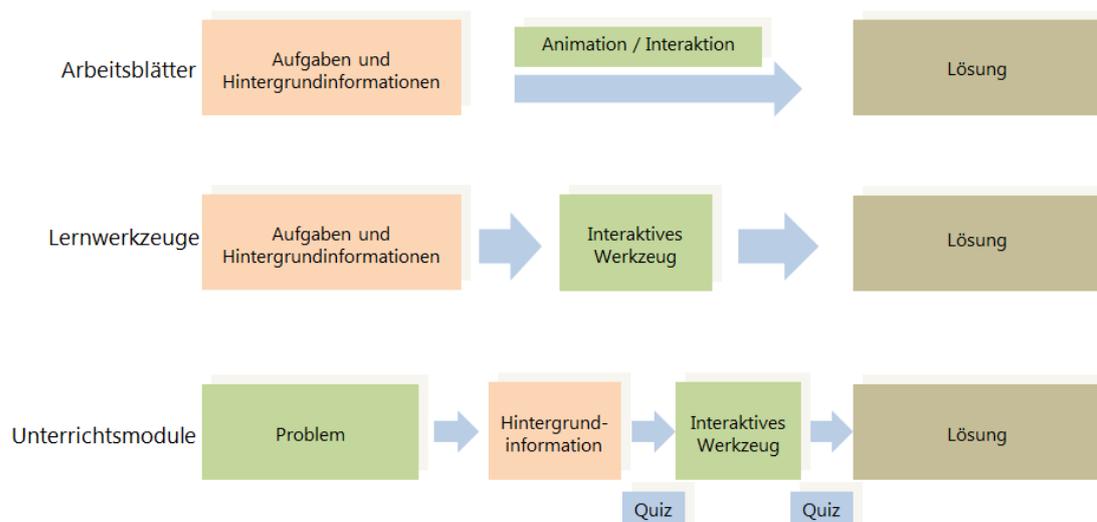


Abb. 2: Die drei Ebenen von Interaktionsgraden der Columbus Eye Unterrichtsmaterialien

Die Arbeitsblätter umfassen u. a. Themengebiete wie die Rayleigh-Streuung (Physik) oder Desertifikationsprozesse (Geographie). Die ISS-Aufnahmen werden zur visuellen Unterstützung des Gelernten genutzt. Die Interaktion erfolgt bei den Arbeitsblättern mit kleinen Bildbearbeitungsanwendungen wie einem Swipe-Tool und lebensnahen Experimenten.

Das „Observatorium“ von Columbus Eye besteht aus interaktiven Lernwerkzeugen zur Klassifikation der HDEV-Aufnahmen. Videos von Überflügen der ISS wurden in Nadir-Panoramen verwandelt. Neben einem Panorama von Westafrika, welches die größte Wüste der Welt – die Sahara – sowie Savannenregionen und den Regenwald abbildet, gibt es auch die Möglichkeit Einblicke in Regionen Kanadas und das Leben in Schnee und Eis zu bekommen.

Ein drittes Tool thematisiert Südamerika und zeigt einen Überflug von der Küste Chiles über die Anden und die weiten Ebenen der Pampa zur Küste Argentiniens. In allen drei Regionen geht es insbesondere auch um die Auseinandersetzung mit unterschiedlichen Klimazonen. Die SuS können mit dem Observatorium eigenständig eine „Minimum-Distance“ Klassifikation der Landschaftsoberflächen an den ISS-Panorama durchführen und werden so mit Methoden der Fernerkundung, aber auch mit den Unterschieden und der Vielfältigkeit unserer Erde vertraut gemacht. Sie lernen, unterschiedliche Landoberflächen zu erkennen, diese einzuordnen und zu benennen und schließlich am Ende ihre eigene Karte zu erstellen. Darüber hinaus setzen sie sich jedoch auch mit der Problematik möglicher Fehlern dieser Methoden auseinander. Die Bewertung der aus der eigenständigen Klassifikation hervorgegangenen Karte hilft, die kritische Reflexion von computerbasierten Produkten zu fördern. Abbildung 3 zeigt eine exemplarische Klassifikation (links) und die daraus resultierende Karte (rechts) am Beispiel von Südamerika. Die dritte und umfassendste Interaktionsebene stellen die Lernmodule von Columbus Eye dar. Es sind „geschlossene“ Lernumgebungen und kombinieren Verarbeitungswerkzeuge der digitalen Bildanalyse mit fachspezifischen Aufgabenstellungen sowie erläuternden Animationen.

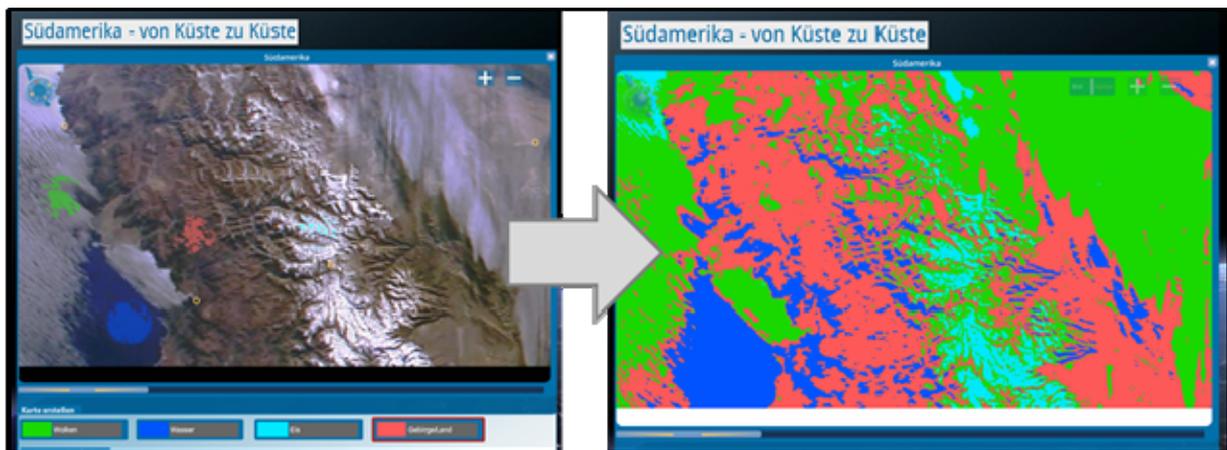


Abb. 3: Das Südamerika-Panorama im Columbus Eye Observatorium: Die SuS führen eigenständig eine Klassifizierung durch und erstellen eine eigene Landbedeckungskarte.

3 Mittelwertberechnung mit Bilddaten von der Internationalen Raumstation

3.1 Mathematikunterricht aus dem All – Einsatz Neuer Medien

Das Lernmodul „Mittelwertberechnung von der ISS“ setzt sich mit der Bildkorrektur auseinander und vermittelt neben dem technischen Verständnis zu digitalen Bildaufnahmen elementare Komponenten aus der Stochastik. Die Stochastik stellt eine zentrale inhaltsbezogene Kompetenz des Mathematikunterrichts in der Jahrgangsstufe 7 dar, welche mit dem zugehörigen Lerntool angesprochen wird. Wichtiger Bestandteil ist hier die Betrachtung und Interpretation relativer Häufigkeiten und Mittelwerte, wie insbesondere das arithmetische Mittel. Dieses Wissen wird schließlich genutzt, um mit Hilfe von Mittelwertfiltern eine Rauschunterdrückung in den ausgewählten Bildern der ISS anzuwenden. Die Unterrichtseinheit bedient sich der Möglichkeit des Computers, um die angesprochene Thematik durch Animation und Interaktion

nachhaltig zu vermitteln. Der Aufbau des Moduls ist wissenschaftsorientiert und fördert somit grundlegend das wissenschaftspropädeutische Lernen. Gemäß den didaktischen Prinzipien des moderaten Konstruktivismus wird so die Problemlösungs-Kompetenz der SuS angesprochen (VOß et al. 2011).

3.2 Hintergrund: Daten und Methodik

Insgesamt wurden vier Bilder aus Überflügen der ISS in das Lernmodul integriert. Zwei Bilder wurden aus einem Überflug über Kanada gewählt, zwei weitere von einem Überflug über die arabische Halbinsel. Eine der Kanada-Aufnahmen enthält einen „Blaustich“, welcher durch die Rayleigh-Streuung in der Atmosphäre verursacht wird. Die beiden Bilder der arabischen Halbinsel wiederum enthalten ein leichtes und ein starkes Rauschen, welches in fehlerhafte Bilddaten resultiert.

Die SuS werden zunächst in die Grundlagen der digitalen Bildbearbeitung eingeführt. Bilder aus dem All bestehen wie alle digitalen Bilder aus einem gleichmäßigen Zellenraster. Jede Rasterzelle, bzw. Pixel, beinhaltet einen Wert, der die von der Erdoberfläche an dieser Stelle reflektierte Strahlung repräsentiert. In dieser Lerneinheit können die SuS feststellen, dass es bei der Aufnahme von erdbeobachtenden Bildern immer wieder zu Bildfehlern, z. B. durch kurzfristige oder systematische Fehlfunktionen des Sensors oder durch atmosphärische Störungen oder Fehler bei der Datenspeicherung, kommen kann. Solche Empfangsstörungen resultieren in fehlerhafte Bildzeilen oder -spalten oder einem Bildrauschen, welches auch ‚noise‘ genannt wird. Mit Hilfe von Mittelwertfiltern können solche Datenfehler jedoch behoben oder zumindest verbessert werden.

Die Grundlage für eine Filterung der Bilddaten bilden gleitende Filtermatrizen, sogenannte Kernel. Innerhalb eines Kernels werden die Pixelwerte abhängig von ihrer Umgebung verändert. Das bedeutet, dass aus allen Pixeln einer definierten Umgebung ein Mittelwert berechnet wird und dieser Wert dem zentralen Pixel zugeordnet wird. (Abb. 4).

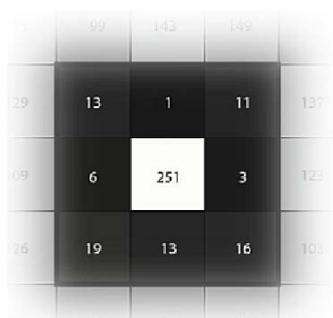


Abb. 4: Beispiel eines 3x3-Pixel Kernels wie auch im Lernmodul dargestellt.

Nach der erfolgreichen Berechnung bewegt sich der Kernel pixelweise weiter, weshalb er auch ‚Moving Window‘ genannt wird. Als statistischer Filter wird hier das arithmetische Mittel angewendet. Bei der Berechnung von diesem wird die Summe aller Werte eines Kernels durch die Anzahl aller Werte des Kernels geteilt. Dieses Vorgehen führt zu einer Glättung des gesamten Bildes und zur Unterdrückung des Rauschens.

3.3 Das Lernmodul: Mittelwertberechnung von der ISS

Die Unterrichtseinheit „Mittelwertberechnung von der ISS“ setzt sich aus zwei Teilen zusammen. Der erste Teil leitet als Einführung auf den zweiten Teil des Moduls, welches sich mit der praktischen Umsetzung beschäftigt, hin. Das Lernmodul wird als eigenständiges Flash-Programm ausgeführt. Mit einer Navigationsleiste, die am unteren Rand angebracht ist, bewegen sich die SuS durch das Modul. Alle zu einem bestimmten Stand verfügbaren Modulteile sind in der Navigationsleiste farbig gekennzeichnet. Weitere Teile werden erst anwählbar, sobald die SuS das erste Quiz gelöst haben.

Mit einem Einführungstext werden die SuS über den Inhalt und den Aufbau des Moduls informiert, sodass sie die Ziele der Lerneinheit eigenständig arbeiten können. Zu jeder Zeit gibt es zudem die Möglichkeit durch Anklicken eines Fragezeichens im Modul Hilfen zu erhalten.

In der einführenden Rubrik ‚Hintergrundwissen‘ werden zwei Komponenten erläutert: die Berechnung des arithmetischen Mittels sowie Bildstörungen und Funktionsweisen des Moving Windows (bzw. Kernels). Hier erleichtern kleinere Animationen das Verständnis der komplexen Thematik. Das Erlernete wird anschließend in einem Quiz abgefragt, das die SuS auch eine Beispielmatrix exemplarisch berechnen lässt. Im zweiten Modulteil erhalten die SuS schließlich die Bilder und das mathematische Werkzeug, um Korrekturen an den ISS-Daten vornehmen zu können. Ein Fenster mit Aufgaben hilft, dass sich die SuS während ihrer Arbeit orientieren können. Im zentralen Schritt können die SuS vier ISS-Bilder via drag & drop in das Hauptfenster gezogen werden.

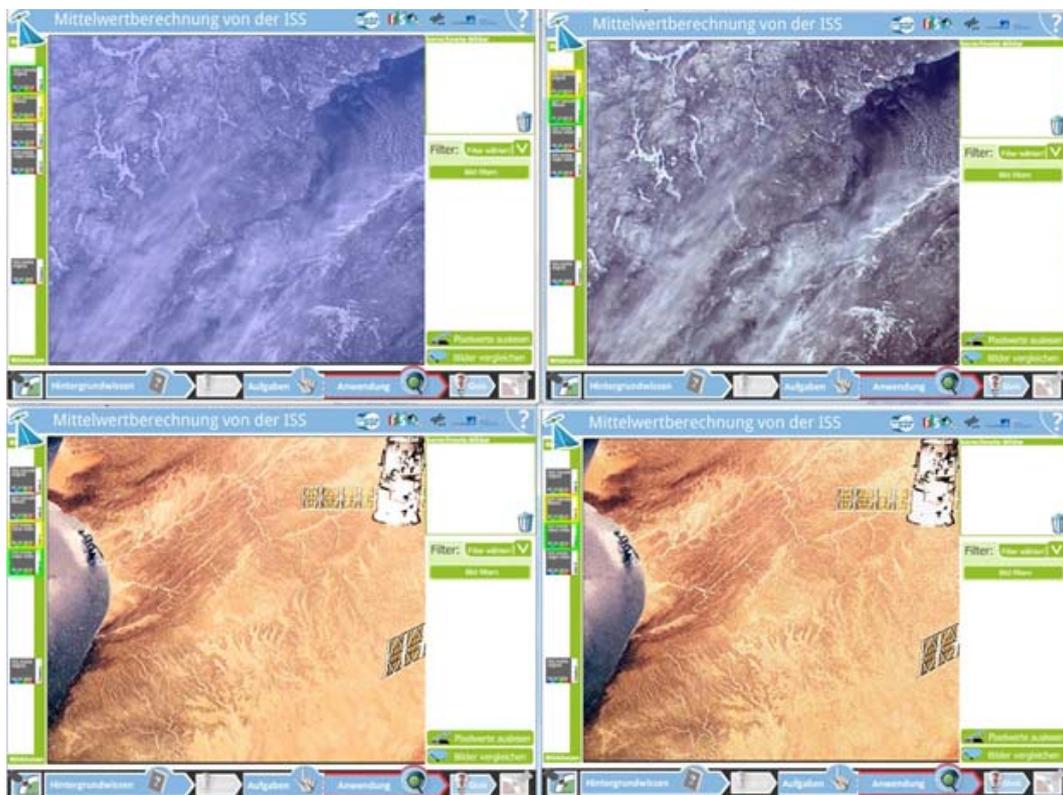


Abb. 5: Die vier zu bearbeitenden Bilddaten von der ISS: Oben ist die Korrektur der Rayleigh-Streuung, unten die Korrektur der fehlerhaften Pixel zu sehen.

Die beiden Bilder in Abb. 5 zeigen die unkorrigierten Aufnahmen über Kanada und der arabischen Halbinsel (links). Zur Erläuterung der Rayleigh-Streuung kann die Original-Aufnahme von Kanada mit dem bereits gefilterten Bild (oben rechts) im Swipe verglichen werden. Die Korrektur der Aufnahme von der arabischen Halbinsel übernehmen die SuS selbst. Mit Auswahl des Mittelwertfilters kann das Bild durch einen Klick geglättet werden. Die Mittelwertberechnungen werden nur auf ein ausgewähltes Fenster angewendet. Auf diese Weise haben die SuS die Möglichkeit die Funktionsweise des Filters nachzuvollziehen und ggf. nachzurechnen. Alle berechneten und gefilterten Bilder können in der rechten Modulleiste eingesehen werden, um die Veränderungen einzusehen und ggf. anpassen zu können. Dieser Teil des Moduls wird wiederum mit einem Quiz beendet, in dem die SuS ihr gerade erlangtes Wissen noch einmal abrufen können.

4 Fazit & Ausblick

Die Erdbeobachtung von der ISS ist in vielen Bereichen der MINT-Fächer einsetzbar. Ein Blick auf unsere Erde aus dem All zeigt nicht nur ihre Schönheit durch beeindruckende Bilder von weiten Landschaften oder komplexen Stadtstrukturen auf, sondern bringt die SuS auch dahin, uns mit der Verletzlichkeit unseres Planeten auseinanderzusetzen. Der Eingriff des Menschen oder die Zerstörungsgewalt von Naturphänomenen sind nur eine Möglichkeit, Erdbeobachtung zur Wissensvermittlung einzusetzen. Neben der Wissensvermittlung von Prozessen und Mustern der Mensch-Umwelt-Systeme im Geographieunterricht, wurde aufgezeigt, welche Möglichkeiten es geben kann, curriculumspezifische Themen weiterer naturwissenschaftlicher Fächer wie Physik und Mathematik mit Hilfe interaktiver Lerneinheiten und der selbstständigen Interaktion mit den ISS-Bilddaten zu bearbeiten.

Die im Columbus Eye- Projekt erstellten Unterrichtsmaterialien fördern neben der Medienkompetenz der SuS auch die praktische Umsetzung von immer noch vorwiegend theoretischen Konzepten. Das hier vorgestellte Beispiel „Mittelwertberechnung von der ISS“ nutzt Bildmaterial von der ISS, um Methoden der Stochastik, welche einen zentralen Bereich des Mathematikunterrichts der Jahrgangsstufe 7 ausmacht, anschaulich zu vermitteln. So werden die SuS in die Erdbeobachtung der ISS eingeführt und erlangen gleichzeitig die notwendigen Kenntnisse in ihrem Fachbereich – hier der Mathematik. Die direkte Anwendung von theoretischen Konzepten fördert zudem das kritische Hinterfragen dieser Konzepte. Beispielsweise können Störungen in Bildern ausgeglichen werden, dieses geschieht jedoch oft auch auf Kosten der Datenqualität. Das Ergebnisbild weist zwar keine Datenfehler oder -lücken mehr auf, jedoch lässt sich die Glättung des gesamten Bildes erkennen, sodass ehemals scharf erscheinende Strukturen unschärfer werden.

Das vorgestellte Lernmodul zur „Mittelwertberechnung von der ISS“ ist nur ein Beispiel, wie vielfältig das Bildmaterial aus der Erdbeobachtung im Allgemeinen und von der ISS im Speziellen in den Schulunterricht integriert werden kann. Die Lernmaterialien und interaktiven Lernumgebungen von Columbus Eye werden in Kürze auch in englischer Sprache verfügbar sein, sodass sie auch im bilingualen Schulunterricht eingesetzt werden können. Zusätzlich wird eine HTML5-Lösung als Substitution für die Flash-Umgebung konzipiert. Der Dreiklang aus ‚Sehen‘, ‚Informieren‘ und ‚Interagieren‘ wird wiederum weiter ausgebaut, um den

eigenverantwortlichen Umgang mit Geomedien zu fördern. Die SuS sollen nicht nur mit dem Blick von Astronauten die Erde betrachten können, sondern auch die Potenziale von alltäglichen Medien erfahren. Hierzu werden Arbeitsmaterialien entwickelt, mit denen die SuS die erdbeobachtenden ISS-Bilder mittels Augmented Reality auf dem eigenen Smartphone erfahren können. So könnte die ISS-Payload weiter zu einer sinnvollen ‚Payload‘ transformiert und die HDEV-Mission didaktisch in Wert gesetzt werden.

5 Literaturverzeichnis

- GOETZKE, R., HODAM, H., RIENOW, A. & VOß, K. 2014: Floods – Dealing with a Constant Thread. Learning and Teaching with Geomedia, Sanchez et al. (Hrsg.), Cambridge, 90-102.
- NASA, 2014: http://www.nasa.gov/mission_pages/station/research/experiments/917.html (letzter Zugriff 30.04.2016).
- RIENOW, A., GRAW, V., HEINEMANN, S., SCHULTZ, J., SELG, F., WEPPLER, J. & MENZ, G., 2015: Experiencing Space by Exploring the Earth – Easy-to-use Image Processing Tools in School Lessons. Proceedings of the 66th International Astronautical Congress 2015, October 12-16, Jerusalem, Israel, 1-7.
- SCHNEIDER, M. & LAPENTA, C.C., 2001: Payload Operations Integration Center Remote Operations Capabilities. Conference and Exhibit on International Space Station Utilization - 2001, Cape Canaveral, FL, October 15-18, 2001, 1-10.
- STEFANOV, W. & EVANS, C.A., 2014: The International Space Station: A Unique Platform for Remote Sensing of Natural Disasters. JSC Biennial Research Report, 108-110.
- VOß, K., GOETZKE, R., HODAM, H. & RIENOW, A., 2011: Remote Sensing, New Media and Scientific Literacy – A New Integrated Learning Portal for Schools Using Satellite Images. Learning with GI 2011 – Implementing Digital Earth in Education, Jekel, T., Koller, A., Donert, K. & Vogler, R. (Hrsg.), Berlin, 172-180.

Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen durch „Blended Learning“

Theresa Meyer¹ & Ansgar Brunn¹

Zusammenfassung: Um Unternehmen bei der Integration aktuell stark im Fokus stehender, berührungsloser Messverfahren zu unterstützen, wird in einem Projekt der Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt (FHWS), das hybride Bildungskonzept „Blended Learning“ verfolgt. Ziel ist es, den Teilnehmern grundlegende Fachkenntnisse in digitaler Photogrammetrie und Bildverarbeitung zu vermitteln. Die Lehr- und Lernstrategien beruhen dabei auf einer Kombination von multimedialen online Lerneinheiten und „klassischen“ Präsenzveranstaltungen, weshalb sie sich besonders für eine Zielgruppe im täglichen Arbeitsleben eignen. Eine stetig wachsende Projektteilnehmerzahl und deren stets große Beteiligung an den Aktivitäten im Projekt, zeigen bisher eine positive Resonanz auf das spezielle Informationsangebot.

1 Motivation

Vor dem Hintergrund stetiger und rasanter Neu- und Weiterentwicklungen photogrammetrischer Aufnahmegeräte und -verfahren, wie z.B. UAVs (unmanned aerial vehicles) und Multikamera-Systeme, ist berufsbegleitendes Lernen in den Fachbereichen digitale Photogrammetrie und Bildverarbeitung derzeit für viele im Arbeitsleben stehende Ingenieure ein relevantes Thema. Web-basierte Lerninhalte sind grundsätzlich eine beliebte Methode für ein flexibles, selbstorganisiertes Wissensmanagement. Sie ermöglichen Berufstätigen die Einarbeitung in neue Messverfahren, die aus zeitlichen, finanziellen oder logistischen Gründen ansonsten keinen Zugang zu geeigneten Informationsangeboten haben.

Im Internet findet sich bereits eine relativ große Palette an frei zugänglichen Informationen zu digitaler Photogrammetrie und Bildverarbeitung. Der Suchbegriff „Photogrammetry“, bringt auf der Videoplattform YouTube mehr als 30.000 Treffer (Stand: April 2016). Dort zeigen u. a. Hersteller und Vertriebspartner photogrammetrischer Mess- und Auswertesysteme ihre Produkte und deren Anwendung in kurzen Tutorials. Ein online-Video „V10 Imaging Rover Processing“ zeigt beispielsweise, wie Fotostationen eines Multikamera-Rovers orientiert werden. Ebenso finden sich dort Videomitschnitte von Vorlesungen an Hochschulen auf der ganzen Welt. Es gibt freie Informationsangebote in Form von Lexika, Links und Literatursammlungen, wie den „Geoinformatik-Service“ der Universität Rostock (<http://www.geoinformatik.uni-rostock.de/>). Genauso bieten viele Universitätsbibliotheken mittlerweile Artikel, Bücher und Publikationen auch als eBooks bzw. ePaper an.

Dieses vielseitige Onlineangebot beinhaltet sehr hilfreiche und wertvolle Informationen, um berufliches Vorwissen zu festigen und zu vertiefen. Es stellt den Lernenden jedoch auch vor die Herausforderung, die große Menge an Fachwissen selbstständig und innerhalb möglichst kurzer

¹ Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Labor für Photogrammetrie & Fernerkundung, Röntgenring 8, D-97070 Würzburg, E-Mail: [theresa.meyer, ansgar.brunn]@fhws.de

Zeit zu überblicken, um daraus nachhaltig Handlungskompetenzen für den Berufsalltag zu gewinnen.

Das Projekt „Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen – DiPhoBi4KMU“ der Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt (FHWS) greift diese Situation auf. Innerhalb des Projekts wird für Teilnehmer in der Berufspraxis ein speziell aufbereitetes Online-Lernangebot entwickelt. Angefangen bei den theoretischen Grundlagen digitaler Photogrammetrie und Bildverarbeitung bis hin zu Best-Practice-Beispielen, die dabei helfen neueste Verfahren in die täglichen Arbeitsabläufe von Ingenieurbüros zu integrieren. Lehr- und Lernkonzepte, die ausschließlich auf Web-Inhalten beruhen, erfordern von den Teilnehmern ein sehr hohes Maß an Eigenmotivation und Disziplin. Deshalb wird das Angebot in DiPhoBi4KMU um einen gewissen Anteil „klassischer“ Präsenzveranstaltungen in Form von z.B. Vorträgen, Seminaren oder Workshops ergänzt. Dabei werden die sozialen Vorzüge traditioneller face-to-face-Kommunikation genutzt. Die Kombination von E-Learning und traditionellem Präsenzlernen in einem Bildungskonzept, wird als „Blended Learning“ oder „integriertes Lernen“ bezeichnet. Die Herausforderung besteht darin, für eine Zielgruppe den optimalen Mix aus Methoden und Angeboten der Online- und Präsenzlehre zu finden und beide Lernformen didaktisch sinnvoll zu verknüpfen. Dieser Beitrag beschreibt die Konzeption und Umsetzung der Blended Learning – Strategie im Projekt DiPhoBi4KMU, die es zum Ziel hat, Teilnehmenden sowohl grundlegende als auch vertiefte Fachkenntnisse in digitaler Photogrammetrie und Bildverarbeitung für den Berufsalltag zu vermitteln.

2 Übersicht über den Artikel

In diesem Beitrag werden zunächst die Entstehung und Projektziele von DiPhoBi4KMU, sowie die Hintergründe zur Projektförderung vorgestellt. Der zweite Abschnitt beinhaltet die ausführliche Erläuterung des spezifischen Lehr- und Lernkonzepts im Projekt. Dazu gehören eine Charakterisierung der Zielgruppe, vorab definierte Lernziele, die konkreten fachlichen Inhalte und Strategie des Blended Learnings. Es folgt daraufhin die Beschreibung der zwei Hauptkomponenten des Bildungskonzepts: Virtuelles Lernen und Präsenzlehre. Die Selbstlehre wird durch einen E-Learning-Kurs realisiert, in dem den Projektpartnern speziell aufbereitete Lernmaterialien bereitgestellt werden. Im Beitrag wird dabei auch auf wesentliche Elemente wie Kommunikation, Lern-erfolgskontrolle und Evaluation beim web-basierten Lernen, eingegangen. Anschließend werden die verschiedenen Gestaltungsformen der Präsenzlehre in DiPhoBi4KMU vorgestellt, sowie der positive Einfluss der sozialen Aspekte „klassischer“ face-to-face-Kommunikation. Letztlich schließt dieser Beitrag mit einem Ausblick auf die weitere Projektentwicklung bis zum Projektabschluss im Sommer 2017 und auf die nachhaltige Projektwirkung, auch über die Projektlaufzeit hinaus.

3 Projekt “DiPhoBi4KMU”

3.1 Entstehung und Ziele

Die Methoden der Vermessung unterliegen derzeit einem Wandel hin zu berührungslosen Messverfahren mit bildgebenden Flächensensoren. Messgeräte wie UAVs, Tachymeter mit integrierten Digitalkameras oder Multikamera-Systeme, stellen Vermessungsingenieure und Geomatiker in der Praxis aktuell vor große Herausforderungen. All diese Verfahren haben gemeinsam, dass sie vom Anwender Spezialkenntnisse in digitaler Photogrammetrie und Bildverarbeitung erfordern. Um Ingenieurbüros dabei zu unterstützen, diese neuen Techniken in bestehende Arbeitsweisen zu integrieren, soll ein Transfer von Fachwissen in Photogrammetrie und Bildverarbeitung zwischen Hochschule und Unternehmen stattfinden. Übergeordnetes Projektziel ist es, das Qualifikationsniveau der Beschäftigten zu erhöhen und die Wettbewerbsfähigkeit kleiner und mittlerer Unternehmen zu steigern.



3.2 Projektförderung



Das Projekt DiPhoBi4KMU wird innerhalb des bayerischen ESF-Programms 2014-2020 „Perspektiven in Bayern – Perspektiven in Europa“ für zwei Jahre gefördert. DiPhoBi4KMU wird Projekten zur „Netzwerk­­tätigkeit zwischen Hochschulen und Unternehmen“ zugeordnet. Die teilnehmenden Unternehmen sollen zu mindestens 80% innerhalb der bayerischen Fördergebietskulisse (s. Abb. 1) angesiedelt sein. Der virtuellen Hochschule Bayern (vhb) ist die Projektträgerschaft übertragen.



EUROPÄISCHE UNION
EUROPÄISCHER SOZIALFONDS

ESF IN BAYERN
WIR INVESTIEREN IN MENSCHEN



Abb. 1: Bayerische Fördergebietskulisse.

4 Lehr- und Lernkonzept

4.1 Zielgruppe und Bedarfsanalyse

Das Lernangebot in DiPhoBi4KMU richtet sich an einzelne Mitarbeiter teilnehmender Unternehmen im Projekt. Sie fungieren als eine Art „Schnittstelle“, um das innovative Wissen aus der Hochschule in ihre jeweiligen Ingenieurbüros zu tragen. Bei der Konzeption des Lehr- und Lernangebots ist es vorrangig wichtig die Gruppe der Lernenden zu charakterisieren, um deren spezifische Anforderungen in der Projektgestaltung angemessen zu berücksichtigen und die Akzeptanz

der Teilnehmenden zu erhöhen. Dazu zählt es die Wissensstände, Interessen, berufliche Situation und voraussichtliche Gruppengröße, im Vorfeld näher zu beleuchten.

Um diese relevanten Informationen einzuholen, zu bündeln und auszuwerten wurde unter den Teilnehmenden zu Projektbeginn eine umfassende Bedarfsanalyse durchgeführt und im Rahmen des Projektaufakttreffens mit allen Projektpartnern diskutiert. Die beteiligten Arbeitnehmer und Arbeitgeber sind zum größten Teil Absolventen eines Studiums im Bereich der Geodäsie (84 %) oder einer einschlägigen Ausbildung (5 %) und arbeiten derzeit alle aktiv als Vermessungsingenieure oder -techniker. Etwa 11 % der Zielgruppe haben Grundlagen angewandter Vermessung ausschließlich als Nebenfach, im Rahmen einer anderen technischen Ausbildung oder Studiums kennengelernt. Entsprechend des Alters der Teilnehmer liegen Ausbildung oder Studium unterschiedlich lange zurück. Sie alle haben aber gemeinsam, dass die Fachrichtungen Photogrammetrie und Bildverarbeitung eher untergeordnet in Ausbildung bzw. Studium gewichtet wurden und dass dort im Hinblick auf die praktische Anwendung aktueller Verfahren erheblicher Nachholbedarf besteht. Hervorzuheben ist das allgemeine Interesse an Genauigkeitsbetrachtungen und -abschätzungen und Methoden der Kamerakalibrierung. Deutliche Trends zeichnen sich außerdem hin zu den Anwendungsschwerpunkten Bestandsaufnahme, Fassadenaufmaß und Mengenermittlung durch Methoden photogrammetrischer Messtechnik ab.

Derzeit (Stand April 2016) sind insgesamt 19 Teilnehmer aus 12 Unternehmen im Projekt verzeichnet.

4.2 Lernziele

In der „klassischen“ Schul-Lehre steht bislang noch meist der Erwerb von reiner Wissenskompetenz im Vordergrund. Der Erfolg des Lerntransfers wird dabei durch theoretische Wissensabfragen gemessen. Das bedeutet der Lernende kennt die Fakten, ist aber nicht zwangsläufig in der Lage Gelerntes zur Lösung eigener Problemstellungen anzuwenden. Im Kontrast dazu zielt das Lehr- und Lernkonzept in DiPhoBi4KMU darauf ab, den Teilnehmenden neben dem theoretischen Wissen über die Thematik, vor allem Handlungs-, Methoden- und Urteilskompetenzen für den Berufsalltag zu vermitteln. Sie sollen zum Projektende in der Lage sein, auf Basis des neu erworbenen Wissens eigene Handlungsabläufe zu entwickeln, um digitale photogrammetrische Messtechniken in die bestehenden Arbeitsweisen ihrer Ingenieurbüros zu integrieren. Dieses komplexe Kompetenzziel setzt sich aus mehreren Teil-Lernzielen zusammen.

Neben den theoretischen Grundlagen lernen die Projektmitglieder aktuelle photogrammetrische Vermessungsinstrumente wie beispielsweise Multikamera-Rover und UAV, kennen. Sie sind in Kenntnis über Aufbau, Funktionsweise und Handhabung der Geräte und überblicken die gängigsten Modelle verschiedener Hersteller auf dem Markt. Darüber hinaus sind sie bei erfolgreicher Projektteilnahme im Stande, auf Basis des eigenen Instrumentariums und individueller Kundenaufträge, Aufnahmekonfigurationen photogrammetrischer Messkampagnen zu planen und dabei stets einen Konsens zwischen Qualität und Wirtschaftlichkeit für das Unternehmen zu finden. Schließlich sind die Teilnehmer geschult im qualifizierten Umgang mit photogrammetrischer Auswertesoftware und kennen auch hier die Vorzüge unterschiedlicher Produkte. Sie können Techniken der Bildvorverarbeitung, -orientierung und -analyse anwenden und Resultate im Hinblick auf die weitere Verarbeitung auch kritisch beurteilen.

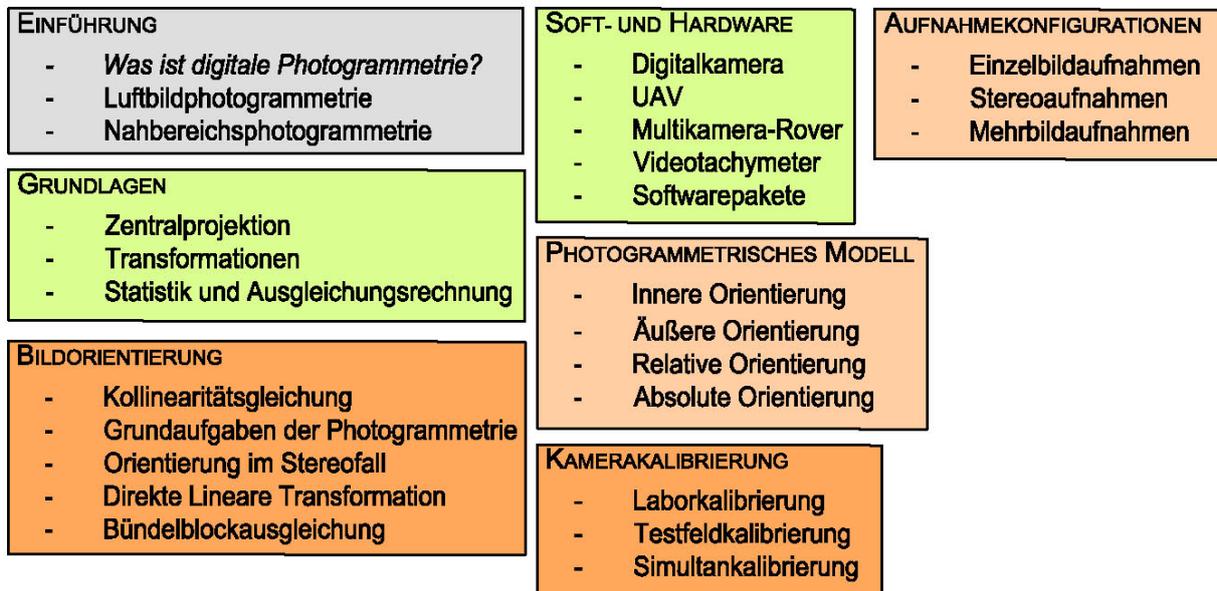
4.3 Fachliche Inhalte

Im ersten Projektjahr werden vor allem allgemeine Grundlagen in Photogrammetrie und Bildverarbeitung vermittelt. Zum Erwerb vertiefter Fachkenntnisse beinhaltet die zweite Projekthälfte komplexere Theorie. Um einen ständigen Praxisbezug zu wahren, werden die beiden Projektabschnitte mit den Anwendungen „Multikamera-Systeme“ und „UAV“ verknüpft (Tab. 1).

Tab. 1: Struktur der Projekthinhalte

Projektjahr	theoretische Information	Anwendung
1. Jahr	Grundlagen	Multikamera-Systeme
2. Jahr	komplexe Themen	UAV

DIGITALE PHOTGRAMMETRIE



DIGITALE BILDVERARBEITUNG

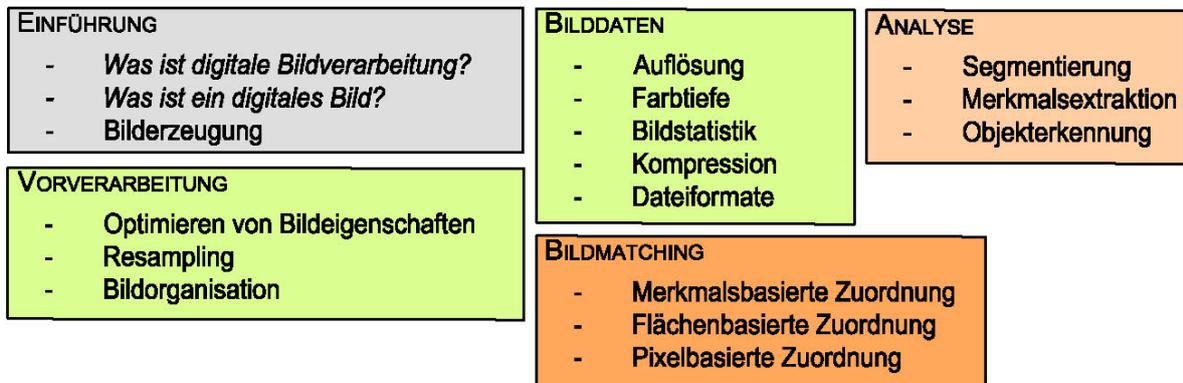


Abb. 2: Übersicht fachliche Projekthinhalte

Das aufbereitete Lernmaterial in DiPhoBi4KMU erhebt nicht den Anspruch eine umfassende Enzyklopädie des Fachwissens in Photogrammetrie und Bildverarbeitung zu sein. Zu einer professionellen Lehr-Konzeption gehört es, den Lernstoff so zusammenzustellen, dass das Lernziel von den Lernenden möglichst effizient und nachhaltig erreicht wird. Deshalb wäre das Durcharbeiten von Standardwerken der Fachliteratur im Projekt nicht geeignet. Stattdessen werden diejenigen Themen ausgewählt, die für das Verständnis der Hintergründe, bis hin zur Planung und praktischen Umsetzung aktueller photogrammetrischer Messmethoden relevant sind. Abb. 2 stellt eine grobe Übersicht über die ausgewählten fachlichen Inhalte im Projekt dar.

4.4 Strategie – Blended Learning

Vor dem Hintergrund einer Zielgruppe, die im täglichen Berufsleben steht und räumlich voneinander und der Hochschule getrennt ist, wird im Projekt DiPhoBi4KMU die Strategie des „Blended Learning“ verfolgt. Blended Learning heißt wörtlich übersetzt „vermisches Lernen“ und wird im Deutschen oft als „integriertes Lernen“ bezeichnet. Es bedeutet zunächst nichts anderes als die Kombination von virtueller und „klassischer“ Präsenzlehre. Die Vorzüge beider Lehr- und Lernformen kommen dabei zum Tragen. Zum einen sind die Lernenden unabhängig in der Gestaltung ihrer Lernphasen. Sie sind zu einem Großteil an keine festen Zeiten und Termine gebunden, die mit Beruf und Familie eventuell nur schwer zu vereinbaren wären. Zum anderen gewährleisten Präsenzveranstaltungen die sozialen Vorzüge der „face-to-face“-Kommunikation, ergänzend zu den online-Inhalten.

Viele Hochschulen nutzen bereits E-Learning-Plattformen zusätzlich zum Vorlesungs- und Übungsangebot. Im Unterschied zum Konzept in DiPhoBi4KMU beruht der überwiegende Teil der Lehre jedoch auf Präsenzveranstaltungen vor Ort. In DiPhoBi4KMU besteht die Herausforderung darin, das online-Angebot so zu gestalten, dass die Teilnehmer über weite Strecken alleine mit dem Lernstoff zurechtkommen. Dazwischen müssen Termine an der Hochschule zeitlich und inhaltlich sinnvoll gewählt werden, um das virtuelle Lernen optimal zu ergänzen.

5 Virtuelles Lernen

5.1 Online-Kurs

5.1.1 Technische Umsetzung

Der Projekt-online-Kurs ist in das bestehende E-Learning der FHWS eingebunden. Dabei handelt es sich um die weit verbreitete open-source Lernplattform „Moodle“. Üblicherweise haben nur Studierende und Mitarbeiter der FHWS personalisierten Zugriff auf das E-Learning. Für die Projektpartner in DiPhoBi4KMU werden externe Zugänge vergeben. Jeder Teilnehmer erhält einen eigenen Benutzernamen mit persönlichem Passwort und ist im Kurs durch ein Nutzerprofil für die Kursleitung und andere Teilnehmer identifizierbar.

5.1.2 Lernumgebung

Die Lernumgebung im E-Learning-Kurs kann als eine Art „Umwelt“, ähnlich einem sozialen Netzwerk betrachtet werden, in der sich die Lernenden während der gesamten Projektlaufzeit immer

wieder aufhalten. Gerade weil die Vorkenntnisse der Projektpartner zum E-Learning sehr unterschiedlich sind, wird auf eine übersichtliche und klar strukturierte Gestaltung der Kursseite geachtet (Abb. 3). Die Motivation der Teilnehmer würde schnell abreißen, wenn sie sich im Kurs nicht zurechtfinden. Deshalb baut der Kursbereich auf der einfachen Gliederung in die zwei Hauptschwerpunkte „digitale Photogrammetrie“ und „digitale Bildverarbeitung“ auf, innerhalb derer sich die hierarchisch strukturierten Lernmaterialien befinden.

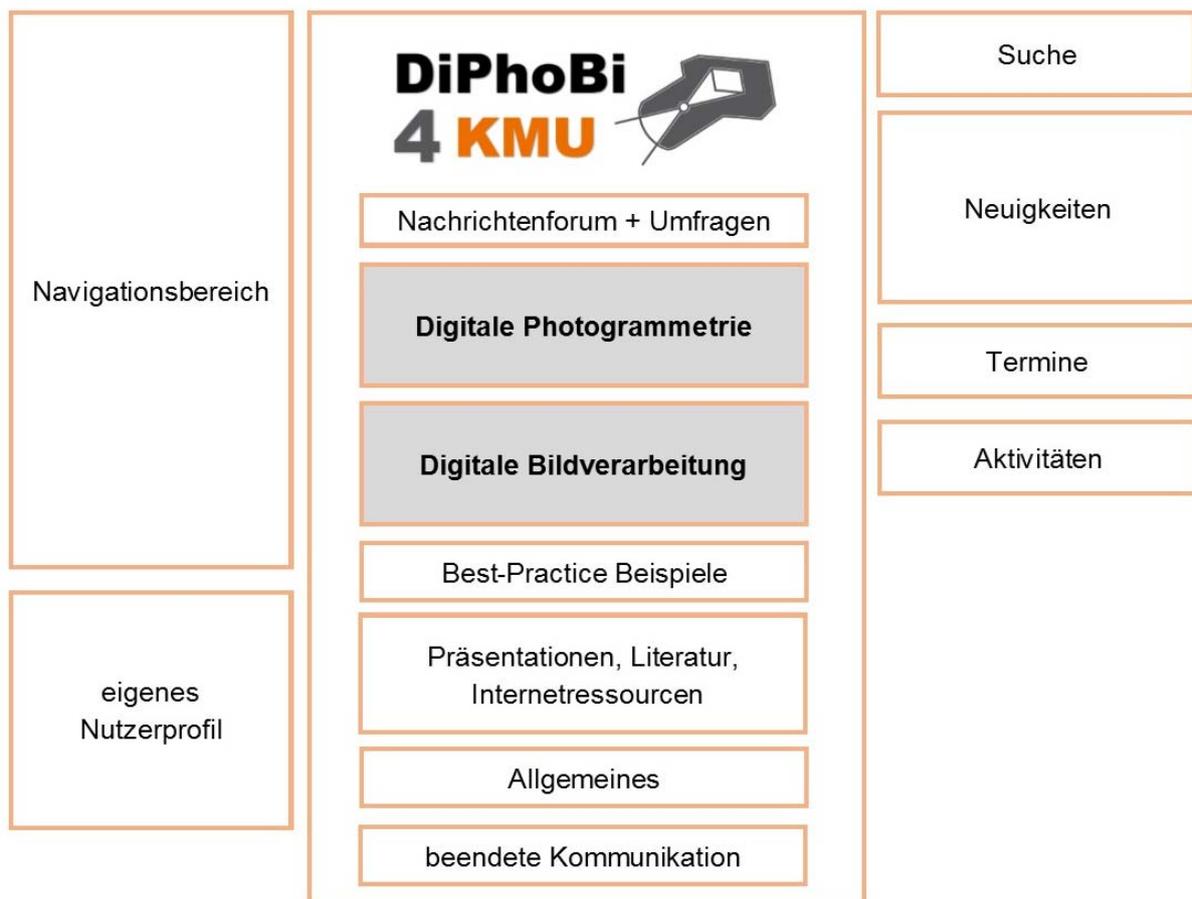


Abb. 3: Struktur des E-Learning-Kurses

Daneben gibt es eine Reihe wertvoller E-Learning-Funktionen und –Werkzeuge. An oberster Stelle im Kursbereich befindet sich ein Nachrichtenforum, sowie Umfrage- oder Abstimmungstools für die Kommunikation mit allen Projektbeteiligten. Eine Literaturquellensammlung und allgemeine Projektunterlagen finden sich im untersten Abschnitt der Seite. Am rechten äußeren Rand werden die aktuellsten Neuigkeiten, Termine und Aktivitäten des Kurses übersichtlich dargestellt. Visuell etwas abgegrenzt sind Navigationsbereich, die allgemeinen Einstellungen, sowie der Zugang zum eigenen Nutzerprofil angeordnet.

5.2 Lernmaterialien

Das Lernkonzept in DiPhoBi4KMU beinhaltet das Angebot einer speziellen „Ordnung“ des Wissens. Es werden hierarchisch strukturierte Lerneinheiten, in Form von einzelnen Kapiteln im E-Learning zur Verfügung gestellt. Dabei handelt es sich um eine Empfehlung für die Teilnehmenden, wo sie anfangen und womit sie thematisch fortfahren sollten. Die Themenschwerpunkte „digitale Photogrammetrie“ und „digitale Bildverarbeitung“ werden möglichst parallel bearbeitet. Die vorab definierten fachlichen Inhalte, werden kapitelweise ausgearbeitet und im PDF-Format zur Ansicht und zum Download im Kurs bereitgestellt (Abb. 4). Über die Projektlaufzeit sollen sich aus den einzelnen Dateien zwei Skripte zur Photogrammetrie und Bildverarbeitung entwickeln. Die Themenabschnitte bauen von Anfang an auf einander auf. Jeder Fachbegriff, der für das richtige Auffassen der Thematik wichtig ist, wird im Text erläutert. Alle Kapitel enthalten Grafiken, Tabellen, Diagramme oder auch kleine Berechnungsbeispiele, die das Verständnis erleichtern sollen und einen ständigen Praxisbezug herstellen. Am Ende jedes Unterkapitels werden dessen wichtigste Aussagen und Inhalte in einer kleinen „Merke-Box“ zusammengefasst.

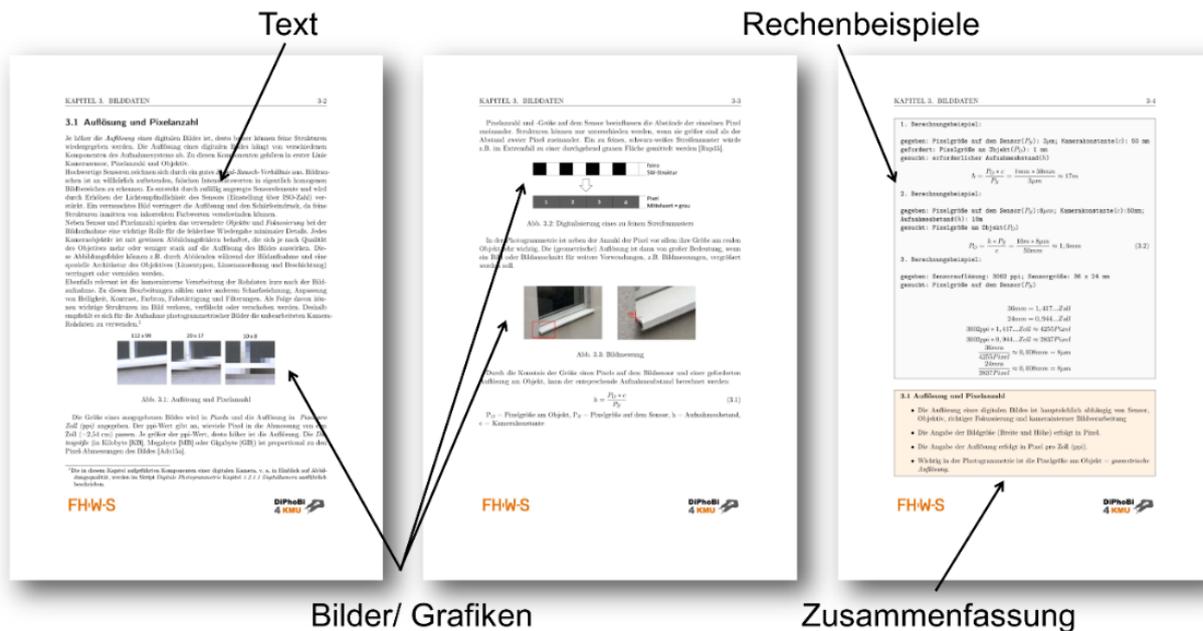


Abb. 4: Beispiel Lernmaterialien.

Ergänzend zu den allgemeinen Kapitel-Ausarbeitungen, werden auch einzelne Kurzberichte zu ausgewählten Themen angeboten. Dabei kann es sich um durchgeführte Testszenerarien an der Hochschule handeln, studentische Projekte oder auch Ergebnisse von Abschlussarbeiten, die thematisch passend und interessant sind.

5.3 Vernetzung und Kommunikation

Bei einem Lernarrangement, das größtenteils auf Web-basierten Inhalten und Methoden beruht, ist es notwendig zwischen Lehrern und Lernenden durch kontinuierliche Kommunikation eine

Verbindung zu schaffen. Soziale Interaktion ist wichtig für den Lernprozess in einer Webgemeinschaft. Daher ist es für den Lernerfolg der Gruppe erforderlich, dass sich die Teilnehmenden untereinander möglichst gut vernetzen und austauschen.

Virtuelle Kommunikation über eine online-Plattform hat die Besonderheit, asynchron zu funktionieren. Die Kommunikationspartner können nur zeitversetzt auf die eingehenden Informationen des jeweils anderen reagieren, denn zwischen ihnen steht ein PC oder Smartphone als Bindeglied. Der online-Kommunikationsprozess wird außerdem dadurch erschwert, dass wichtige „Werkzeuge“ der face-to-face-Kommunikation wegfallen. Mimik, Gestik, Betonungen (durch Stimmmodulationen) und allgemeine Körpersprache, tragen beim Kommunizieren von physisch präsenten Kommunikationspartnern wesentlich dazu bei, dass Informationen richtig übermittelt und verstanden werden. Der Informationsaustausch im E-Learning erfordert daher von allen Beteiligten spezielle Kommunikationskompetenzen.

Ein allgemein bekanntes Axiom der Kommunikationstheorie von Paul Watzlawick lautet: „One cannot *not* communicate“ (WATZLAWICK 1967: 51). Kommunikation findet schon allein dadurch statt, dass jeder Kursteilnehmer, sowohl Dozenten, als auch die Projektpartner, durch ein persönliches Nutzerprofil im Kurs vertreten sind. Diese Profile enthalten mindestens Vorname, Nachname, E-Mail-Adresse und Unternehmen der Person. Optional können ein Nutzerbild, weitere Kontaktdaten, Beschreibungen, Webseiten, Privat- oder Firmenadresse angegeben werden.

Die aktive Kommunikation im Projekt läuft hauptsächlich über ein Nachrichtenforum im online-Kurs ab. Aktuelle Neuigkeiten bezüglich der Projektentwicklung, z.B. Hinweise zu neu eingestellten Inhalten im Kurs oder Ankündigungen zu nächsten Präsenzveranstaltungen, werden darüber verbreitet. Eine besonders wertvolle Eigenschaft des Nachrichtenforums im E-Learning-Kurs ist die Tatsache, dass jeder Teilnehmer die Nachrichten auch direkt als E-Mail erhält. So ist weitgehend sichergestellt, dass alle Beteiligten Informationen unabhängig davon erhalten, ob sie nun gerade im Kurs eingeloggt sind oder nicht.

In DiPhoBi4KMU haben sich außerdem online-Umfragen oder –Abstimmungen als festes Kommunikationswerkzeug etabliert. Sie werden überwiegend für die inhaltliche und terminliche Planung von Projekttreffen oder Workshops an der Hochschule genutzt. Die Gruppe kann gemeinschaftlich abstimmen und die Projektleitung erhält relativ schnell aussagekräftige Ergebnisse. Um die Interaktion der Teilnehmenden weiter zu fördern, ist im online-Kurs ein Bereich für die Sammlung von fachlich interessanten Internetlinks eingerichtet. Die Projektpartner sind dazu aufgerufen, auf diese Weise Informationen mit der Gruppe zu teilen und im Idealfall darüber online oder bei Präsenzveranstaltungen zu diskutieren.

5.4 Lernerfolgskontrolle

Gemäß dem vereinbarten Lernziel (s. Abs.4.2) werden den Projektteilnehmenden Kompetenzen für die praktische Anwendung von Methoden digitaler Photogrammetrie und Bildverarbeitung vermittelt. Das Lernziel definiert den idealen IST-Zustand nachdem das Projekt beendet wurde. Um den Erfolg des Lerntransfers schon während der Projekt-/Lernphase zu messen, werden nach der Freischaltung von fertigen Themenblöcken im Kurs, kleine online-Lernerfolgskontrollen durchgeführt. Diese Wissenstests werden durch kurze Aufgaben, z. B. in Form von Multiple-Choice-Fragen, im E-Learning realisiert. Dadurch wird gewährleistet, dass sich die Projektpartner mit den Inhalten auseinandersetzen und die Anregung von Denkprozessen, während der Selbstlernphasen

im Projekt, wird gefördert. Sie dienen den Lernenden außerdem zur Selbsteinschätzung und Motivation. Jedoch ist die ausschließliche Abfrage von reinem Faktenwissen für die absolute Beurteilung von Lernerfolg, in einem kompetenzorientierten Lernangebot wenig aussagekräftig. Ob die Teilnehmenden durch den Projektverlauf tatsächlich zusätzliche Kompetenzen erwerben ist vielleicht nicht immer sofort ersichtlich und erfordert wiederum vom Lehrenden entsprechend diagnostische Kompetenzen.

In DiPhoBi4KMU wird es, etwa zur Hälfte der Projektlaufzeit (06/2016), eine Projektveranstaltung zum Thema „Anwendungen“ geben. Bei dieser Veranstaltung sollen sich die Teilnehmenden durch eigene Beiträge aus der Praxis aktiv einbringen. Jedem Projektbeteiligten wird die Möglichkeit gegeben, in einer kleinen Präsentation von persönlichen Berufserfahrungen und aktuellen (Test-) Projekten im Bereich Photogrammetrie und Bildverarbeitung zu berichten. In diesem Rahmen können Erfolge oder auch Schwierigkeiten in der Gruppe gemeinschaftlich analysiert werden.

5.5 Evaluationen

Evaluationen in der online-Lehre dienen der Kontrolle von Qualität, Funktionalität, Wirkung und Nutzen des Angebots. Kontinuierliche Wirkungskontrollen sind ein wesentlicher Erfolgsfaktor bei virtuellem Lehren und Lernen. Durch sie lässt sich ein Bildungsangebot kurzfristig oder auf lange Sicht verbessern. Unterschieden werden grundsätzlich zwei Arten von Evaluationen. Zum einen werden im Projekt *formative* Evaluationen durchgeführt. Sie dienen der entwicklungsbegleitenden Kontrolle während der Bildungsmaßnahme. Diese Art von Evaluation reflektiert Zwischenstände im Projektverlauf und wird jeweils nach abgeschlossenen Themenschwerpunkten oder Projektmeilensteinen durchgeführt. Die Teilnehmenden sollen dabei eine Reihe von Fragen zu Struktur, Aufbereitung (Bild und Text), inhaltliche Auswahl und Schwierigkeitsgrad der Lerninhalte, beantworten. Außerdem werden sie dazu befragt, wie sie mit den Werkzeugen und Funktionen des E-Learning zurechtkommen, z.B. Navigation im Kursbereich, Nutzung von Umfragetools, etc. An dieser Stelle wird der Gruppe auch die Möglichkeit für Anmerkungen, Lob und Kritik, in Form von anonymen Nachrichten an die Projektleitung gegeben.

Summative Evaluationen dagegen finden zum Ende der Bildungsmaßnahme statt. Sie dokumentieren die individuellen Einschätzungen der Teilnehmenden zum Projektendzustand. Inhaltlich geht es dann um zusammenfassende, abschließende Bewertungen, wie z.B. das Erreichen von persönlichen Lernzielen und Abschätzung zu zeitlichem und finanziellem Aufwand für die Maßnahme.

6 Präsenzlehre

Das Blended-Learning-Bildungskonzept im Projekt DiPhoBi4KMU beinhaltet neben den online-Selbstlernphasen, regelmäßige Präsenzveranstaltungen an der Hochschule in Würzburg. Sie sind als gleichwertiger Bestandteil des Lehr- und Lernarrangements zu betrachten und können z.B. als Workshop oder Seminar gestaltet werden. Pro Projektjahr sind etwa fünf Präsenztermine mit allen Projektpartnern vorgesehen.

6.1 Gestaltungselemente der Präsenzlehre

6.1.1 Vorträge

Die Projektpartner hören im Rahmen von Präsenzveranstaltungen an der FHWS, Fachvorträge zu ausgewählten Themen aus Photogrammetrie und Bildverarbeitung. Diese Vorträge sind in Inhalt, Länge und Aufbau mit traditionellen Hochschul-Vorlesungen zu vergleichen und nehmen stets Bezug zu den Lernmaterialien im E-Learning-Kurs. Auf diese Weise wird das (Vor-) Wissen aus den Selbstlernphasen gefestigt. Die Teilnehmenden erhalten die vorgetragenen Präsentationen als Handout zum Mitlesen, sie können sich während des Vortrages Notizen machen und natürlich Fragen stellen. So wurden bis zu diesem Zeitpunkt (April 2016) in DiPhoBi4KMU die Themen „UAV“, „Auswertestrategien“, sowie „Grundlagen der Photogrammetrie“ und „Grundlagen der Bildverarbeitung“ behandelt.



Abb. 5: Vortrag bei einem Präsenztermin



Abb. 6: UAV- Flugvorführung

6.1.2 Geräte-Demonstrationen

Bestandteile und Funktionsweisen von Digitalkameras, Multikamera-Rover, Videotachymetern und UAVs werden in den bereitgestellten Lernmaterialien im E-Learning ausführlich beschrieben. Doch Aufbau, Inbetriebnahme und Handhabung können erst durch Live-Vorfürungen wirklich praxisnah vermittelt werden. Während eines Projektworkshops im Februar 2016, konnten die Teilnehmenden bereits den praktischen Umgang mit dem Multikamera-Rover Trimble® V10 üben indem sie nach einer kurzen Geräte-Einweisung selbst einige 360°-Panoramen auf dem Hochschulgelände aufnehmen durften. Im Oktober 2015 fand ein Anwenderforum an der FHWS statt, bei dem den Projektpartnern UAV-Flugvorführungen geboten wurden (Abb. 6).

6.1.3 Übungen

Die vermittelten Lerninhalte werden bei praktischen Übungen im Labor für Photogrammetrie und Fernerkundung der Hochschule, vertieft und gefestigt. Innerhalb der Übungsblöcke bei Präsenzveranstaltungen, konnten die Teilnehmenden beispielsweise einen Luftbildstreifen des Würzburger Stadtbereichs orientieren und anschließend Gebäude im Luftbildblock dreidimensional erfassen. Außerdem wurden die Aufnahmen einer Geräte-Vorführung des Trimble V10 genutzt, um daraus die automatische Punktwolkengenerierung aus Fotos im Nahbereich zu demonstrieren.

6.2 Sozialer Aspekt

Kontakte, die zwischen den Projektmitgliedern während Präsenzterminen an der FHWS entstehen, bilden die wichtigste Grundlage für dauerhafte Vernetzung und Kommunikation über das Internet und mindern maßgeblich die Abbruchquote, eines hauptsächlich Web-basierten Bildungsangebots. Erfolgreiches Blended Learning erfordert ein Mindestmaß an sozialer Interaktion unter den Lernenden. Durch den direkten Kontakt beim Präsenzlernen, werden die Vorteile der face-to-face-Kommunikation, für den Lernerfolg voll ausgeschöpft. Die synchrone Art der Kommunikation ermöglicht es den Teilnehmenden spontan auf ihre Gesprächspartner zu reagieren um Informationen auf diese Weise wesentlich effizienter auszutauschen, als zeitversetzt über das Internet. Die Entstehung einer sozialen Gruppe fördert die Motivation und Bindung an das Projekt.

6.3 Evaluationen

Auch für Präsenzveranstaltungen sind Evaluationen wichtig, um deren kurz-, mittel oder langfristige Qualität zu optimieren. Den Teilnehmern eines Präsenztermins, z.B. Workshop, werden zum Veranstaltungsende Evaluationsbögen zum anonymen Ausfüllen ausgeteilt. Dabei werden Einschätzungen zur persönlichen Zufriedenheit mit Anforderungsniveau, Tempo, Inhalten, Medien und Materialien, Lerneffekt und Organisation, anhand einer Antwortskala abgefragt. Zusätzlich können die Teilnehmenden Lob und Kritik anbringen, sowie Wunsch-Inhalte für nächste Veranstaltungen äußern. Von den Projektpartnern favorisierte Themen sind *Kamerakalibrierung*, *Genauigkeitsbetrachtungen*, *Bildvorverarbeitung* und *–Interpretation*. Für die Planung zukünftiger Präsenztreffen wünschen sie sich noch mehr Raum für den persönlichen Austausch untereinander.

7 Ausblick

Derzeit befindet sich DiPhoBi4KMU noch mitten in der Förderphase. Daher steht die Gesamtevaluation der Bildungsmaßnahme noch aus. Die Lernmaterialien für den online-Kurs werden kontinuierlich weiterentwickelt und den Projektmitgliedern bereitgestellt. Bis zum Projektende (06/2017) werden die fachlichen Inhalte vollständig ausgearbeitet sein, um den Projektpartnern auch für die Zeit nach DiPhoBi4KMU, als kompaktes Nachschlagewerk zur digitalen Photogrammetrie und Bildverarbeitung in der Berufspraxis zu dienen.

Bis dahin werden noch einige Präsenzveranstaltungen zu verschiedensten Themen an der FHWS stattfinden. Die Kontakte, die während der Projektphase zwischen den einzelnen Teilnehmenden sowie den Teilnehmenden und der Hochschule entstehen, werden in einer angedachten Interessensgemeinschaft „Photogrammetrie und digitale Bildverarbeitung in der Vermessung“ nachhaltig gepflegt, weiter fortbestehen und für zukünftige Projekte genutzt.

8 Literaturverzeichnis

- ESCHNER, A., 2009: Didaktik für Ingenieure, Erkenntnisse und Tools. Didaktiknachrichten (DiNa 05) – Publikation des DiZ – Zentrum für Hochschuldidaktik der bayerischen Fachhochschulen (Hrsg.), https://www.diz-bayern.de/images/documents/82/DiNa_05-2009_web.pdf
- HASSELHORN, M. & GOLD, A., 2006: Pädagogische Psychologie, Erfolgreiches Lernen und Lehren. Kohlhammer Verlag, 488 S.
- KERRES, M., 2013: Mediendidaktik, Konzeption und Entwicklung mediengestützter Lernangebote. 3., vollständig, überarbeitete Auflage, De Gruyter Oldenbourg, 543 S.
- KERRES, M., DE WITT, C. & STRATMANN, J., 2002: E-Learning. Didaktische Konzepte für erfolgreiches Lernen. Jahrbuch Personalentwicklung & Weiterbildung 2003, von Schwuchow, K. & Guttman, J. (Hrsg.), Hermann Luchterhand Verlag.
- KRAPP, A. & WEIDEMANN, B., 2006: Pädagogische Psychologie. 5., vollständig überarbeitete Auflage, Beltz Verlag, 864 S.
- NIEGEMANN, H., DOMAGK, S., HESSEL, S., HEIN, A., HUPFER, M. & ZOBEL, A., 2007: Kompendium multimediales Lernen. Springer-Verlag Berlin Heidelberg, 661 S.
- ROST, D.H., 2010: Handwörterbuch Pädagogische Psychologie. 4. Auflage, Beltz Verlag, 1034 S.
- WATZLAWICK, P., BEAVIN, J.H., JACKSON, D.D., 1967: Pragmatics of Human Communication. W. W. Norton, New York, 296 S.

Les drones au service de l'enseignement

BERTRAND CANNELLE¹

Résumé : Pendant de nombreuses années, l'apprentissage de la photogrammétrie s'est fait grâce à des images argentiques, des stéréoscopes et des restituteurs. La prise de vue aérienne ainsi que les calculs de plan de vols ne se faisaient qu'à l'aide de manuels et d'illustrations. Le processus photogramétrique complet, de l'acquisition à la production, d'une carte était difficilement réalisable : il était coûteux et prenait également beaucoup de temps (acquisition des images, développement des clichés, mise en place des couples...).

Aujourd'hui, le plan de vol, calculé en classe, est réalisé quelques instants plus tard par un drone. Les données sont directement exploitables dès la plateforme posée : contrôle du recouvrement effectif, contrôle qualité des images et aérotriangulation. Là aussi, le travail a considérablement changé : fini les six points manuels de von Gruber par couple et vive les points SIFT par milliers ! La production de modèle de terrain, limitée à des courbes de niveaux, est réalisable aujourd'hui en quelques minutes. La production d'ortho-photos est également réalisée en quelques instants en utilisant le modèle numérique issu de la corrélation.

Ces nouvelles techniques n'ont pas que des avantages : l'explication des algorithmes (SIFT, aérotriangulation, modèles numériques) n'est pas toujours aisée et il n'y a plus d'explications réalistes possibles : les restituteurs ont été remplacés par des fichiers contenant les rapports des calculs beaucoup moins intuitifs.

Cet article présente les évolutions de l'enseignement à la fois d'un point de vue pratique mais également d'un point de vue théorique. Il met également l'accent sur les aspects historiques qui sont encore visibles aujourd'hui à travers les logiciels.

1 Introduction

La photogrammétrie a été, pendant des dizaines d'années, une discipline confidentielle réservée aux instituts géographiques nationaux parce que les matériels utilisés étaient extrêmement coûteux : caméras argentiques calibrées précisément, mesure de points sur le terrain afin de caler les images dans un référentiel connu et restitutions sur des appareils dotés de qualités optiques et mécaniques extrêmement efficaces.

Depuis quelques années, l'arrivée du numérique mais surtout des drones ont permis de démocratiser cette discipline (KASSER & EGELS 2001). Par contre, cela a entraîné quelques bouleversements sur l'ensemble de la chaîne de production et également dans l'enseignement.

Dans cet article, nous allons présenter comment le drone a révolutionné l'apprentissage de la photogrammétrie, puis nous expliciterons l'intérêt de celui-ci à chaque étape du processus : préparation, acquisition, traitement des données et production.

¹ HEIG-VD, route de Cheseaux 1, CP 521, CH-1401 Yverdon-les-Bains,
E-Mail bertrand.cannelle@heig-vd.ch

2 L'initiation à la photogrammétrie

La plupart des personnes ayant fait de la photogrammétrie se souviennent de leur premiers cours : démonstration des équations de la parallaxe, passage sur un restituteur pas toujours adapté à la vision de la personne (d'où un léger mal de tête...), apprentissage du vocabulaire : convergence, site, déversement... (ALBERTZ & KREILING 1980). En général, il a marqué les esprits mais, malheureusement, pas forcément de façon positive. Certains ont même découvert tardivement que les mesures se faisaient sur des clichés et que les formules permettant les mesures tridimensionnelles n'étaient en fait que des vis de réglages des instruments.

Aujourd'hui, la première prise de contact est beaucoup plus facile : il suffit de présenter le drone (figure 1), la camera qu'il emporte et le tour est joué. Il est possible en une demi-journée de réaliser toutes les étapes :

- Acquisition d'images par drones
- Visualisation des données
- Réalisation d'une aérotriangulation
- Production d'un modèle numérique de surface ainsi qu'une orthophoto.
- Restitution d'objets ou de bâtiments en utilisant la saisie multi-images.

Après ce premier contact, il est beaucoup plus facile de parler de points de liaison ou d'appui, de position et d'orientation d'images dans l'espace...



Fig. 1: Exemple de drones pouvant être utilisés. (a) et (b) drones professionnels, (c) et (d) drones amateurs

3 Préparation de chantiers, plan de vol et acquisition

La préparation de chantiers photogrammétriques était réservée aux sociétés de prises de vues. Ce point n'était pas forcément utile dans une formation traditionnelle et n'était pas toujours abordé. Aujourd'hui le vol est souvent réalisé par la personne qui souhaite exploiter le chantier. Elle doit donc être sensibilisée aux problématiques de l'acquisition : autonomie de la plateforme, choix de la résolution, contraintes météorologiques, adaptation au terrain.

Sur la figure 2, on peut voir un exemple de plan de vol complexe. En 2.a, l'altitude de vol a dû être adapté afin d'assurer une résolution adaptée sur toute la zone. En 2.b, on se rend compte des difficultés liées à l'acquisition avec un drone. La zone en bleue indique une reprise de bande : la surface étant importante, le vol a été découpé en deux zones et, afin d'éviter les problèmes de recouvrement, une bande commune entre les deux vols a été réalisés. Par contre, un des problèmes qui arrive de temps en temps et l'absence d'une ou plusieurs photos (en rouge sur la figure 2.b).

En effet, sur les drones amateurs, il n’y a aucun acquittement lorsqu’une photo est prise. Il faut donc adapter les paramètres de recouvrement afin de pallier ce problème, par exemple en indiquant un recouvrement de 60% par 60%.

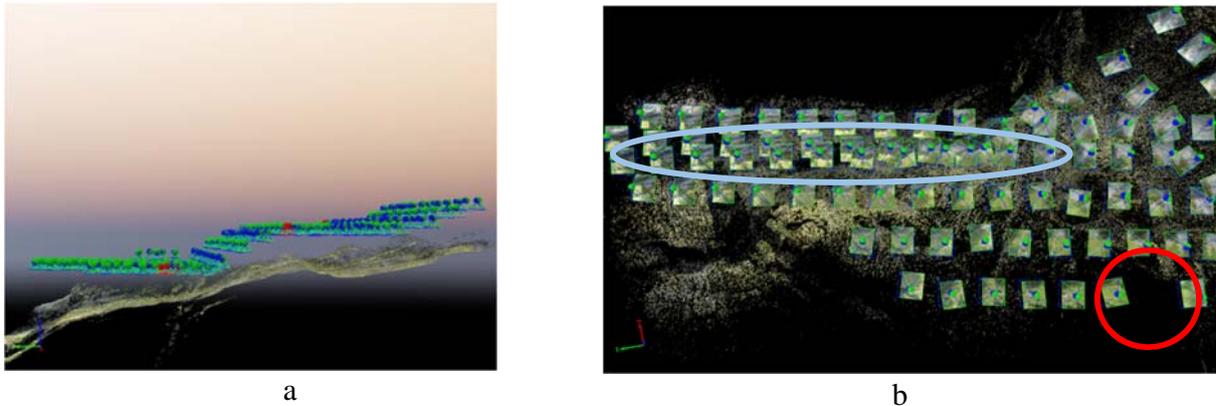


Fig. 2: Plan de vol complexe. (a) gestion des altitudes pour adapter la résolution. (b) bleu : reprise de bande rouge : absence de données.

4 Dépouillement des données

Après l’acquisition des données, il faut passer au post-traitement, étape qui permet d’associer les méta-données de vol aux images afin que chacune ait une position et une orientation. Ces métadonnées associées aux images se trouvent dans les tags exifs, informations visibles par l’explorateur windows (voir figure 3) ou grâce à certains logiciels spécifiques (exiv2 (EXIV2, 2016)).

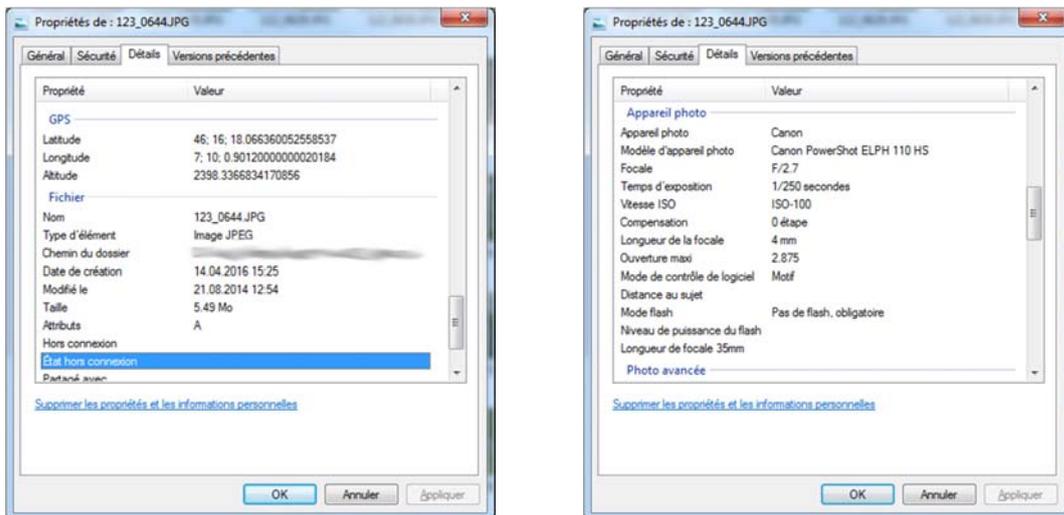


Fig. 3: Exemple d'informations de positionnement d'une image.

5 Aérotriangulation

Tous les drones ne possèdent pas de système de positionnement, ou il n'est pas assez précis. C'est pourquoi il est toujours nécessaire d'effectuer l'aérotriangulation, étape qui permet de retrouver de manière précise, la position et l'orientation des images au moment de la prise de vue. Nous allons détailler les étapes importantes qui sont arrivées depuis le numérique.

5.1 Points de liaison

Les points de liaison en photogrammétrie sont un élément essentiel pour assembler les images les unes par rapport aux autres. La première chose enseignée était les 6 points de von Gruber (figure 4) qui consiste à déterminer 6 points homologues par couple (5 étant le minimum + 1 en contrôle), répartis de façon uniforme dans la zone de recouvrement. Le numéro et l'emplacement des points sont indiqués sur la figure 4.

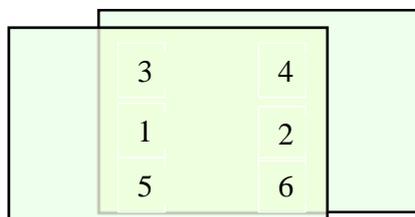


Fig. 4: Points de von Gruber

Avec les drones, nous n'avons pas (ou difficilement) de bandes rectilignes avec des recouvrements rectangulaires (voir partie 4) et le travail ne se fait plus manuellement. Tous les logiciels actuels utilisent des algorithmes (par exemple, SIFT (LOWE 2004) ou encore FAST (ROSTEN & DRUMMOND 2006)) qui détectent et appariant les points automatiquement. Le principal avantage est le nombre de points extraits, pouvant aller jusqu'à plusieurs milliers par couple, mais malheureusement comportant un nombre de fautes non négligeable. Il a donc fallu mettre en place des stratégies (comme RANSAC (FISCHLER & BOLLES 1981), qui consiste à tester plusieurs sous ensemble de données pour calculer l'orientation relative des images et de conserver la meilleure) basées sur les statistiques, ou encore sur la position approchée des images afin de ne garder que les bonnes mesures. Un deuxième inconvénient est la répartition non homogène dans la zone de recouvrement ce qui peut entraîner des mises en places erronées. La figure 5 présente les points homologues extraite (5.b) à partir d'une image (5.a). Certaines zones sont plus propices à la détection de points (changement de contraste, de couleurs, détails dans l'image : pierre...) et d'autres moins (zones uniformes, neige...)



Fig. 5: Exemples de points homologues automatiques. (a) Image initiale et (b) points de liaison extraits

Le numérique a permis d'effectuer un grand nombre de mesures dans un temps relativement court mais dont la qualité dépend : du contenu de l'image et des algorithmes de détection de fautes qui peuvent être perfectibles : c'est aujourd'hui un problème récurrent lorsque les chantiers ne fonctionnent pas et dont il faut essayer de prendre en compte lors de l'acquisition, notamment en augmentant le recouvrement dans les zones complexes ou en réalisant des bandes à différentes altitudes afin d'augmenter les zones propices à la détection.

5.2 Compensation

La compensation a elle aussi évoluée, notamment à cause des nombres de mesures par couple (cf partie précédente) mais elle a aussi dû s'adapter. Le but de la compensation est de déterminer précisément la position et l'orientation des images lors de la prise de vue. Malheureusement, les drones ne peuvent pas garantir les acquisitions identiques aux plans de vols initiaux. Il a donc fallu mettre en place des stratégies afin de pallier cette difficulté mettre l'accent dans la formation sur les étapes automatiques actuelles et leurs limitations.

5.2.1 Avec solution initiale

La solution initiale provient généralement d'un GNSS qui sert également au pilotage du drone. Dans la majorité des cas la précision de position est métrique. Concernant l'orientation, elle est soit déduite de la trajectoire du drone, soit issue de capteur d'orientation dont la précision est de quelques degrés. Ces informations servent, dans un premier temps, de prédicteurs pour l'appariement d'images : inutile d'appareiller une image en début de bande avec une image située à la fin. Ensuite, cette solution initiale peut servir à la détection de faute sur les mesures homologues elles-mêmes en se basant sur des considérations géométriques (résidus images...) afin d'aider les algorithmes de détection de fautes. Malheureusement, dans certains cas, lorsque la précision est trop dégradée (mauvaise réception GNSS à cause de masques ou de multi trajets par exemple), on peut être contraint de supprimer ces informations. En effet, des informations erronées ne permettent plus de détecter les mesures vraies ou fausses et entraînent une divergence du calcul. Le processus devient donc identique à celui expliqué juste après.

5.2.2 Sans solution initiale

Tous les drones ne possèdent pas de système de positionnement. Des stratégies ont donc été mises en place pour déterminer une solution initiale avant d'effectuer la compensation mais avec certaines limites.

Tout d'abord, la détection des points homologues (voir partie 5.1) a lieu entre tous les couples d'images possibles dans la mesure où le logiciel ne sait pas quelles images sont proches les unes des autres. Il y a, pour N images, $N \times (N-1)$ couples à calculer, ce qui peut être très long... Ensuite, le programme choisi une image, généralement l'image qui a le plus points homologues, comme référence, puis calcule les poses relatives des images voisines par rapport à cette image et ainsi de suite.

Malheureusement, ces hypothèses ne se basent que sur des statistiques et peuvent parfois arriver à des mises en place erronées : deux images sont mises en place avec uniquement des points de liaisons faux, mais qui n'ont pas été détectés comme tels. C'est le cas notamment en photogrammétrie de façade ou les fenêtres se ressemblent entre les étages. On peut donc arriver à un géoréférencement totalement aberrant mais avec des indicateurs relativement bons

Il faut donc que l'opérateur ait un regard critique afin de trouver l'erreur et de remédier au problème. En pratique, il faut invalider toutes les images mal géoréférencées, puis les réactiver les unes après les autres en fonction de leurs emplacements par rapport au bloc déjà mis en place.

5.3 Validation et export des résultats

Les données d'aérotriangulation ne sont pas très intéressantes en soit. Ce sont les produits qui en découlent qui le sont (voir partie suivante). Par contre, il est nécessaire que les résultats soient valides afin d'avoir des produits de qualité. Il faut donc connaître les outils permettant de savoir si un calcul est cohérent ou non et si toutes les données sont exploitables. Nous présenterons également les difficultés des formats d'exports, qui doivent permettre de passer d'un logiciel à l'autre mais qui ne sont pas toujours simples.

5.3.1 Validation des résultats

Nous allons présenter et détailler différents indicateurs qui sont généralement fournis par les logiciels et qui permettent de savoir si un calcul est bon ou non.

5.3.1.1 Indicateurs statistiques

Après une compensation, il y a différentes grandeurs qui sont calculées et qui doivent faire l'objet d'une analyse :

- Résidus moyens / erreur moyenne / erreur de reprojection : correspond à l'écart entre la mesure faite dans l'image (points de liaison) et la reprojection du point 3D correspondant (mesure en orange et reprojection en vert figure 6.a). Elle est normalement exprimée en pixel (car on travaille avec des images numériques) mais peut encore être exprimée en millimètre. Normalement cette valeur doit être inférieure à 1 pixel.
- Taux de recouvrement (figure 6.b) : cette valeur permet de savoir si une zone a été acquise par 1, 2, 3 images ou plus. En photogrammétrie, il faut qu'une zone soit photographiée par au moins 2 images pour être mesurable par stéréoscopie. Cette carte ne prend pas forcément en compte les zones de masques provoquées par les bâtiments notamment en milieu urbain, ni les masques dûs au relief accidenté. Il faut donc analyser cette carte avec soin.
- Nombre de points de liaison entre les images (figure 6.c) : dans la mesure où les points de liaison sont calculés de manière automatique, nous ne maîtrisons ni le nombre (ou uniquement le nombre maximum) ni la répartition. Il faut donc être vigilant sur ce type de graphique afin d'éviter les zones avec peu de points de liaison : zone où le bloc est moins bon géométriquement et là où il peut y avoir des anomalies géométriques comme par exemple sur la zone en rouge de la figure 6.c. En fonction des logiciels, on peut : relancer les calculs en augmentant le nombre de points de liaison, ajouter des points de liaisons manuels ou encore refaire un vol si nécessaire.

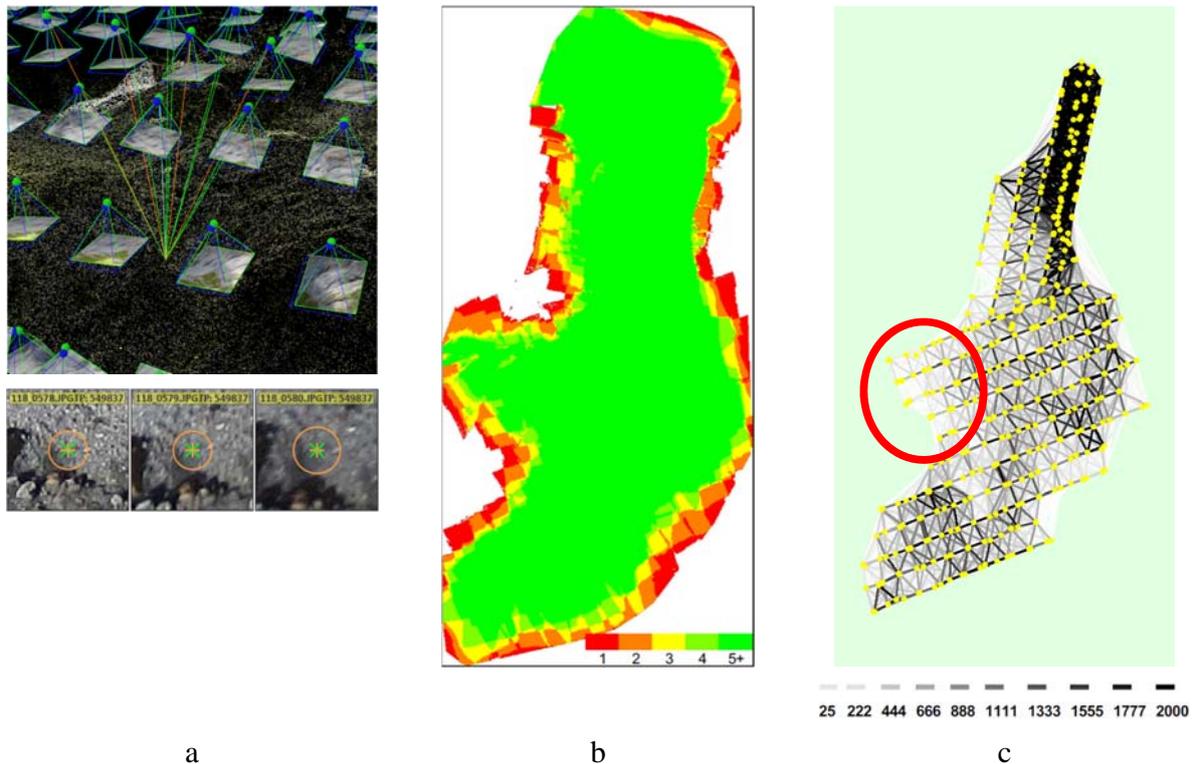


Fig. 6: Indicateurs après compensation

- Résidu sur les points d'appui (GCP : Ground Control Point) : deux grandeurs sont généralement fournies :
 - Erreur sur les coordonnées 3D : grandeur exprimée dans l'unité du système terrain (m ou cm). L'erreur doit être en adéquation avec la précision des coordonnées 3D mais également avec la précision de pointé de l'objet dans l'image (précision de l'opérateur ou de l'algorithme d'extraction).
 - Résidu image / erreur de reprojection : correspond à l'écart entre la mesure faite dans l'image et la reprojection du point 3D correspondant. Là encore, l'erreur doit être en adéquation avec la précision de pointé de l'objet dans l'image (précision de l'opérateur ou de l'algorithme d'extraction). Pour un opérateur humain, elle est de l'ordre du pixel. Pour un algorithme, elle peut être de 0.1 pixel voire mieux.

Une erreur sur les résidus images est souvent révélateur d'une erreur de pointé : l'opérateur s'est trompé de cible, de détail... entre les différentes images. De forts résidus sur les points terrains peuvent venir soit : d'une erreur de coordonnées/de nom ou d'une mauvaise aérotriangulation.

Il faut également prendre garde que l'erreur ne vienne pas forcément du point ayant les plus forts résidus parce que le processus de compensation consiste à répartir les écarts entre toutes les mesures sur les points de liaison et d'appui.

Une fois tous ces critères validés, si l'opérateur souhaite changer de logiciel, exporter les données.

5.3.2 Exports des résultats

L'export de résultats devrait être une étape relativement simple et triviale, mais en pratique ne l'est pas forcément. Voici les principales sources de problèmes lors de la conversion :

- L'unité des mesures images : aujourd'hui nous travaillons avec des images numériques et la plupart des logiciels travaille avec comme unité le pixel. Seulement, certains logiciels historiques travaillent encore en millimètre. Il faut donc convertir les données en connaissant la taille d'un pixel (souvent quelques microns...). Attention, la focale sur les appareils photos compacts utilisés peut être de quelques millimètres alors que les caméras aériennes étaient plutôt de l'ordre d'une centaine de millimètres.
- Le repère image : les images numériques ont traditionnellement un repère comme sur la figure 7.a : origine en haut à gauche, axe des x correspondant aux colonnes dans l'image et axe des y vers le bas correspondant aux lignes. Le repère présenté sur la figure 7.b correspond au repère image historique : origine au centre de l'image et axe des x horizontal et axe des y vertical. Le repère présenté en 7.c a déjà été rencontré : l'origine est située en bas à gauche et axe des x horizontal et axe des y vertical. Malheureusement, chaque logiciel utilise ses propres conventions qui sont plus ou moins bien décrites. Lors de l'export des mesures images, il faudra donc faire attention et au repère et à l'unité...

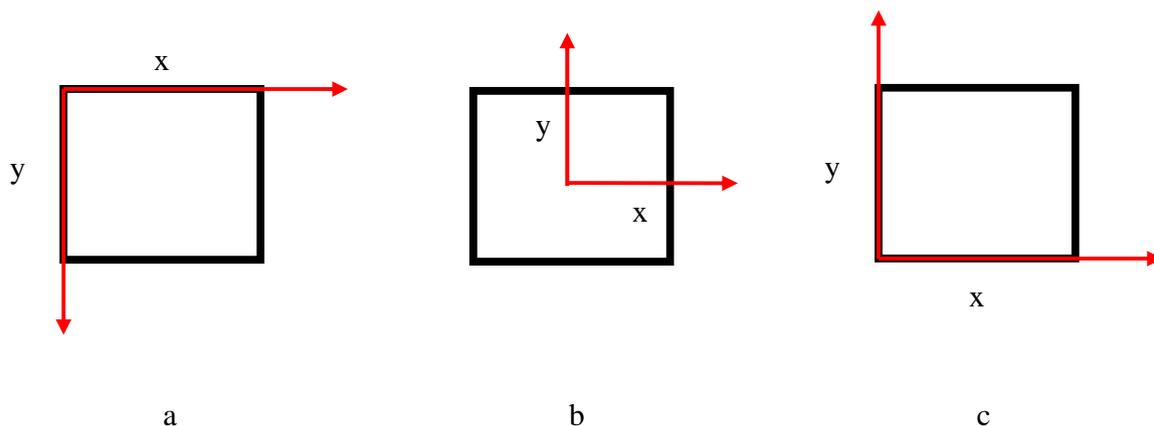


Fig. 7: Exemples de conventions de repères images existants à l'heure actuelle

- Géoréférencement d'images :
 - Orientation du cliché : positionner un objet dans l'espace n'est pas très compliqué, il suffit de connaître l'unité du repère et de donner les trois coordonnées X, Y et Z.
 - Orientation du cliché : historiquement les orientations étaient calculées à partir de 3 angles (convergence, site, déversement (ALBERTZ & KREILING 1980) ou encore oméga, phi et kappa). Pour obtenir la rotation globale, il suffit de faire le produit des 3 rotations. Malheureusement, le produit matriciel étant non commutatif, il convient de savoir dans quel sens les multiplier, sens qui est rarement communiqué par les logiciels... Ensuite, en fonction du repère image choisi (cf point précédent), la rotation ne sera pas exprimée de la même manière (et donc les angles non plus).

Il convient ensuite d'être au clair sur l'unité angulaire choisie : degré, radian, grade... là encore la documentation n'est pas toujours adaptée. Enfin il faut ensuite savoir si la rotation transmise permet de passer du repère terrain au repère image ou l'inverse. Il ne s'agit que de la transposée au sens mathématique mais rajoute une combinaison supplémentaire à toutes les options énumérées précédemment. Il faut généralement passer beaucoup de temps pour échanger des données d'un logiciel l'autre si l'export n'est pas prévu.

6 Produits de la photogrammétrie

Après avoir acquis puis géoréférencé précisément les images, nous allons maintenant détailler les différents produits réalisables : tout d'abord la restitution, les modèles numériques et enfin l'orthophoto.

6.1 Restitution

L'une des difficultés de l'acquisition par drone est qu'elle est peu propice à une restitution stéréoscopique traditionnelle. En effet, les images ne sont toujours parallèles et il peut y avoir des changements d'altitude qui provoquent un inconfort lors de la saisie. Par contre, il existe une autre solution : la saisie multi-images. Cela consiste à saisir l'objet pas uniquement dans 2 images mais dans beaucoup plus (figure 8). La saisie est faite dans chacune des images séparément : cela permet d'augmenter la précision de saisie dans la mesure où l'opérateur réalise la saisie dans des images prises avec de fortes différences de points de vue. Sur la figure 8.a, on peut voir les différents faisceaux (en orange) qui servent à la mesure du point et ceux qui pourraient être mesurés (en vert). Il faut veiller à ce que les faisceaux servant à la mesure soient répartis afin d'avoir une intersection franche qui permet une bonne détermination du point 3D.

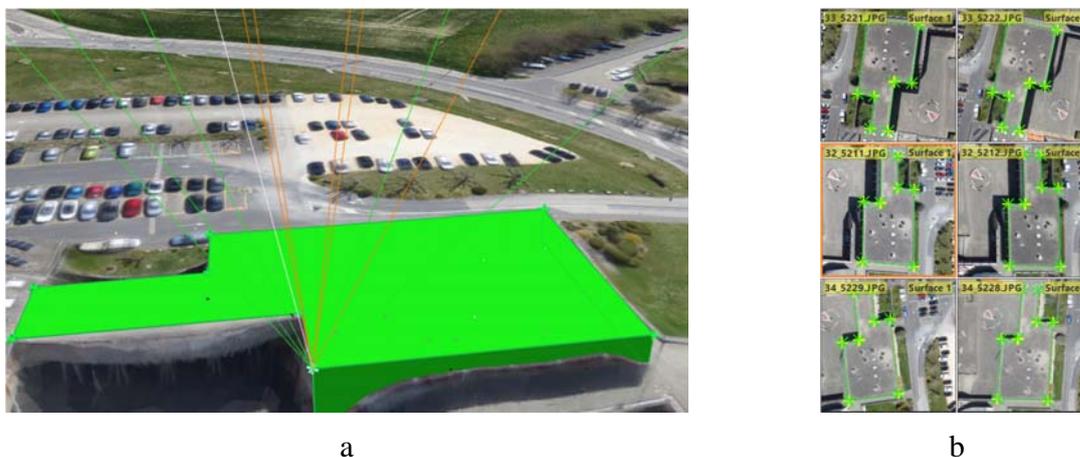


Fig. 8: Principe de restitution multi-images. (a) Faisceaux servant à la reconstruction, (b) Saisie de la surface dans 6 images

6.2 Modèle Numérique de Surface

Un des produits de la photogrammétrie par drone est le modèle numérique. Il s'agit d'un modèle numérique de surface car il décrit le terrain et les objets qui se trouvent dessus. La génération se

fait automatiquement grâce à des algorithmes de reconstruction comme par exemple (ROY & COX 1998).

Il existe plusieurs stratégies de génération mais le principe de base est de tester, pour une position planimétrique (X,Y) plusieurs valeurs d'altitude Z et de ne conserver que la valeur de Z pour laquelle la reprojection dans chacune des images est identique. Il peut donc y avoir des difficultés lorsque les zones sont homogènes, des motifs répétitifs ou encore lorsque les points de vues entre les images sont trop différents.

On peut, en général, définir le type de modèle numérique que l'on cherche à reconstruire : surface lisse (figure 9.b), surface normale (figure 9.c) ou surface accidentée (figure 9.d). Il faudra adapter le paramètre en fonction du résultat souhaité : calcul de volumétrie, calcul hydraulique, orthophoto... Un modèle « lisse » aura tendance à éliminer les petits reliefs (figure 9.b) alors qu'un modèle accidenté présentera des artefacts liés aux calculs et non pas à la géométrie du terrain (figure 9.d).

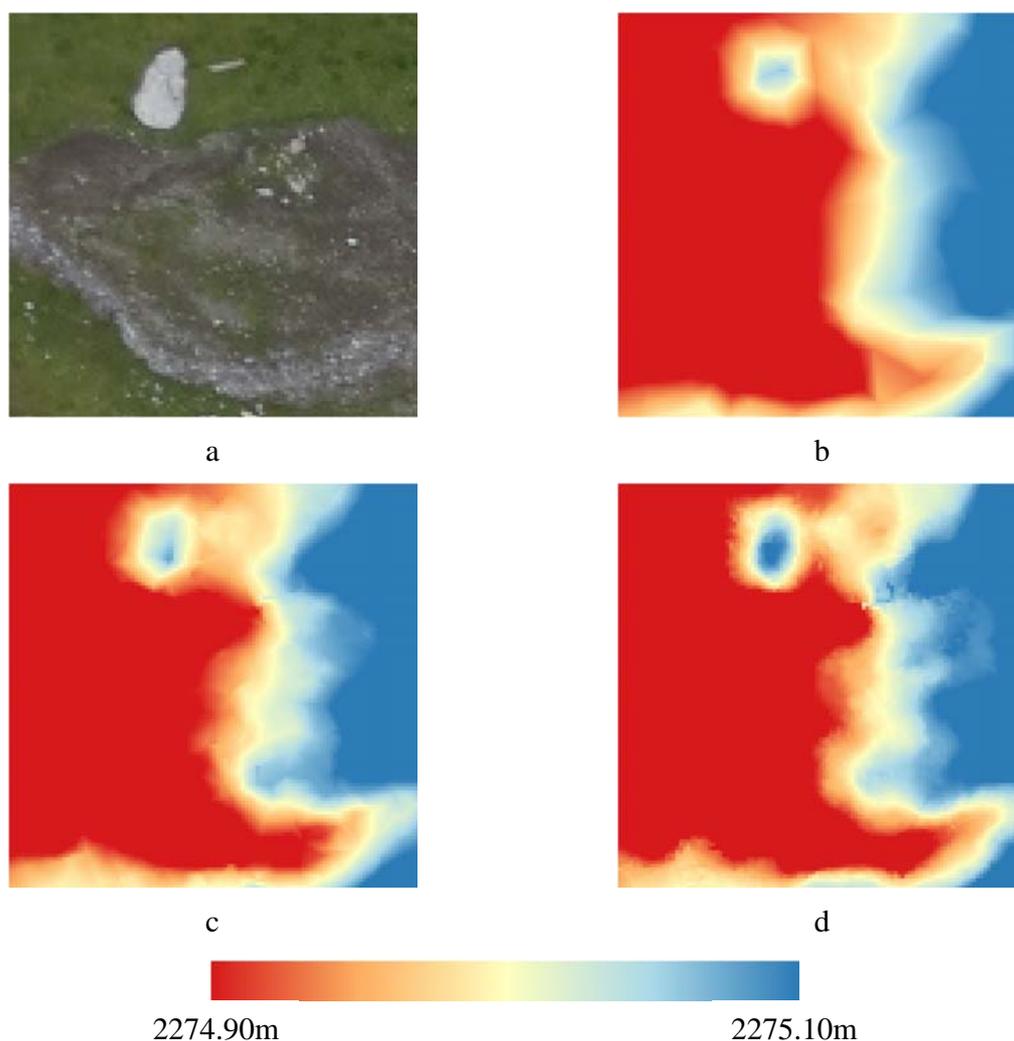


Fig. 9: Influence des paramètres sur le modèle numérique de surface. (a) orthophoto, (b) modèle lisse, (c) modèle moyen, (d) modèle accidenté

8 References

- AIBOTIX, 2016: <https://www.aibotix.com/de/aibot-x6.html>, accédé le 28/04/2016.
- ALBERTZ, J. & KREILING, W., 1980: Photogrammetrisches Taschenbuch. Photogrammetric Guide. Guide Photogrammetrique. Manual Fotogrametrico de Bolsillo. Wichmann, Karlsruhe.
- DJI, 2016: <http://www.dji.com/fr>, accédé le 28/04/2016.
- EXIV2, 2016: <http://www.exiv2.org/getting-started.html>, accédé le 28/04/2016.
- FISCHLER, M.A. & BOLLES, R.C., 1981: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *ACM* **24**, (6), 381-395.
- KASSER, M. & EGELS, Y., 2001: Photogrammétrie numérique. Paris, Hermès sciences publications, 2001, 379 p.
- LOWE, D.G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* **60** (2), 91-110.
- PARROT, 2016: <https://store.parrot.com/fr/drones/515-bebop-2.html#/couleur-white>, accédé le 28/04/2016.
- ROSTEN, E. & DRUMMOND, T., 2006: Machine learning for high-speed corner detection. *European Conference on Computer Vision*, 430-443.
- ROY, S. & COX, I., 1998: A Maximum-Flow Formulation of the N-camera Stereo Correspondence Problem, dans : *International Conference on Computer Vision*, Bombai 1998, 142.
- SENSEFLY, 2016: <https://www.sensefly.com/home.html>, accédé le 28/04/2016.

Photogrammetric Recording and Modelling of a Valley Floor using RPAS-mounted RGB and TIR Cameras

KONRAD EDER¹, LUDWIG HOEGNER¹, WOLFGANG RIEGER², MARIA KAISER² & UWE STILLA¹

Zusammenfassung: In diesem Beitrag wird ein Projekt zur 3D Modellierung und Auswertung eines Talbodens für die hydrologische Abflusssimulation vorgestellt. Hierfür wurde ein RPAS (Remotely Piloted Aerial System) System eingesetzt, um in aufeinander folgenden Überflügen eine Bildserie jeweils mit einer hochauflösenden photogrammetrischen Kamera und einer thermischen Infrarotkamera aufzuzeichnen. Aus den Bildern der photogrammetrischen Kamera wird ein Digitales Geländemodell erstellt und anschließend ein Ortho imagesaiki gerechnet. Die Bilder der Thermalkamera werden anschließend auf dieses Ortho image koregistriert. Hierdurch ist eine Lokalisierung und Identifikation von im Geländemodell nicht erfassten Oberflächenabflüssen möglich, was zur Verbesserung und Validierung hydrologischer Modelle eingesetzt werden kann.

Abstract: This contribution presents a project for 3D modelling and analysis of a valley floor for hydrological simulations. An RPAS (Remotely Piloted Aerial System) is used to mount either a high-resolution photogrammetric camera or a thermal infrared camera in two different, consecutive flights recording an image sequence. The images of the photogrammetric camera are used to generate both a digital terrain model and an orthoimage mosaic. The images of the thermal camera are coregistered to this orthoimage. This allows the localization and identification of surface water flow that is not derivable from the terrain model only. The results can be used to improve and evaluate process based hydrological models.

1 Introduction

Aerial Photogrammetry for generation of digital terrain models and orthoimages is a widely used method. Tie points are automatically detected to connect the images within a set to an image bundle block which is adjusted optimizing the exterior orientation parameters of the images. This procedure is transferable from the visible domain to the thermal infrared domain where the lower geometric resolution reduces the accuracy of the 3D reconstruction. Methods for joint analysis of images from the visible domain and thermal infrared have been introduced to remote sensing (NICHOL & LEE 2004; NICHOL 2005). Most of these methods rely more or less on a flat earth scenario and are invalid for mountainous areas as they do not include digital terrain models. The reconstruction of 3D geometry has been introduced to UAV / RPAS (unmanned aerial vehicle / remotely piloted aerial system) mounted cameras using feature points and onboard inertial measurement units for building reconstruction and cultural heritage (MAYER ET AL. 2012). This work also includes semi-global matching to derive dense point clouds from the whole pixel set of

¹ Technische Universität München, Photogrammetry and Remote Sensing, Arcisstrasse 21, D-80333 München, Germany, E-Mail: [Konrad.Eder, Ludwig.Hoegner, stilla]@tum.de

² Technische Universität München, Hydrology and River Basin Management, Arcisstrasse 21, D-80333 München, Germany, E-Mail: [Wolfgang.Rieger, Maria.Kaiser]@tum.de

the images instead of only using 3D coordinates for homologous points (HIRSCHMÜLLER 2008). 3D reconstruction from thermal images is investigated for terrestrial (HOEGNER & STILLA 2015) and RPAS mounted (WESTFELD et al. 2015) infrared cameras. The coregistration of the thermal infrared images to the RGB ortho image allows the generation of a thermal orthoimage and the interpolation of thermal infrared intensities for all pixels of RGB orthoimage. The resulting four-channel orthoimage can be used for segmentation and object detection.

In the field of hydrology, the question arises, in which part of a catchment surface runoff is generated and what ways it chooses to concentrate down to the valley ground and catchment outlet. Most of the water follows small runlets down the slopes. These runlets are mainly not included in standard digital terrain models, but are visible in high resolution aerial images. This allows generating a terrain model with a few centimeters grid size. Additionally, many rills show slightly differences in radiometric behavior and can be classified. A relevant percentage of the water is flowing down directly on the surface and these very small flows are mainly invisible in the geometric reconstruction and radiometric classification in the visible domain. In the thermal infrared, the cooling effect of surface water flow shows areas with water present or higher moisture caused by prior water flow. Integrating this information to the RGB image increases and refines the classification accuracy.

To identify the areas of surface runoff generation in a catchment and to quantify the amount of generated surface runoff, physically based hydrological models are used in hydrology. In this context, a suitable method is the Water Balance Simulation Model WaSiM (SCHULLA 1997; 2015), which has been extended in a way that the runoff concentration of surface flow can be determined by a cinematic wave approach (WINTER 2013). This contribution deals with the question how high-resolution aerial images and thermal infrared pictures can help to validate the predicted flow paths of WaSiM in the Sachenbach catchment.

2 Methodology

The integration of the thermal image information into the RGB orthoimage and the joint usage for hydrological surveys is split into four parts: In part one, the RGB images are combined into a bundle block. Part two uses detected feature points visible in both the RGB orthoimage and the thermal infrared images to overlay the infrared images to the orthoimage. Part three is a segmentation of possible water flow areas on the slopes and meadows. Part four is about the integration into hydrological modelling.

2.1 Ortho Image Generation

It is assumed that all images are taken with an end lap of at least 60% and a side lap of around 40%. The recording positions are given by GPS/INS information. On the ground, additional tie points are given to avoid mismatches caused by repeating structures or areas with a low number of strong natural tie points. These points are additionally used in chapter 2.2 for the co-registration with the thermal infrared images. A bundle adjustment is done using GPS/INS exterior orientations as estimates for the unknown exterior orientations, 3D object coordinates of ground control points (GCPs) from GPS measurements in the field and the 2D image coordinates of the tie points as observations (TRIGGS et al. 1999). The interior orientation is meant to be known and introduced as

additional observations that stay constant during the recordings. Having minimized the mean squares error of the back projection of the 3D tie points to the images, a dense matching is performed using the resulting improved exterior orientations of the images to derive a grid based digital terrain model. The digital terrain model is fit to the existing digital terrain model given from the Bavarian Authorities (BLBDV).

2.2 Coregistration of Thermal Images and Ortho Image

A coregistration of 2D image edges implies that the difference in the field of view and the exterior orientation of the RGB and the TIR image is small enough to ignore differences in the visibility and perspective of both images. It is also assumed that prominent edges appear in the RGB image as well as in the TIR image. The coregistration is a two-step process. In the first step, corresponding feature candidates have to be found. An affine transformation is calculated to minimize the search space for corresponding features in the TIR image and its corresponding RGB image. In a second step, the found corresponding feature points are taken to calculate the projection of the TIR image to the ortho image plane.

For the affine transformation, an edge detection based on the Sobel operator is done in both the scaled TIR and the RGB image followed by a nonmaximum-suppression. The different radiometric appearance of edges is mainly eliminated by this step. Based on the first estimation of the scale factor of the TIR image and the corresponding part of the RGB image, assuming a limited error in the exterior orientation, the coregistration is done using iterative estimation of the best transformation parameters in a similarity metric (STYNER et al. 2000). The algorithm results in an affine transformation of the TIR image to the RGB image. Based on this transformation it is assumed that the remaining differences are small enough to have a very reduced search space for corresponding point features in the TIR image and the RGB image. Following the argument, that prominent edges appear in both images, the Förstner operator (FÖRSTNER & GÜLCH 1987) is used to detect point features. The correspondences are found only based on the point positions looking for smallest distances of a homologous point candidate in the two images.

Being aware that an affine transformation is only a first estimation of the correct projection, especially in a mountainous area, the resulting homologous points of step one are now used to refine the exterior orientations of the TIR images. From step one feature points are known in the RGB images. Their 2D coordinates are projected onto the digital terrain model using the estimated exterior orientations of the RGB images from chapter 2.1. These 3D coordinates are now projected into the TIR image plane to minimize the distance of the projected RGB feature points and their corresponding TIR points in an adjustment using the 2D points as observations and the exterior orientation of the TIR image as unknowns. RANSAC (FISCHLER & BOLLES 1981) is used to remove outliers. The resulting estimated exterior orientations are used to do an ortho projection of the TIR image to add the TIR channel to the RGB ortho image.

2.3 Image Segmentation and Classification

The image classification is used to extract cold areas on the meadows with an orientation along the slopes. A supervised maximum-likelihood classification is used to classify meadow and forest from the RGB channels. A segmentation is done for the classified meadow areas to detect cold areas in the TIR channel. These cold areas are supposed to be areas with high moisture, which is

an indicator for water flow. These cold segments are supposed to be thin, long structures following the slope. Morphological operators are applied to remove gaps.

2.4 Hydrological Modelling

The Water Balance Simulation Model WaSiM is a deterministic, mainly physically based hydrologic catchment model. It allows the simulation of the water cycle above and below the land surface based on simplified physical process descriptions. WaSiM is applicable to numerous spatial and temporal scales. The model is hence suitable for both small and very large catchments. Furthermore, WaSiM can be used for short-term and long-term simulations. As input data information on precipitation, temperature, topography, land use and soil properties are needed. Due to the extensive selection of modules, WaSiM is suitable for solving manifold hydrologic questions (SCHULLA 1997). WaSiM has been extended by a surface routing module developed and presented by WINTER (2013). This module determines the runoff concentration of the generated surface runoff according to the kinematic wave approach. The surface routing module is suitable for micro and lower meso scale applications. The flow velocity is calculated by the Manning-Strickler-equation. Gradient and surface roughness highly influence the runoff concentration in small scale applications. Due to soil compaction or exceeding of storage capacity infiltration excess is ponding on the land surface. When the depression storages are exhausted, free water starts to flow off following the slope. Then the surface routing module passes the surface runoff from cell to cell depending on the flow direction. The surface water is routed per cell until the subbasin outlet is reached. Finally the runoff is transferred to the discharge routing of the river system (Fig. 1). The surface routing module even allows the occurrence of multiple flow paths.

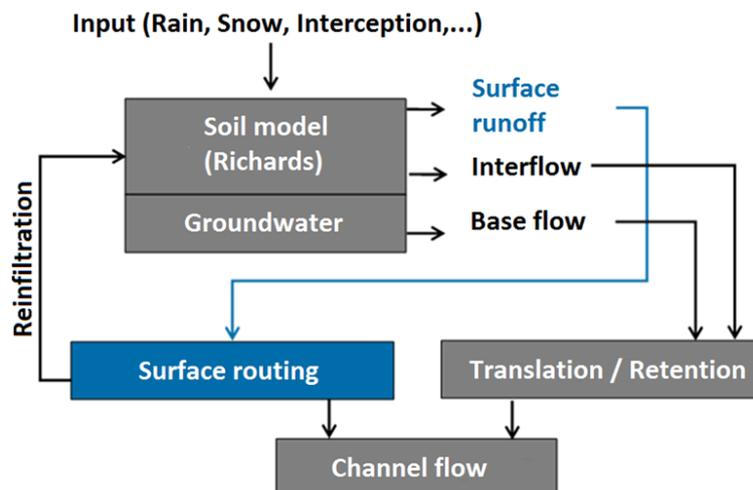


Fig. 1: Model approach: WaSiM with the implemented surface routing module (WINTER 2013)

The curvature of the area specifies whether the surface runoff is propagating in one or more flow directions. The flow velocity is computed using energy gradient, water film height and surface roughness. Subsequently, the amount of surface water flowing into each direction can be calculated. The type of cultivation determines the surface roughness and hence the flow resistance of the area. The delineation of the roughness coefficient is based on different empirical approaches.

The SR module even enables the simulation of re-infiltration and the influence of dynamic roughness and flow resistances on the flow velocity (WINTER 2013).

3 Experimental Setup

A flight campaign was carried out at the Sachenbach valley in the south of Bavaria (Fig. 2), where the meadow slope areas to both sides of the Sachenbach have been recorded. Figure 3a shows one exemplary flight path. For the RGB images a Sony Nex 7 was used. The Thermal infrared images were recorded using a FLIR Tau 640. Both cameras were mounted one after another onto an Ascending Technologies Falcon 8 (Fig. 3b) flying the same predefined path with both cameras.



Fig. 2: Overview of the test site. Left is the Walchensee lake. Only the meadow areas close to the valley floor are recorded. Forest areas are skipped.



Fig. 3: a) Overview of the test site. Left is the Walchensee lake. Only the meadow areas close to the valley ground are recorded. Forest areas are skipped. b) Falcon 8 with Sony Nex 7 mounted.

4 Results

The resulting image sequences contains image pairs in RGB (Fig. 4a) and TIR (Fig. 4b) for every recording position. Figure 3 shows one image pair of a meadow with two little brooks. The brook in the upper part of the images is clearly visible in both RGB and TIR. The one in the lower part of the images is hard to see in the RGB image but clearly visible in the TIR image. In the center of the TIR image, a few small dark structures following the slope can be seen that are moisture areas.

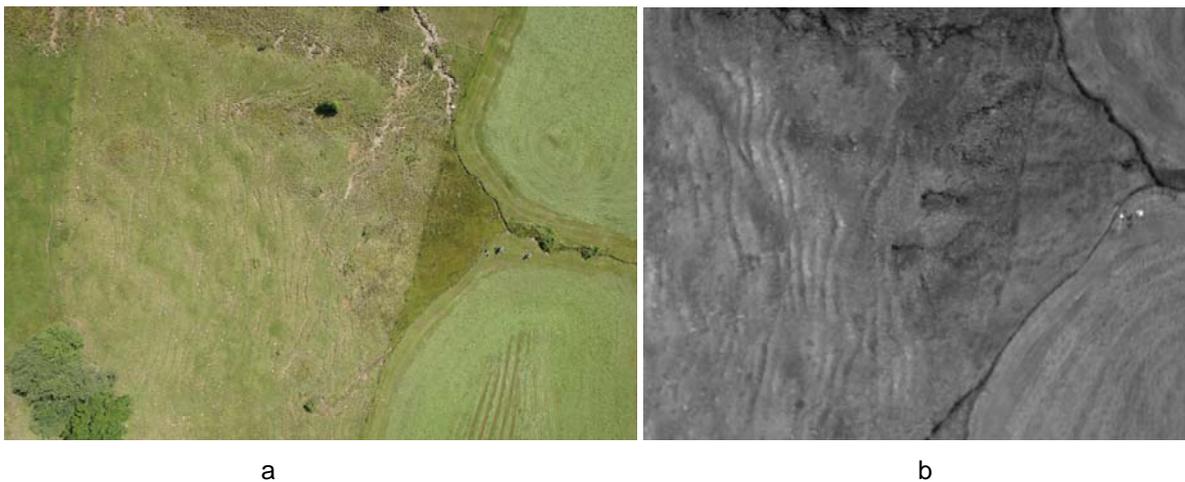


Fig. 4: Image pair a) RGB image: Only the brook in the upper part of the image is visible. b) TIR image: two brooks are visible and some additional moisture areas in the center of the image following the slope in the meadow.

4.1 Digital Terrain Model and Ortho Image Generation

A digital terrain model and an ortho image are derived from the bundle block of the RGB images. Figure 5 shows a 3D Reconstruction of the scene generated from the bundle block with dense matching (HIRSCHMÜLLER, 2008).

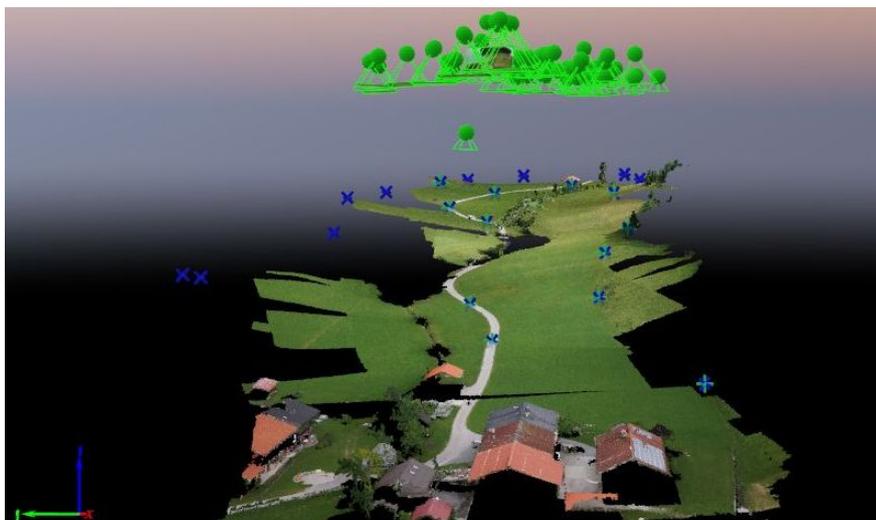


Fig. 5: 3D Reconstruction of the valley floor.

Gaps in the model are caused by trees and their different occlusions from different viewing positions. The digital terrain model is generated from the bundle block and fit to the existing thin terrain model obtained from the Bavarian Geodata Service. The resulting fused terrain model is shown in figure 6. In figure 6a the calculated dense terrain model parts derived from the bundle are shown as they fit into the given terrain model. Figure 6b shows the final fused terrain model as shaded relief model. The small squared blobs along the river are manually measured profiles of the river bed. The shaded relief model shows no artefacts along the borderlines due to the high accuracy of the generated point cloud.

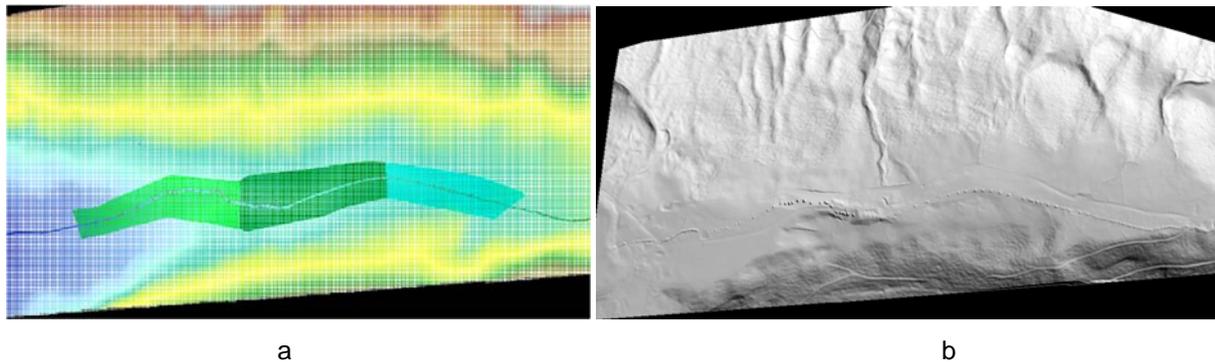


Fig. 6: Digital Terrain Model. a) Dense terrain model stripes fit into the thin terrain model. b) Fused terrain model including measured river bed profiles (small blobs).

4.2 Coregistration of Infrared Images and Classification

The coregistration is based on two steps. Step one is the estimation of an affine transformation to detect corresponding feature points in RGB and TIR images. A result is shown in figure 7a, where the RGB image is drawn in red and the coregistered TIR image in green. The brooks show a much higher contrast in the TIR image compared to the RGB image. River elements fit together quite well. It can be seen that in the meadow, RGB and TIR show slightly different structures. In the second step, the found corresponding feature points are used to refine the exterior orientations of the TIR images and to project the TIR images into the ortho image. Figure 7b shows the result of the second step for the image pair of figure 7a. The ortho image is drawn in RGB and overlaid with the projected TIR image in blue.

4.3 Hydrologically Modelling and Interpretation

The aerial image shows that the grassland is very heterogeneous. On the basis of structure and color different areas can be distinguished and severely eroded areas on the grassland can be discovered. Soil is removed by massive surface runoff and return flow. Thermal remote sensing provides information about the runoff behavior of an area. Due to the radiant temperature not only flowing waters but also water saturated soils and water outlets in the ground can be made visible. In this way, flow structures and runoff contributing areas can be visualized which are otherwise difficult or impossible to detect in the field or on aerial photographs. Therefore, areas with high surface runoff generation and surface runoff flow paths can be detected and defined as erosion prone areas (Fig. 8).

The surface runoff simulation of WaSiM with surface routing module is analyzed based on different precipitation events. Selected convective and advective precipitation events, as well as a flood event from June 2013 help evaluating the simulation of surface runoff. The surface runoff result grids provide interesting information about the runoff behavior of the basin during rain events. Convective precipitation events cause immediately high amounts of surface runoff in the Sachenbach catchment. One hour later, however, the overland flow is already significantly declined. During advective events it takes much longer until surface flow occurs on extensive areas. In return, the decline of the surface runoff is slow. WaSiM does not predict overland flow on the entire grassland area (Fig. 8). The largest part of the generated surface runoff is less than $0.009 \text{ m}^3/\text{s}$. Flow paths on which surface runoff accumulates to larger amounts are clearly recognizable by the darker colors. WaSiM predicts surface runoff on all plotted erosion prone areas (bordered in red). Shape and runoff amount of the two largest and most important erosion sites are remarkably well simulated. Furthermore, the lengthening or temporary formation of branches is pictured. In this context, a long temporary branch highlighted in dark purple is particular noteworthy. According to the WaSiM simulation, the highest amount of surface runoff flows off in this flow path. In summary, the surface runoff situation in the grassland area is well predicted by the surface routing model. WaSiM was even able to picture the largest erosion sites which were previously identified on aerial and thermal images.

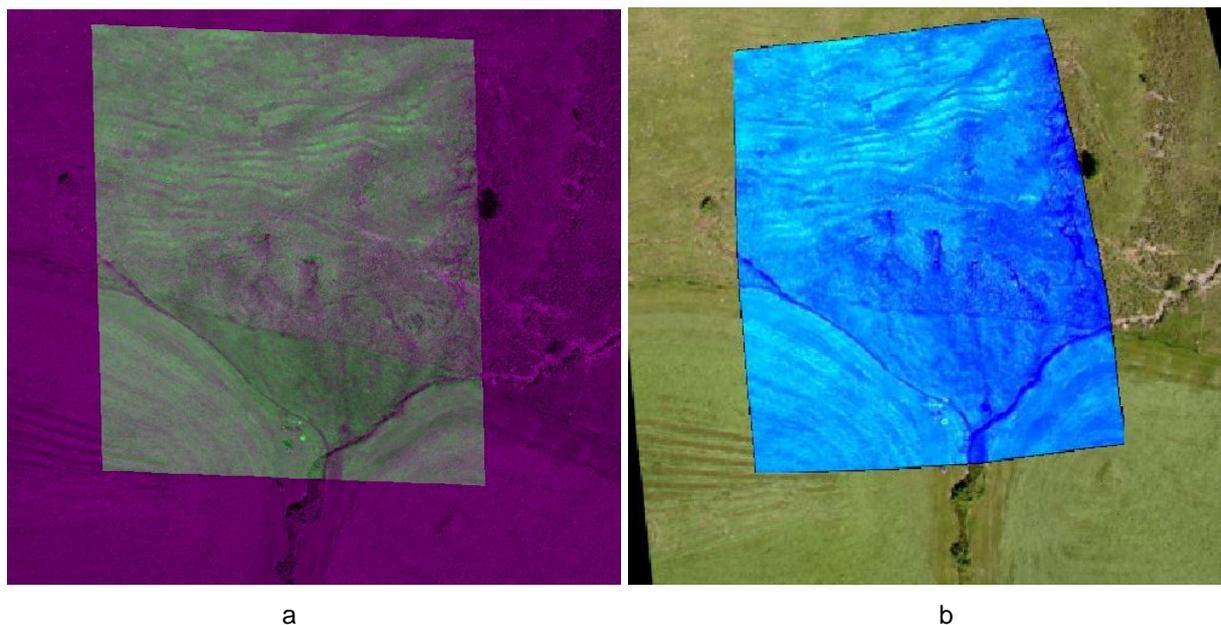


Fig. 7: RGB / TIR coregistration: a) Overlay of RGB image (red) and TIR image (green). The brooks fit together quite well. Structures on the meadow show slightly different intensities. b) TIR image (blue) projected onto the ortho image (RGB).

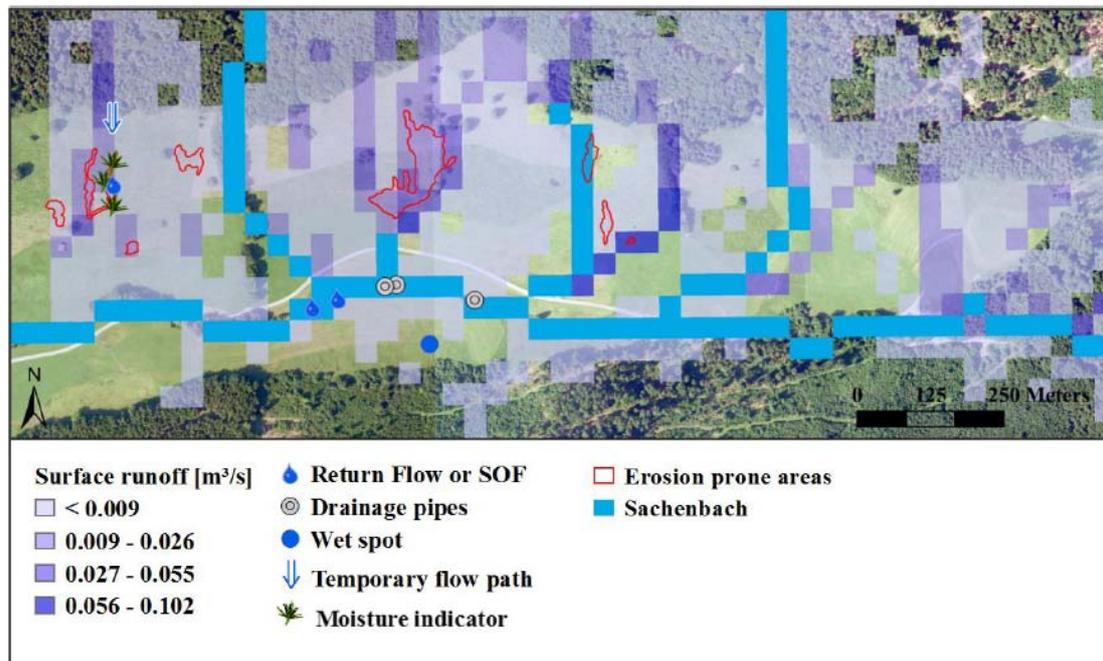


Fig. 8: Simulation of surface runoff combined with the results of the field investigation. The grassland area is shown at the time of the highest simulated surface runoff peak during the extreme weather event in June 2013 (02.06.201312:00)

5 Discussion and Outlook

In general, the coregistration of RGB and TIR images of landscape scenes is more complex compared to manmade scenarios with rectangular structures and straight lines. Nevertheless, a two-step process of coregistration based on feature correspondence search is possible. The number of reliable features visible in both the RGB images and the TIR images was relatively small so that the redundancy in the estimation of the exterior orientation of the TIR images is quite low. In the scenario used for this investigation the parameters of the affine transformation had to be adopted manually to the scene. This limits the methods to small bundle blocks. The flight path and recording orientation have been optimized for the RGB camera with sufficient overlap. As the TIR camera has a smaller field of view, the overlap for the TIR images was too small to directly calculate a bundle block of the TIR images, which could further improve the quality of the projection in two ways: First, the accuracy of the estimated orientation from the bundle block adjustment is better than the original recorded orientation parameters. Second, it would be possible to do a 3D reconstruction from the RGB and TIR images separately and coregister the whole image blocks using the 3D point clouds. Thermal photography enables the visualization of waterlogged areas, flow paths and sources. The application of thermal photography for hydrological issues is still unexplored but seems to be very promising. Using thermal images it could be confirmed that surface runoff occurs on the erosion sites in the Sachenbach basin.

6 References

- FISCHLER, M. & BOLLES, R., 1981. Random sample consensus: A paradigm for model fit-ting with applications to image analysis and automated cartography. *Communications of the ACM* **24** (6), 381-395.
- FÖRSTNER, W. & GÜLCH E., 1987. A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centers of Circular Features. *Proceedings of the ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data 1987*, 281-305.
- HIRSCHMUELLER, H., 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30** (2), 328-341.
- HOEGNER, L. & STILLA, U., 2015. Building facade object detection from terrestrial thermal infrared image sequences combining different views. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **II** (3/W4), 55-62.
- MAYER, H., BARTELTSEN, J., HIRSCHMUELLER, 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*, F. Dellaert, J.-M. Frahm, M. Pollefeys, L. Leal-Taixe & B. Rosenhahn (eds), Dagstuhl Castle, Germany, June 26-July 1, 2011, **7474**, Springer, Berlin, 285-304.
- NICHOL, J. & LEE, C.M., 2004. Urban vegetation monitoring in Hong Kong using high resolution multispectral images. *International Journal of Applied Earth Observation and Geoinformation* **26** (5), 903-918.
- NICHOL, J., 2005. Remote Sensing of Urban Heat Island by Day and Night. *Photogrammetric Engineering & Remote Sensing* **71** (5), 613-621.
- SCHULLA, J., 1997: Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaänderungen. ETH Zürich, Dissertation.
- SCHULLA, J., 2015: Model Description WaSiM. Water balance Simulation Model. Zürich.
- STYNER, M., BRECHBUEHLER, C., SZKELY, G. & GERIG, G., 2000. Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Transactions on Medical Imaging* **19** (3), 153-165.
- TRIGGS, B., MCLAUCHLAN, P., HARTLEY, R. & FITZGIBBON, A., 1999. Bundle Adjustment – A Modern Synthesis. *Proceedings of the International Workshop on Vision Algorithms ICCV '99*, Springer-Verlag, 298-372.
- WESTFELD, P., MADER, D. & MAAS, H., 2015. Generation of TIR-attributed 3D Point Clouds from UAV-based Thermal Imagery. *Photogrammetrie Fernerkundung Geoinformation* (5), 381-394.
- WINTER, F., 2013: Prozessorientierte Modellierung der Abflussbildung und -konzentration auf verschlammungsgefährdeten landwirtschaftlichen Nutzflächen. Universität der Bundeswehr, München, Dissertation.

Entwicklung einer zeitoptimierten automatischen Extraktion von Wasserflächen aus hochauflösenden Luftbilddaten (True Ortho Mosaik, Digitales Oberflächenmodell)

ANJA GÄNSRICH¹, STEVEN BAYER² & TILMAN BUCHER²

Zusammenfassung: Für die pixelscharfe Erfassung und Korrektur von Wasserflächen in photogrammetrisch erzeugten hochauflösenden digitalen Oberflächenmodellen wurde ein automatisiertes objektbasiertes Verfahren entwickelt. Zur Verkürzung der Laufzeit wurde diesem die automatische Berechnung einer Wassererwartungsmaske vorangestellt, welches Regionen, die potentiell Wasser beinhalten, selektiert und somit das Datenvolumen bis >60 % reduziert.

Der objektbasierte Ansatz nutzt Informationen aus dem Bildmatching (Zuverlässigkeit, Höhe), spektrale Indizes und Textur- und Formparameter, es extrahiert die Wasserflächen mit einer Genauigkeit von >90%. Durch die Vielzahl von Eingangsdaten ist das Ruleset sehr robust und sensorunabhängig. Um die Laufzeit weiter zu verringern (>50%), werden vereinfachte Segmentierungs- und Stapelverarbeitungsalgorithmen auch auf niedrigeren Bildpyramiden eingesetzt.

1 Einleitung

Die Berechnung und Weiterverarbeitung von kostengünstigen photogrammetrisch erzeugten Digitalen Oberflächenmodellen (DOM) stehen in zunehmendem Interesse (vgl. MEIJERINK & MANNAERTS 2000; MATTIKALLI & ENGMAN 2000). Bei der photogrammetrischen Erstellung von DOMs kommt es zu Artefakten (ZHANG & GRUEN 2006): Gewässer sind meist sehr homogen, haben keine eindeutige Textur, sie unterliegen ständiger Bewegung und ihre Reflektions-eigenschaften sind richtungsabhängig. Dies führt im Bildzuordnungsprozess zu Fehlern und Lücken. Die Erstellung einer Wassermaske kann diese Fehler bereinigen und nachfolgende Analysen erleichtern (vgl. MAUNE & DEWBERRY 2010). Ein automatisiertes Verfahren zur Klassifizierung von Wasseroberflächen, das uferkantengenau und mit möglichst wenig Prozessier- und Zeitaufwand arbeitet, wurde bisher in der Literatur kaum behandelt. Es ist stets ein hoher manueller Aufwand zur Bereinigung von Fehlklassifizierungen in Schatten- und Straßenbereichen notwendig, um eine gute Wassermaske zu erhalten. Die Modelle besitzen in der Regel nur regionale Gültigkeit und sind nicht global übertragbar.

2 Kontext und Ziel

Zur Wasserklassifikation werden vor allem Satellitenbilddaten unterschiedlicher Auflösungen genutzt. Ansätze von HERRERA-CRUZ & KOUDOGBO (2009), HAHMANN et al.

¹ Beuth Hochschule für Technik Berlin, Luxemburger Straße 10, D-13353 Berlin, E-Mail: anja-gaensrich@gmx.de

² German Aerospace Center (DLR), Institute of Optical Sensor Systems, Rutherfordstraße 2, D-12489 Berlin, E-Mail: [steven.bayer, tilman.bucher]@dlr.de

(2009), UCA AVCI et al. (2014), DRONOVA et al. (2011) u.a. sollen für die vorliegenden hochauflösenden Luftbildaufnahmen bewertet und verfeinert werden. Die Übertragbarkeit der in diesen Studien vorgestellten Ansätze ist nicht gegeben. Bis auf bei HAHMANN et al. (2009) wurden die entwickelten Verfahren nur regional eingesetzt und nicht mit anderen Sensorsystemen getestet.

Ziel dieser Arbeit ist es, einen objektbasierten Workflow zu entwickeln, der automatisiert pixelscharfe Wassermasken in einem beliebigen Untersuchungsgebiet auf Basis von hochauflösenden True Ortho Luftbild Mosaiken (TOM) und photogrammetrisch erzeugter DOM ableitet. Der Workflow soll sensorunabhängig ohne Anpassung funktionieren.

Aufgrund großer Datenvolumen wird zusätzlich ein Modell entwickelt, das initial das Untersuchungsgebiet analysiert und eine Wassererwartungsmaske (WEM) identifiziert. Mit Hilfe dieser Maske kann der Bearbeitungsraum und somit der Zeit- und Prozessierungsaufwand für die objektbasierte Bildauswertung verringert werden. Die Eignung von Texturparametern soll evaluiert werden, um möglichst unabhängig von den spektralen Eigenschaften der Sensoren einen einheitlichen und robusten Workflow aufzubauen.

3 Datengrundlage

Für die Analyse von Gebieten mit unterschiedlichen Typen von Oberflächengewässern (Fließ- & Stillgewässer, Küste) wurden vom Deutschen Zentrum für Luft und Raumfahrt, Abteilung Sensorkonzepte und Anwendungen, Daten zur Verfügung gestellt. Ausgewählt wurden drei Kamerasysteme und unterschiedliche Geländeformen: eine Digital Modular Camera II (DMC II, NEUMANN 2011) – voralpines Gebiet München, eine UltraCamX (UCX, GRUBER et al. 2008) - Urstromtal Berlin und ein Modular Airborne Camera System (MACS, LEHMANN et al. 2011) – Nordseeinsel Helgoland. Die Datensätze (Beispiel Berlin in Abb. 1) bestehen jeweils aus einem TOM (RGBI), dem DOM und dessen Nebenprodukt, der Matching Count Mask (MCM). Die MCM wird bei der Generierung eines DOMs mit dem Semi-Global-Matching Verfahren ausgegeben und gibt die Anzahl der korrespondierenden Punktepaare eines DOM-Pixels wieder, was wiederum ein Maß für die Verlässlichkeit ist. Die MCM ist sowohl von der Befliegungsplanung wie auch vom Untergrund abhängig.

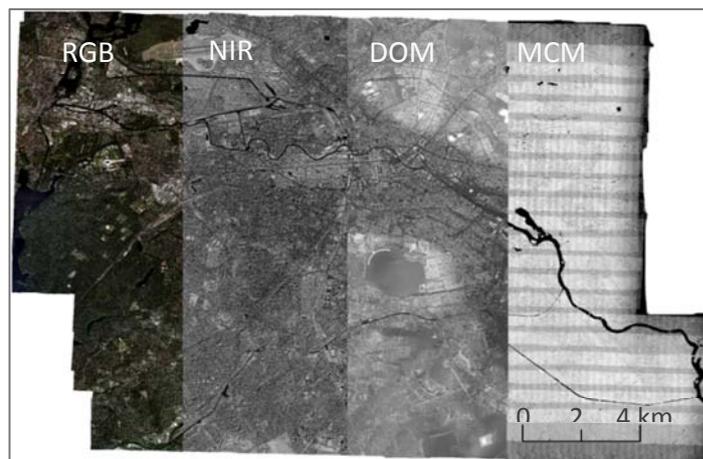


Abb. 1: Datensatz Berlin (UCX, 2010, DLR) bestehend aus TOM (RGBI), DOM und MCM

4 Methodik

4.1 Wassererwartungsmaske (WEM)

Die Erstellung einer WEM kann das Datenvolumen und die Verarbeitungszeit für eine Bildanalyse deutlich reduzieren. Das Modell, das sich in eine hydrologische Simulation und eine Auswertung der Matching-Statistik aufteilt, arbeitet automatisch, so dass der Aufwand gering ist. Um potentielle Artefakte bei der Generierung möglicher Wasserläufe zu bereinigen werden die DOM geglättet. Das Glättungsverfahren nach JACKSON (2014) erwies sich als am geeignetsten. Auf Grundlage der Strömungsrichtung wird die Abflussakkumulation bestimmt. Nach STRAHLER (1957) werden die Abflusslinien mit den meisten Zuflüssen selektiert.

Die MCM gibt Aufschluss darüber, an welchen Pixeln keine korrespondierenden Punktepaare gefunden worden sind. Fehlende Matchings deuten u.a. auf bewegte Objekte (z. B. Autos, Wellenkämme) oder sehr homogene Flächen (z.B. ruhige Wasserflächen oder Schattenbereiche) hin. Als mögliche Wasserflächen werden Pixelgruppen $> 300 \text{ m}^2$ kategorisiert.

Auf Grundlage der Verschneidung dieser Informationen wird eine WEM generiert (Abb.2).

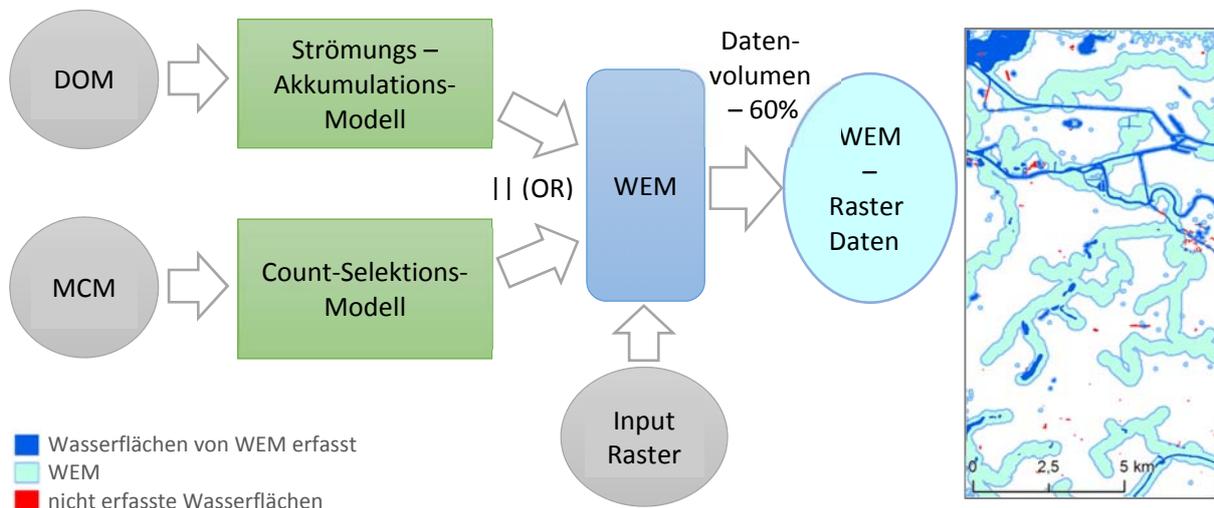


Abb. 2: Modell der Wasserwartungsmaske mit Beispielausschnitt aus dem Datensatz Berlin

4.2 Objektbasierte Bildanalyse

In der Arbeit wurde die Software Trimble eCognition Developer 8.64 genutzt. Um große Daten in möglichst kurzer Zeit auf mehreren Computerkernen zu analysieren, werden die Daten automatisch in Kacheln zerlegt, segmentiert, klassifiziert und wieder zusammengefügt.

Um rechenintensive Algorithmen zu vermeiden, wurde eine Kombination einfacher Segmentierungsverfahren (Quadtree Segmentierung, Spectral Difference Segmentierung) mit unterschiedlicher Gewichtung komplexeren Segmentierungen vorgezogen. Weiterhin werden einzelne temporäre Arbeitsschritte in niedrigeren Bildpyramiden (Faktor 4) durchgeführt. Dadurch konnten in der Hälfte der Prozessierungszeit kompakte Segmente generiert werden. Durch erweiterte Regiongrowing-Verfahren werden die Objektgrenzen in voller Auflösung nach der Klassifizierung verfeinert, so dass uferkantengenaue Ergebnisse die manuelle

Nachbereinigung auf ein Minimum reduzieren. Um falsch positive Klassifizierungen auf spektral ähnlichen Oberflächen zu vermeiden, werden temporäre Hilfsklassen (Schatten, Asphalt, Gebäude, Vegetation) angelegt.

Als wichtigste Eigenschaften wurden genutzt (Tab. 1):

Tab. 1: Überblick der Objekteigenschaften, die für Segmentierung und Klassifizierung genutzt wurden

Objekt	Eigenschaften / Index
Wasser dunkel	<ul style="list-style-type: none"> • Große Homogenität (Texturparameter, GLCM) • hohe Absorption im NIR (NDWI (MCFEETERS 1996), LWM (UDDIN 2010)) • niedriges MCM
Wasser hell (Gischt, Schaum, spiegelnd)	<ul style="list-style-type: none"> • Nachbarschaftsbeziehungen Wasser dunkel • Hohe Varianz und Helligkeit • HSI Intensität (IR, G und B) ↔ helle Dächer
Schatten	<ul style="list-style-type: none"> • Blue-Red Ratio • Nachbarschaft zu erhöhten Objekten
Gebäude	<ul style="list-style-type: none"> • Höhe im DOM, Fläche • helle Dächer, Glasreflektion: HSI Intensität (IR, G und B)
Hohe Vegetation	<ul style="list-style-type: none"> • NDVI (TUCKER 1979) , DOM

5 Ergebnisse und Diskussion

Es wurde objektbasierter Ansatz entwickelt, der ohne individuelle Anpassung Wasserflächen mit einer Nutzergenauigkeit von 93-95% aus hochauflösenden Fernerkundungsdaten unterschiedlicher Sensoren (UltraCamX, DMC II, MACS) extrahiert (siehe Tab. 2 und Abb. 3). Diese Ergebnisse übertreffen die Ergebnisse aus den Studien von HAHMANN et al. (2009), UCA AVCI et al. (2014) oder HERRERA-CRUZ & KOUDOGBO (2009), die Satellitenbilddaten als Grundlage nutzen, jedoch sensorspezifisch ausgelegt sind.

Tab. 2: Ergebnisse der objektbasierten Klassifikation von Wasser

Datensatz	Nutzergenauigkeit in %	Produzenten- genauigkeit in %	Gesamtklassifikations- genauigkeit in %	Kappa- Koeffizient
MACS Hale	97,78	93,62	95,56	0.9111
UCX	95,10	100	97,37	0.9473
DMC II	96,77	100	98,33	0.9667

Mit wenigen Einschränkungen ist es gelungen, Wasserflächen nahezu uferkantengenau zu extrahieren, auch flache Gewässer mit hohem Vegetations- und Schwebstoffanteil wurden größtenteils korrekt erfasst. Zu den Fehlklassifikationen zählen Schatten und spektral und textuell ähnliche versiegelte Flächen und Dächer. Vereinzelt nicht erfasst wurden zudem sehr flache Uferregionen, kleine überschattete oder stark mit Gischt und Schaum bedeckte Wasserflächen. Um noch bessere Klassifizierungsergebnisse zu generieren, wäre es notwendig,

die Regelbasis an das jeweilige Untersuchungsgebiet und die Eigenschaften der verwendeten Sensorik direkt anzupassen. Zugunsten einer sensorübergreifenden Kompatibilität wurde darauf verzichtet.

Um die die Algorithmen sensorunabhängig einsetzen zu können, standen die Texturparameter und Nachbarschaftsbeziehungen bei der Entwicklung der Regelbasis im Vordergrund. Es konnte jedoch nicht gänzlich auf multispektrale Eigenschaften verzichtet werden. Insbesondere die Absorption im NIR ist für eine klare Abgrenzung von Wasser hilfreich.

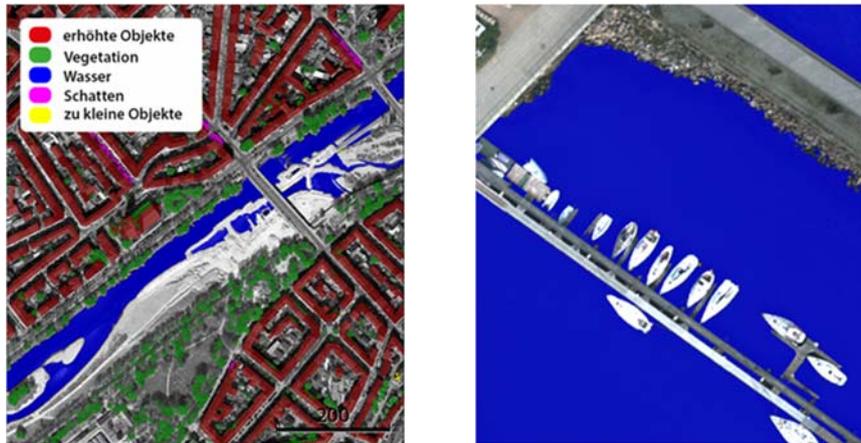


Abb. 3: Beispielhafte Klassifizierungsergebnisse mit pixelscharfen Grenzen

Die Prozessierungszeit wurde durch ein kombiniertes Modell aus vollautomatischer hydrologischer Simulation in einem GIS und einer statistischen Auswertung der MCM zur Generierung einer Wassererwartungsmaske (WEM) reduziert. Im Datensatz von Berlin werden 96,5% der Gewässer durch die WEM erfasst, die fehlenden Gewässer sind überwiegend anthropogen beeinflusst. Die Datengrundlage wurde mittels der WEM durchschnittlich um 60% reduziert, was eine Verkürzung der Laufzeit der objektbasierten Analyse um mehr als 50% bewirkt.

6 Acknowledgement

Für die Bereitstellung der Daten und die Kooperation bedankt sich die erste Autorin beim DLR-Institut für Optische Sensorsysteme, Abteilung Sensorkonzepte und Anwendungen, Berlin.

7 Literaturverzeichnis

- DRONOVA, I., GONG, P. & WANG, L., 2011: Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at Poyang Lake, China. *Remote Sensing of Environment* **115** (12), 3220-3236.
- GRUBER, M., PONTICELLI, M., BERNÖGGER, S. & LEBERL, F., 2008: UltracamX, the Large Format Digital Aerial Camera System by Vexcel Imaging / Microsoft. *The International Archives*

- of the Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B1), 665-670.
- HAHMANN, T., MARTINIS, S., TWELE, A. & BUCHROITHNER, M., 2009: Strategies for the automatic mapping of flooded areas and other water bodies from high resolution TerraSAR-X data. *Cartography and Geoinformatics for Early Warning and Emergency Management: Towards Better Solutions*, 207-214.
- HERRERA-CRUZ, V., KOUDOGBO, F. & HERRERA, V., 2009: TerraSAR-X rapid mapping for flood events. *Proc. Int. Soc. Photogramm. Rem. Sens. (Earth Imaging for Geospatial Information)*, 170-175.
- JACKSON, S., 2014: Optimized Pit Removal. Center for Research in Water Resources. <http://tools.crwr.utexas.edu/OptimizedPitRemoval/CRWR%20Tools%20Optimized%20Pit%20Removal.html>, letzter Zugriff 2015-05-10.
- LEHMANN, F., BERGER, R., BRAUCHLE, J., HEIN, D., MEIBNER, H., PLESS, S., Strackenbrock, B. & Wieden, A., 2011: MACS - Modular Airborne Camera System for generating photogrammetric high-resolution products. *Photogrammetrie Fernerkundung Geoinformation*, (6), 435-446.
- MATTIKALLI, N.M. & ENGMAN, E.T., 2000: Integration of remotely sensed data into geographical information systems. *Remote Sensing in Hydrology and Water Management*, Springer Berlin Heidelberg, 65-82.
- MAUNE, D.F. & DEWBERRY, C., 2010: Digital Elevation Model (DEM) Whitepaper NRCS High Resolution Elevation Data. Dewberry, Arlington, 3-19.
- MCFEETERS, S.K., 1996: The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing* **17** (7), 1425-1432.
- MEIJERINK, A.M.J. & MANNAERTS, C.M., 2000: Introduction to and general aspects of water management with the aid of remote sensing. *Remote sensing in hydrology and water management*, Schultz, G.A. & Engman, E.T. (eds.), Springer, Berlin, 329-356.
- NEUMANN, K., 2011: The Z/I DMC II – Imaging Revolution. *PhotogrammetricWeek '11*, Fritsch, D. (ed.), Wichmann Verlag, 97-101.
- STRAHLER, A. N., 1957: Quantitative analysis of watershed geomorphology. *Transactions of the American Geophysical Union* **38** (6), 913-920.
- UCA AVCI, Z.D., KARAMAN, M., OZELKAN, E., KUMRAL, M. & BUDAKOGLU, M., 2014: OBIA based hierarchical image classification for industrial lake water. *Science of the Total Environment* **487**, 565-573.
- UDDIN, K., 2010: Introduce variable indices using landsat image. *GIS and Remote Sensing Analyst at International Centre for Integrated Mountain Development*. <http://www.slideshare.net/kabiruddin/introduce-variable>, letzter Zugriff 2015-02-14.
- ZHANG, L. & GRUEN, A., 2006: Multi-image matching for DSM generation from IKONOS imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* **60** (3), 195-211.

Konzeption und geometrische Bewertung der flächenhaften Berechnung von Digitalen Oberflächenmodellen aus Luftbildern

MARIUS RÖDER¹, ANNETT HEUSINGER² & ANSGAR BRUNN¹

Zusammenfassung: Das Land Bayern erstellt zurzeit einen flächendeckenden 3D-Gebäudebestand. Die Erfassung erfolgt aktuell aus Punktwolken des flugzeuggetragenen Laserscannings. Die Verwendung von Digitalen Oberflächenmodellen (DOM) aus photogrammetrischen Befliegungen wird zurzeit diskutiert. In diesem Beitrag wird daher die Generierung von digitalen Oberflächenmodellen aus Luftbilddaten der regelmäßigen Befliegung der Stadt Würzburg untersucht. Zum Einsatz kommt die Software „Erda Imagine 2014“, die den Semi-Global Matching Algorithmus (SGM) implementiert. Das Verfahren wird unter wirtschaftlichen, die erzielten Punktwolken unter geometrischen Aspekten beurteilt. Zur Bewertung der geometrischen Qualität der Ergebnisse werden die erzeugten Punktwolken mit LoD2-Daten der Bayerischen Vermessungsverwaltung überlagert und verglichen. Die Ergebnisse zeigen eine hohe geometrische Qualität in radiometrisch gut strukturierten Bereichen. Defizite sind in verschatteten Modellbereichen und teilweise in der Laufzeit der Berechnungen vorhanden.

1 Motivation

„Die Zukunft ist dreidimensional“ – unter diesem Titel erschien ein Artikel in der Schriftenreihe des Deutschen Vereins für Vermessungswesen (DVW) im Februar 2014 (HÜMMER & ROSCHLAUB 2014). Darin wird auf die Initiative zum Aufbau eines landesweiten 3D-Gebäudebestands der Bayerischen Vermessungsverwaltung eingegangen. In Zukunft soll gar das komplette Liegenschaftskataster Bayerns dreidimensional geführt werden. Dabei spielt die Erstellung von Digitalen Oberflächenmodellen (DOM) aus Luftbildern eine wichtige Rolle. Daraus können Folgeprodukte wie 3D-Gebäudemodelle abgeleitet oder DGM partiell fortgeführt (KREY et al. 2016) werden. Dies mindert wiederum die Durchführung des kostenintensiveren (TILSNER 2012) flugzeuggetragenen Laserscannings (engl.: Airborne Laserscanning (ALS)). Aber nicht nur für die Bayerische Vermessungsverwaltung, sondern auch für Ingenieurbüros und Städte bzw. Kommunen werden die DOM aus Luftbildern als Planungsgrundlage immer wichtiger. Ohne die dreidimensionalen Geoinformationen eines DOM können beispielsweise bauliche Aspekte mit Höhenbezug (z. B. Lärmschutz, etc.) in kommunalen Planungen nur unzureichend behandelt werden (HÜMMER & ROSCHLAUB 2014). Nicht zuletzt deshalb soll in diesem Beitrag auf die Erstellung eines Konzepts zur Generierung von DOM aus Luftbildern am Beispiel Würzburg eingegangen werden und die Genauigkeit der DOM bewertet werden.

¹ Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Labor für Photogrammetrie & Laserscanning, Röntgenring 8, D-97070 Würzburg, E-Mail: marius.roeder@web.de, ansgar.brunn@fhws.de

² Stadt Würzburg, Tiefbauamt, Geodaten und Vermessung, Veitshöchheimer Straße 1, D-97080 Würzburg, E-Mail: annett.heusinger@stadt-wuerzburg.de

2 Digitale Oberflächenmodelle

Die in der Vergangenheit am meisten verwendete Methode, große Punktmengen der Erdoberfläche zur Erstellung von DOM zu erfassen, ist das ALS. Für diesen Beitrag ist das ALS insofern von Bedeutung, als dass die geometrische Bewertung der DOM aus Luftbildern durch einen Vergleich mit LoD2-Daten stattfindet. Diese wurden hauptsächlich aus ALS-Daten abgeleitet (LDBV 2015).

Eine weitere, häufig angewandte Methode, große Punktmengen von der Erdoberfläche zu erzeugen, ist die Punktgenerierung durch Bildkorrelation (KERSTEN et al. 2001). Die Erzeugung von DOM aus Luftbildern basiert im Wesentlichen auf dem Verfahren der Stereophotogrammetrie. Ausgangsdatensatz ist hier im Allgemeinen ein Verband von digitalen Luftbildern, der mit speziellen Luftbildkameras von einem Messflugzeug aus aufgenommen wurde (STOLZ 2013).

Eine wichtige Komponente der Bildkorrelation ist die Einbeziehung der Epipolargeometrie. Dadurch wird versucht, die zweidimensionale Korrelation in eine eindimensionale Korrelation zurückzuführen (KRAUS 2004). Die Begriffe Disparität und Disparitäts- bzw. Tiefenbild stehen in engem Zusammenhang zur Epipolargeometrie und sind für den später erläuterten SGM-Algorithmus von besonderer Bedeutung.

Die Identifikation von korrespondierenden Pixeln in zwei oder mehreren Bildern ist ein Zuordnungsproblem. Die Zuordnung vieler Millionen korrespondierender Pixel wird heute durch effiziente Bildkorrelationsverfahren ermöglicht. Jedes Stereobildpaar wird dabei zunächst in ein Ausgangsbild (Referenzbild) und ein Suchbild eingeteilt. Dann wird versucht, jedem Pixel des Ausgangsbildes das korrespondierende Pixel im Suchbild zuzuordnen. Die Zuordnung basiert auf der Berechnung eines Ähnlichkeitswertes. Daraus kann abgeleitet werden, wie ähnlich sich die beiden untersuchten Pixel sind (STOLZ 2013).

Für die Berechnung eines Ähnlichkeitswertes gibt es verschiedene Korrelationsstrategien. Die populärsten Zuordnungsverfahren sind die flächenbasierte Zuordnung (area-based matching) und die merkmalsbasierte Zuordnung (feature-based matching). Vor wenigen Jahren erfolgte dann ein maßgeblicher technischer Fortschritt mit der Zuordnung auf Basis globaler Kostenfunktionen. Dabei berechnet eine Kostenfunktion einen Kostenwert für ein einzelnes Pixel im Ausgangsbild und ein Pixel im Suchbild (pixelwise matching). Im Rahmen der Bachelorarbeit (RÖDER 2015) wurden die DOM für diesen Beitrag ausschließlich mit Semi Global Matching (SGM) berechnet - einem pixelbasierten Bildzuordnungsverfahren. Der SGM-Algorithmus wurde erst vor wenigen Jahren von Dr. Heiko Hirschmüller (HIRSCHMÜLLER 2008) entwickelt und ist heute schon in vielen kommerziellen Softwarepaketen implementiert. Der Grund für diesen Erfolg des SGM ist die effiziente Verbindung von dichter Punktkorrelation mit geringer Rechenzeit (STOLZ 2013). Dadurch ist er dem Facetten-Stereosehen (SCHLÜTER M. 1999) vorzuziehen.

Im ersten Schritt des SGM-Algorithmus werden die Matchingkosten mit „Mutual Information“ (MI) berechnet. MI gibt für jede Grauwertkombination die Stärke des statistischen Zusammenhangs an und ergibt sich aus der jeweiligen Entropie H (Informationsgehalt) der beiden Bilder und der gemeinsamen Entropie der Bilder. Ein Vorteil von MI besteht darin, dass radiometrische Transformationen möglich sind, während die lokale Abfolge der Grauwerte

erhalten bleibt (ARAND 2010). Die Matchingkosten $C(p, D_p)$ für ein Pixel p und der Disparität D_p ergeben sich aus der negativen m_i (vgl. Formel (1)) (HIRSCHMÜLLER 2008).

$$C_{MI}(p, d) = -m_{I_b, f_D(I_m)}(I_{bp}I_{mq}) \quad (1)$$

Sind für jedes Pixel die Matchingkosten berechnet, kommt es im zweiten Schritt zur Aggregation bzw. Anhäufung der Matchingkosten. Die pixelweisen Matchingkosten sowie die Glattheitsbedingung werden in eine vom Disparitätsbild abhängige Energiefunktion zusammengefasst (vgl. Formel (2)) (HIRSCHMÜLLER 2008).

$$E(D) = \sum_p \left(C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{q \in N_p} P_2 T[|D_p - D_q| > 1] \right) \quad (2)$$

Das Hauptproblem besteht darin, das Disparitätsbild zu finden, welches die Energiefunktion minimiert. Die Energiefunktion ist global, d.h. sie müsste für alle möglichen Disparitäten eines jeden Pixels ausgewertet werden. Es würde folglich sehr lange dauern, das Disparitätsbild zu berechnen. Die entscheidende Innovation von Hirschmüller bestand darin, dies semiglobal durch verschiedene eindimensionale Pfade entlang verschiedener Richtungen vorzunehmen. Im dritten Schritt wird das Disparitätsbild berechnet, indem für jedes Pixel die Disparität mit den geringsten Summenkosten ausgewählt wird. Durch die bestimmte Disparität kann für jede Position im ersten Bild die Position im zweiten Bild berechnet werden. In einem weiteren Schritt kommt es noch zur Verfeinerung des Disparitätsbildes (HIRSCHMÜLLER 2008)

3 Konzept zur flächenhaften Berechnung von Digitalen Oberflächenmodellen

Die Erarbeitung eines Konzepts zur flächenhaften Berechnung von DOM ist neben der geometrischen Bewertung der Ergebnis-Punktwolke das Hauptthema dieses Beitrags. Bei der verwendeten Software handelt es sich um „Erdas Imagine 2014“ mit dem Produkt „Imagine Photogrammetry“ und dem darin enthaltenen Tool zur Punktwolkenberechnung „SGM“. Es sollen qualitativ bestmögliche Ergebnisse unter Beachtung wirtschaftlicher Aspekte erzielt werden. Zu den wirtschaftlichen Aspekten zählt vor allem die reibungslose Berechnung der Punktwolken. Fehlermeldungen während des Workflows sollen vermieden werden. Dadurch können Zeit und Aufwand für die Prozesse minimiert werden. Für Unternehmen und Verwaltungen bedeutet dies auch gleichzeitig die Einsparung von Kosten. Als Datenbasis dienen die Luftbilder der Befliegung des Stadtgebietes Würzburg aus dem Jahr 2012, welche von der Stadt Würzburg in Auftrag gegeben wurde. Zum Zeitpunkt der Befliegung trugen die Bäume noch kein Laub, so dass dadurch entstehende Störeffekte vermieden wurden. Tabelle 1 zeigt die Parameter des Bildflugs.

Tab. 1: Parameter des Bildflugs

Bodenauflösung	5 cm
Längs- und Querüberlappung	60 %
Abmessungen Luftbild	12096 x 11200 Pixel (605 x 560 m)
Anzahl der verwendeten Luftbilder	49 (7 Bilder in jedem der 7 Flugstreifen)

Zur dreidimensionalen Generierung von Punktwolken müssen die Luftbilder orientiert werden (Bestimmung der äußeren Orientierung). Dazu wurde ebenfalls das Programm „Imagine Photogrammetry“ der Software „Erdas Imagine 2014“ verwendet (RÖDER 2014).

Zu Beginn der Untersuchungen wurden die DOM mit den Standardeinstellungen der Software berechnet. Bei Verwendung der Standardeinstellungen kam es jedoch häufig zu Fehlermeldungen und zum Abbruch der Prozesse. Zur Verbesserung der Wirtschaftlichkeit war es notwendig, die Standardeinstellungen im Laufe der Berechnungen Schritt für Schritt zu variieren und dadurch die Ergebnisse zu optimieren. Dadurch soll ein reibungsloser Ablauf der Berechnungen möglich sein.

Der in der Software enthaltene Parameter „Urban Processing“ erwies sich als der entscheidende Faktor bei den Berechnungen der DOM aus Luftbildern mit dem „SGM“-Tool. Ist der Parameter aktiv, kommt es häufig zu Prozessabstürzen. Bei aktiviertem „Urban Processing“ werden die Einstellungen für die Korrelation in Stadtgebieten optimiert. Welche Parameter im Algorithmus genau verändert werden, lässt sich aus der Online Hilfe der Software nicht ableiten (SGM 2014). Bei aktivem „Urban Processing“ ist scheinbar eine ausreichend vielfältige Textur bis zum Rand der Stereomodelle notwendig. Aus dieser Erkenntnis lässt sich Folgendes ableiten: sind in Überlappungsbereichen zweier Luftbilder Gebiete vorhanden, die keine vielfältige Textur aufweisen, so ist die Gefahr hoch, dass die Berechnung fehlschlägt. Solche Bereiche ohne vielfältige Textur können z.B. Schattenbereiche neben hohen Gebäuden sein. Im Innenstadtbereich Würzburgs kann dies v.a. in der Nähe von Kirchen auftreten (vgl. Kapitel 4). Durch „Urban Processing“ werden in Bereichen ohne ausreichend vielfältige Textur sehr leicht Ausreißer produziert. Durch diese Ausreißer kommt es im Prozess wiederum zu Fehlermeldungen und Prozessabstürzen, da in der Software festgelegte Grenzwerte überschritten werden.

Leider ist es mit der Software nicht möglich, pauschal die Überlappungsbereiche herauszufiltern, bei denen es zu einem Absturz mit aktiviertem „Urban Processing“ kommen könnte. Somit ergeben sich für den Anwender drei mögliche Vorgehensweisen:

- 1. Variante: Die Punktwolken werden ohne aktiviertes „Urban Processing“ berechnet und somit auf qualitative Verbesserungen im Stadtbereich verzichtet.
- 2. Variante: Der Benutzer aktiviert den Parameter „Urban Processing“ und geht somit das Risiko ein, dass es zu Fehlermeldungen und zum Abbruch des Batch-Prozesses kommt. Anschließend müssen die fehlerhaften Prozesse nochmals ohne aktiviertes „Urban Processing“ durchgeführt werden.
- 3. Variante: Aufteilung der Projekte in separate Bereiche (für großflächige texturarme Gebiete)

Die Qualität der Punktwolken ohne aktives „Urban Processing“ stellt sich als nicht wesentlich schlechter als die der Punktwolken mit aktiviertem „Urban Processing“ heraus. Gerade bei großen Gebieten ist es daher wirtschaftlicher, die Punktwolken ohne „Urban Processing“ zu berechnen. Mit aktivem Parameter müsste der Anwender zunächst alle Überlappungsbereiche berechnen lassen und die fehlerhaften Überlappungsbereiche notieren. Die Fehlermeldung während eines Batch-Prozesses muss vom Nutzer selbst geschlossen werden, damit der nächste Teilprozess ausgeführt werden kann. Dadurch ist es auch nicht möglich, die Prozesse z.B. über Nacht vom Anwender unbeobachtet berechnen zu lassen. Dies bedeutet wiederum eine erhebliche Steigerung des Zeitaufwands. Der Zeitgewinn bei deaktiviertem „Urban Processing“ soll mit konkreten Zahlen wiedergegeben werden: wird für die Berechnung eines Überlappungsbereiches eine Zeit von 30 Minuten angesetzt, so würde sich - bei reibungslosem Verlauf der Berechnungen ohne Fehlermeldungen (d.h. „Urban Processing“ deaktiviert) - eine Berechnungszeit von ca. 50 Stunden für den orientierten Luftbildverband ergeben (RÖDER 2015). Bei aktiviertem „Urban Processing“ zogen sich die Berechnungen jedoch über mehrere Wochen hin. Aus rein wirtschaftlicher Sicht ist also die Deaktivierung von „Urban Processing“ definitiv zu empfehlen.

4 Geometrische Bewertung der Ergebnisse

Um die erzeugten Punktwolken geometrisch bewerten zu können, d.h. ihre Qualität und Genauigkeit zu bestimmen, bedarf es Vergleichsdaten. Dazu werden LoD2-Daten der Bayerischen Vermessungsverwaltung verwendet. Die Lagegenauigkeit der LoD2-Modelle entspricht der Genauigkeit der zugrundeliegenden Gebäudegrundrisse aus der digitalen Flurkarte. Da Standarddachformen verwendet werden, entspricht die Höhengenaugigkeit der Dachpunkte ca. 1 m, bei Einzelfällen und komplexen Dachformen kann diese auch größer sein (LDBV 2015). Zunächst werden die Punktwolken ohne Einbeziehung der Vergleichsdaten hinsichtlich ihrer Geometrie bewertet. Dazu werden alle erzeugten Punktwolken gemeinsam in einer Gesamt-Punktwolke betrachtet. Am Rand der Ergebnis-Punktwolke ist die Punktdichte deutlich geringer als im Zentrum. Grund hierfür ist das Fehlen von Überlappungsbereichen unterschiedlicher Flugstreifen in den Randgebieten. Außerdem klaffen im Bereich des Flusses (Main) große Datenlücken. Hier werden wenige Punkte zugeordnet, da der SGM-Algorithmus eine ausreichend vielfältige Textur benötigt. Im Bereich von Flüssen ist die jedoch - ähnlich wie in Schattenbereichen - nicht der Fall. Im Bereich des Würzburger Ringparks existieren ebenfalls Datenlücken. Hier liegt durch die Bäume und deren Schatten ebenfalls eine nicht ausreichende Textur vor. Außerdem treten durch die Vegetation Verdeckungen auf, die es dem SGM-Algorithmus schwer machen, identische Punkte zu finden. Die Bewegung der Bäume im Wind erschwert ebenfalls die Identifikation korrespondierender Pixel in mehreren Bildern. Der Innenstadtbereich in der Mitte des Luftbildverbandes wurde mit einer hohen Punktdichte generiert.

Es soll auf einige Gebiete in der Ergebnis-Punktwolke speziell eingegangen werden. Der erste Detailausschnitt beinhaltet enge Bebauung mit großen Gebäudehöhen im Innenstadtbereich (vgl. Abb. 1).



Abb. 1: Punktwolke im Innenstadtbereich (Datenlücken: gelb)

V.a. im Schatten hoher Gebäude (Kirchen) treten hier Punktlücken auf (rote Ellipsen). Dies liegt zum einen an der nicht ausreichenden Texturierung im Schattenbereich und zum anderen an den Verdeckungen von identischen Punkten in den Luftbildern. Ähnlich verhält es sich in den engen Straßen im Innenstadtbereich. Besonders deutlich werden die Punktlücken bei der Vermaschung als 3D-Objekte mit der Software „Erdas Imagine“. Auf der südlichen, sonnenbeschienenen Seite der Gebäude ist eine sehr feine Vermaschungsstruktur zu erkennen, während in auf der schattigen Nordseite und in den senkrecht zur Flugrichtung verlaufenden Gassen eine grobe Struktur mit vielen Ausreißern zu erkennen ist. Im zweiten Gebiet wechseln sich bebaute Areale und Vegetationsflächen ab. Auch hier klaffen teilweise große Punktlücken. Die Gründe für die Datenlücken in Vegetationsgebieten wurden bereits genannt (s.o.). Wird die Punktwolke in diesem Gebiet vermascht, ist die eindeutige Trennung von Gebäude und Vegetation nicht mehr eindeutig erkennbar. Im dritten Bereich wird ein Gebiet mit weiträumiger Bebauung (Wohngebiet) und moderater Gebäudehöhe am Rande des Luftbildverbandes betrachtet (vgl. Abb. 2).



Abb. 2: Detailausschnitt der Punktwolke am Rand des Luftbildverbandes (Datenlücken: gelb)

Dort klaffen an den Gebäudeseiten, die in Richtung des Randes des Luftbildverbandes zeigen, Punktlücken (rote Ellipsen). Auf den gegenüberliegenden, in die Mitte des Luftbildverbandes gerichteten Seiten der Häuser werden Punkte bis an den Gebäudegrundriss berechnet (blaue Ellipsen). In der 3D-Vermaschung dieser Gebäude sind demnach auf der einen Seite gerade, auf der anderen Seite schräge Hauswände zu erkennen. Bei der flächenhaften Berechnung von DOM aus Luftbildern ist es demnach auch wichtig, vor der Berechnung genau zu definieren, von welchem Bereich eine lückenlose Punktwolke erstellt werden soll. Die Luftbilder, die für die Berechnung anschließend herangezogen werden, sollten das Gebiet dementsprechend großräumig abdecken. Dadurch werden solche Punktlücken, wie sie bei den angesprochenen Gebäuden vorkommen, vermieden.

Zusammenfassend für die Bewertung der Geometrie ohne Vergleichsdaten kann festgehalten werden, dass die Punktwolken insgesamt eine sehr gute Qualität aufweisen. Gerade durch die Berechnung von Überlappungsbereichen unterschiedlicher Flugstreifen wird eine hohe Punktdichte gewährleistet. Lediglich in den Randbereichen des Luftbildverbandes treten häufiger Punktlücken auf. Hier ist auch die Punktdichte geringer, da die Überlappungsbereiche unterschiedlicher Flugstreifen fehlen. Im Innenstadtbereich werden gerade auf den Dächern sehr viele Punkte zugeordnet. Auch an einigen Hausfassaden, die in zwei Luftbildern abgebildet sind, werden viele Punkte generiert (vgl. Profil in Abb. 3). Zu beachten ist außerdem, dass es gerade bei hohen Gebäuden (Kirchen), in engen Gassen, in Vegetationsgebieten und im Flussgebiet zu Punktlücken kommen kann.



Abb. 3: Ausschnitt eines Querprofils einer Punktwolke im Innenstadtbereich

Nach der Analyse der Punktwolkenqualität kann der Vergleich mit den LoD2-Daten stattfinden. Die LoD2-Daten sollen mit den Punktwolken in einer 2D-Szene überlagert werden, um herauszufinden, wie gut die Grundrisse des DOM zu denen der LoD2-Daten passen. Stichprobenartig sollen hier auch Maße abgegriffen werden, um die Differenzen zwischen LoD2-Daten und Punktwolken quantitativ darzustellen.

Es fällt auf, dass bei vielen Gebäuden der Grundriss gut zu den LoD2-Daten passt. Vor allem bei freistehenden Gebäuden ohne Dachüberhang ist dies der Fall (vgl. Abb. 4).



Abb. 4: Überlagerung der Punktwolke und der LoD2-Daten (grün)

Bei Gebäuden, die sich in unmittelbarer Nähe von hohen Bäumen oder anderer Vegetation befinden, kann in der Punktwolke nicht eindeutig nachgewiesen werden wie der Grundriss der Gebäude verläuft. Hier kommt es zu einer regelrechten „Verschmelzung“ von Gebäude und Vegetation. Auch bei sehr enger Bebauung im Innenstadtbereich treten zwischen den Gebäuden (v.a. in engen Gassen) häufig Ausreißer auf, wodurch der Grundriss der Gebäude nicht nachvollzogen werden kann. Abbildung 5 zeigt solch ein Gebiet. Am oberen und unteren Bildrand sind zwei Gebäudeblöcke und in der Mitte der Abbildung eine dazwischen liegende enge Straße zu erkennen.

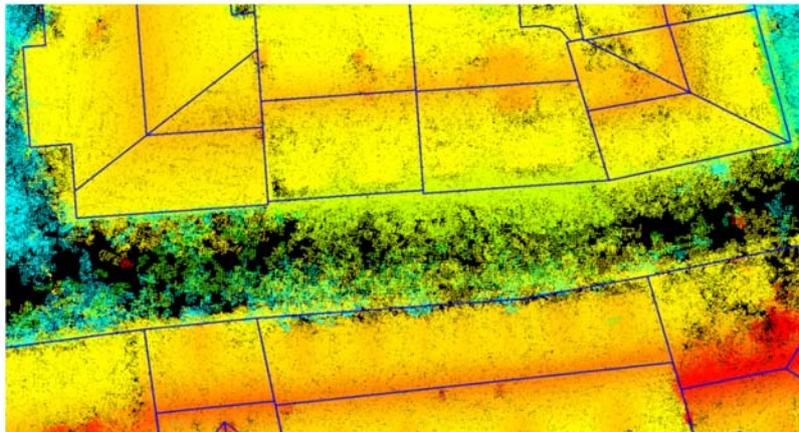


Abb. 5: Vergleich Punktwolke - LoD2-Daten (nach Höhe koloriert)

Im Bereich der Straße fällt auf, dass die Punkte ungefähr dasselbe Höhenniveau aufweisen wie die Dächer (gelbe Punkte). Eigentlich müssten hier die Punkte in der Gasse blau koloriert sein, da sie sich auf dem Höhenniveau der Straße befinden. Der Grundriss der Gebäude ist in der Punktwolke nicht genau erkennbar. Zumal wirkt es so, als seien die beiden Gebäudeblöcke miteinander „verschmolzen“.

In Abbildung 6 ist ein Gebäude in der Innenstadt (enge Bebauung) zu sehen, bei dem die Traufpunkte um ca. 80 cm und der Firstpunkt um ca. 150 cm lagemäßig gegenüber den LoD2-Daten verschoben sind. Die Differenzen wurden mit roten Pfeilen markiert.



Abb. 6: Vergleich Punktwolke – LoD2 Daten

Die durch Bildkorrelation erzeugte Punktwolke gibt in Bezug auf die Dachform wesentlich mehr Details wieder. Die LoD2-Daten stellen nur Standarddachformen dar. So wurde der Giebel links oben im Bild bei den LoD2-Daten erst gar nicht erfasst, sondern nur ein simplifiziertes Walmdach erzeugt. Ähnliche Vereinfachungen wurden bei der Erstellung der Dachformen durch die ALS-Messung bei den meisten Gebäuden im Stadtbereich durchgeführt.

Bei vielen Gebäuden passt der Grundriss gut zu den LoD2-Daten. V.a. bei freistehenden Gebäuden ohne Dachüberhang ist dies der Fall. Gebäude, die sich in unmittelbarer Nähe zu Vegetation oder anderer Bebauung befinden, zeigen das Gegenteil. In Bezug auf die Erstellung von LoD2-Daten/3D-Stadtmodellen ist die Berechnung von DOM aus Bildkorrelation durchaus mit dem ALS-Verfahren gleichzusetzen. Aufgrund der Tatsache, dass bei der Berechnung von DOM aus Bildkorrelation gleichzeitig radiometrische Informationen in den Punktwolken enthalten sind, ist diese Vorgehensweise dem ALS sogar vorzuziehen.

5 Fazit

Das gesamte Liegenschaftskataster der Bayerischen Vermessungsverwaltung soll in naher Zukunft dreidimensional geführt werden. Dazu müssen zwangsläufig flächendeckend von ganz Bayern dreidimensionale, zeitlich aktuelle Geoinformationen zur Verfügung stehen. Konventionell wurde dazu das teure ALS-Verfahren verwendet. Dieser Beitrag zeigt jedoch, dass sich für die flächenhafte Gewinnung von Punktwolken der Erdoberfläche die kostengünstigere Berechnung von DOM durch Bildkorrelation ebenfalls sehr gut eignet. Mit dem Konzept steht eine ausführliche Beschreibung der Erstellung von DOM mit der Software „Erda Imagine 2014“ zur Verfügung (RÖDER 2015). In Zukunft sollte der SGM-Algorithmus der Software in Bezug auf den Parameter „Urban Processing“ überarbeitet werden, damit ein reibungsloser Workflow nicht gestört wird. Die Einbeziehung von Überlappungsbereichen unterschiedlicher Flugstreifen in die Berechnungen ist unumgänglich. Durch den SGM-Algorithmus kann eine sehr dichte, detailreiche Punktwolke erzeugt werden. Jedoch waren auch in den Ergebnis-Punktwolken dieser Arbeit die für die Bildkorrelation typischen Probleme zu erkennen (z. B. wenige Punkte in Schatten- und Vegetationsarealen). Im Vergleich mit den LoD2-Daten der Bayerischen Vermessungsverwaltung passt gerade bei freistehenden Gebäuden ohne Dachüberhang der Grundriss aus der Punktwolke sehr gut zu den LoD2-Daten.

Zusammenfassend ist festzustellen, dass in der Berechnung von DOM durch Bildkorrelationsverfahren ein enormes Potenzial steckt. Sollen flächendeckend für ein großes Gebiet Geoinformationen erfasst werden, so ist dieses Verfahren eine kostengünstige und gleichzeitig effektive Variante der Datenbeschaffung. Zur weiteren Optimierung des DOMs sollen bei zukünftigen Befliegungen die Längs- und Querüberlappung erhöht werden. Eine Anpassung der Flugplanung an die Hauptausrichtungen des Straßensystems wäre optimal für dichte Punktwolken in den Straßen und auf den Fassaden.

6 Literaturverzeichnis

- ARAND, S., 2010: Genauigkeitsuntersuchung von digitalen Oberflächenmodellen aus Bildzuordnungsverfahren in Stadtgebieten. Bachelorarbeit, Institut für Photogrammetrie und GeoInformation, Leibniz Universität Hannover.
- HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on pattern analysis and machine intelligence*, **30** (2), 328-341.
- HÜMMER, F. & ROSCHLAUB, R., 2014: Die Zukunft ist dreidimensional - 3D Gebäudemodelle in Bayern. *Mitteilungen des DVW-Bayern*, Februar.
- KERSTEN, T., O'SULLIVAN, W. & ZANINI, M., 2001: DOM/10 und DOM/2.5 – Digitale Oberflächenmodelle durch Bildkorrelation. *VPK – Vermessung, Photogrammetrie, Kulturtechnik*, (2), 60-66.
- KRAUS, K., 2004: *Photogrammetrie Band 1 - Geometrische Informationen aus Photographien und Laserscanneraufnahmen*, de Gruyter Verlag, Berlin.
- KREY, T., MÄRKL, P., MÖST, K. & ROSCHLAUB, R., 2016: Partielle Fortführung eines Digitalen Geländemodells aus Laserscanning mittels bildbasierter Digitaler Oberflächenmodelle. *Mitteilungen des DVW-Bayern*, Januar.
- LDBV, 2015: *Kundeninformation LoD2*. Landesamt für Digitalisierung, Breitband und Vermessung Bayern, November.
- RÖDER, M., 2014: Erstellung von Digitalen Oberflächenmodellen aus Luftbildern im Innenstadtbereich. Projektarbeit, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, September.
- RÖDER, M., 2015: Konzeption und geometrische Bewertung der flächenhaften Berechnung von Digitalen Oberflächenmodellen aus Luftbildern. Bachelorthesis, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, März.
- SCHLÜTER, M., 1999: Photogrammetrische Oberflächenrekonstruktion durch Mehrbildzuordnung mit dem Facetten-Stereosehen. *Mitteilungen des Bundesamtes für Kartographie und Geodäsie* **4**, 7-21
- SGM, 2014: *Intergraph: SGM Online-Hilfe*.
- STOLZ, M., 2013: Digitale Oberflächenmodelle aus Bildkorrelation. *Mitteilungen des DVW-Bayern*, Februar.
- TILSNER, A., 2012: 3D-Messdaten. Informationsveranstaltung „3D-Geobasisinformation – die Produkte des LGLN“. Landesamt für Geoinformation und Landentwicklung Niedersachsen, September 2012.

La Photogrammétrie Peut-elle Remplacer La Topométrie Traditionnelle Pour du Levé de Bâtiments ?

GREGORY STEINER¹ & BERTRAND CANNELLE²

Résumé : Depuis une quinzaine d'années, l'émergence des techniques d'acquisition et de modélisation 3D pousse les professionnels de la construction à se tourner vers les géomètres, professionnels de la mesure, pour leurs besoins de modélisation ou mesures. Jusqu'alors et très bien implantée, la topométrie traditionnelle, tachéométrie et lasergrammétrie, a fait ses preuves pour du levé de bâtiments. Toutefois ces techniques peuvent s'avérer longues et coûteuses et l'investissement de départ peut en décourager plus d'un.

L'arrivée récente sur le marché public des drones, les capacités informatiques grandissantes, l'avancée des capteurs numériques et l'implémentation de procédés issus de la communauté de la vision par ordinateur créent un cocktail tridimensionnellement explosif ! La photogrammétrie numérique moderne est née, se démocratise à grands pas et devient accessible à un grand nombre de professionnels. Ces nouvelles technologies sont certes "sexy" mais peuvent-elles rivaliser avec les méthodes de topométrie traditionnelles ? Et à quel prix ? Le marché reste marginal et ces récentes technologies sont sujettes à de nombreuses interrogations en terme de résultats (précision, référencement, produits).

Les travaux et recherches effectués dans cet article démontrent qu'avec des méthodes simples et un équipement lowcost il est possible, à certaines conditions, de traiter un bloc photogrammétrique dont sa robustesse permet une modélisation et des mesures subcentimétriques par géoréférencement ou simple mise à l'échelle. Autre avantage certain de cette technologie : les images permettent des mesures stéréoscopiques sûres et indépendantes d'un nuage de points ou d'une modélisation. Ces résultats permettent de générer différents produits selon les besoins du client qui lui serviront de base d'analyse et de mesure. Nous présenterons, sur un même bâtiment, différents protocoles d'acquisition (acquisition en géométrie panoramique, utilisation de position GPS sur les images, mise à l'échelle, etc) ainsi que leurs précisions respectives.

1 Introduction

Depuis une quinzaine d'années, l'émergence des techniques d'acquisition et de modélisation 3D pousse les professionnels de la construction à se tourner vers les géomètres, professionnels de la mesure, pour leurs besoins de modélisation ou mesures. Jusqu'alors et très bien implantée, la topométrie traditionnelle, tachéométrie et lasergrammétrie, a fait ses preuves pour du levé de bâtiments. Toutefois ces techniques peuvent s'avérer longues et coûteuses et l'investissement de départ peut en décourager plus d'un.

Dans cet article, nous proposons d'étudier différentes configurations de chantiers photogrammétriques aussi bien au niveau de la matérialisation qu'au niveau des capteurs avec des méthodes simples et peu coûteuses qui pourraient être utilisées en complément ou à la place des méthodes habituelles.

Dans une première partie, nous présenterons le matériel utilisé, puis dans une seconde nous présenterons différentes configurations pouvant être utilisées pour atteindre des précisions centimétriques.

¹ Jean Hodler SA, Rte du Château d'Affry 6, CH-1762 Givisiez, www.hodlersa.ch

² HEIG-VD, Route de Cheseaux 1, CP 521, CH-1401 Yverdon-les Bains,
E-Mail bertrand.cannelle@heig-vd.ch

2 Matériels utilisés

2.1 Appareil photographique

L'appareil photo numérique utilisé est un Canon EOS 5D Mark III. Il possède une résolution de 22 MégaPixels, une très bonne dynamique d'image (14bits) ainsi qu'un temps de pose allant de 30 secondes à 1/8000 de seconde (CANON 2016). Il permet également de prendre des rafales de 6 images par seconde pour de l'acquisition rapide (ALLAZ & KASSER 2006). Dans notre cas, il est parfaitement utilisable pour prendre des photos dans un environnement sombre ou éclairé ainsi que pour des scènes avec des objets en mouvement.



Fig 1 : EOS 5D Mark III (WikipediaEOS 2016)

Tab. 1 : Taille du pixel en fonction de la focale et de la distance

	Objectifs (mm)					
	600	400	100	50	28	14
distance (m)	taille d'un pixel (mm)					
10	0,1	0,2	0,6	1,3	2,2	4,5
20	0,2	0,3	1,3	2,5	4,5	8,9
50	0,5	0,8	3,1	6,3	11,2	22,3
100	1,0	1,6	6,3	12,5	22,3	44,6



Fig. 2 : Tête panoramique avec l'appareil photo

Un autre intérêt de l'appareil photographique reflex est de pouvoir adapter l'objectif en fonction de l'objet à photographier et de la résolution souhaitée. Le tableau 1 ci-dessous donne quelques tailles de pixels en fonction de la distance à l'objet et de la focale choisie.

2.2 Plateforme d'acquisition : une tête panoramique motorisée

Le RODEON piXpert, commercialisé par la société Clauss (CLAUSS 2016) (figure 2), est une tête panoramique motorisée entièrement automatisée. Son plateau inclinable (pan-tilt-head en anglais) permet une rotation verticale et horizontale de 360°. La puissance de ses moteurs et son frein magnétique permettent de fixer, en toute sécurité, toute sorte de reflex numériques (même dotés de lourds téléobjectifs (CANNELLE 2015)), de caméras moyen format et de les positionner avec grande précision.

2.3 Matérialisation

Dans la photogrammétrie traditionnelle, il est d'usage d'utiliser des cibles connues en 3 dimensions dans la scène ou sur l'objet. Ces coordonnées sont souvent mesurées par des stations totales ou par GNSS. Nous avons choisi ici de tester des alternatives : mire de nivellement pour la mise à l'échelle ainsi que GNSS pour mesurer les sommets de prise de vue.

2.3.1 Mire de Nivellement

Le géomètre, dans son matériel traditionnel, possède des mires, utilisées habituellement pour faire du nivellement. Une mire présente l'avantage d'avoir une distance connue avec une grande précision (submillimétrique) ainsi que des graduations facilement identifiables dans les images. Une ou plusieurs mires peuvent être disposées dans la photographie afin de mettre le bloc photogramétrique à l'échelle.



Fig. 3 : Mire de nivellement pouvant être utilisée pour la mise à l'échelle

2.3.2 GNSS

Le GNSS fait également partie du matériel traditionnel du géomètre. Il est habituellement utilisé pour du levé d'objet, pour de l'implantation ou encore pour de la mesure de réseaux géodésiques. Dans notre cas, nous allons l'utiliser pour mesurer la position des images ou de la tête panoramique présentée au 1.2. Il suffit alors de substituer la tête panoramique à l'antenne GNSS. La partie 3.2 décrit le processus complet.

2.4 Logiciel de compensation

Pour toute cette étude, nous avons choisi d'utiliser le logiciel Agisoft PhotoScan (AGISOFT 2016). Il permet de mesurer les points homologues automatiquement et de calculer les positions des images sans solution initiale. Il permet également, entre autre, de : mettre un modèle à l'échelle sans point d'appui, de fixer les sommets des images, et surtout de contraindre la position d'un groupe d'images : condition remplie lors de l'acquisition en géométrie panoramique. Cette dernière condition n'est, pour l'instant, pas remplie par les autres logiciels du marché comme Micmac (PIERROT-DESEILLIGNY & CLERY 2011) ou encore Pix4D (PIX4D 2016).

La saisie des points dans les images a été réalisée par mesure multi-images. En effet, ce type de chantier n'est pas propice à la mesure stéréoscopique car les images ne sont pas toujours prises dans les conditions standards : images parallèles entre-elles et plan image parallèle au « plan » de l'objet (MCGLONE et al. 2004). Un autre avantage à ce type de mesure et qu'elle ne nécessite pas de savoir-faire particulier au niveau de l'opérateur. Enfin la saisie peut être faite dans un grand nombre d'image, ce qui permet d'augmenter la précision du pointé : le point 3D correspond à l'intersection de tous les faisceaux issus de toutes les images.

Après avoir présenté le matériel utilisé nous allons maintenant nous intéresser aux différentes techniques permettant un relevé

3 Protocoles d'acquisitions pour du levé de bâtiments

Dans cette partie, nous allons présenter plusieurs façons de géoréférencer des images sans utiliser de cible ou de point naturel et dans des configurations historiquement peu recommandées en photogrammétrie.

3.1 Imagerie panoramique et levé de grande hauteur

La première configuration est l'utilisation de panoramiques (cônes verts sur la figure 4) avec une simple mise à l'échelle pour du levé de façade en utilisant une mire de nivellement présente sur les photos. Le principe est de faire 3 stations panoramiques afin d'avoir de la tri-stéréo sur tout l'ouvrage. Deux focales différentes sont choisies (cône vert clair et vert foncé) afin d'avoir la même résolution sur l'ouvrage : une focale deux fois plus longue pour une panoramique deux fois plus loin (voir figure 4). Les positions des stations panoramiques ont été choisies afin d'éviter les configurations « dégénérées » : trois stations alignées par exemple.

Ensuite, nous avons choisi d'utiliser une mire de nivellement (cf partie 1.3.1), en rouge sur la figure 4 afin d'effectuer la mise à l'échelle en la disposant sur la partie basse du chantier (trait rouge 1-2 sur la figure 4).

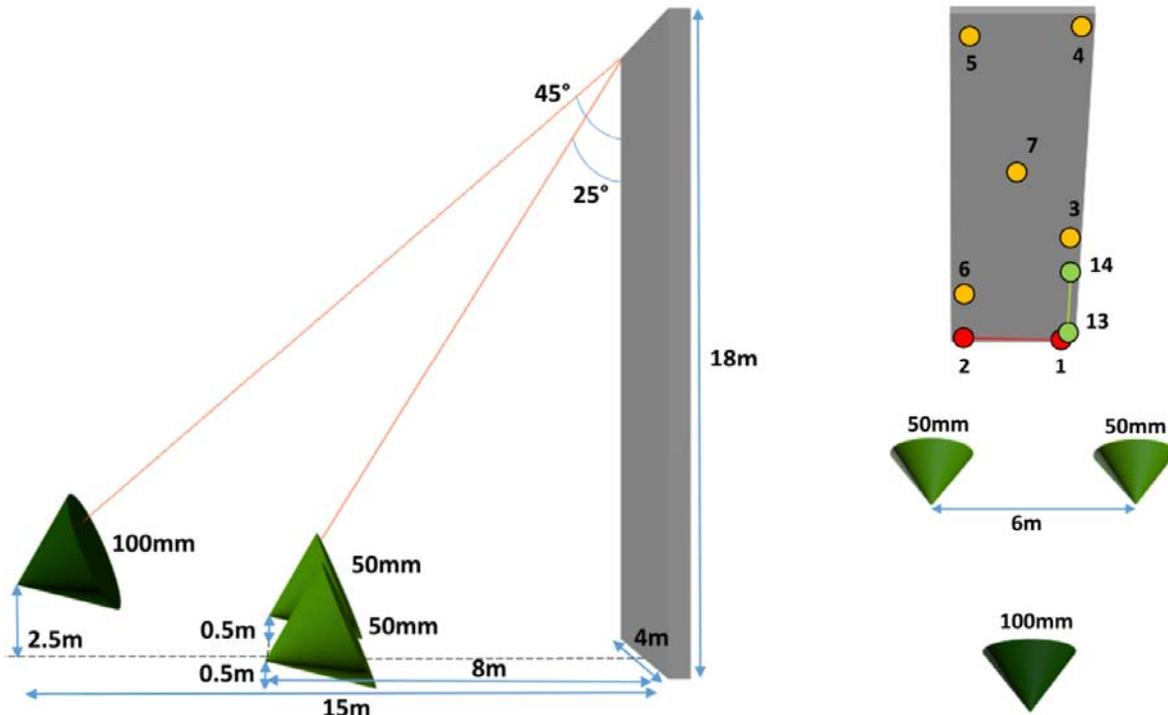


Fig. 4 : Protocole de mesure de façade par station panoramique et mire de nivellement

Le tableau 2 et la figure 5 présentent les résultats de l'expérimentation. La première constatation est que le recouvrement des images au sein d'une même panoramique n'est pas significatif : on peut donc choisir un recouvrement minimal (une dizaine de pourcents) afin de garantir des liens entre les images et de limiter le temps d'acquisition et de calculs.

Ensuite, en faisant une mise à l'échelle au bas du bloc, nous arrivons à des erreurs sur les mesures de distances sur tout le reste de l'ouvrage en moyenne à 5mm près, même sur les distances au sommet de la façade. Ce protocole est donc tout à fait adapté pour du relevé centimétrique géoréférencé : ce qui convient à certains travaux d'architecture et de relevés patrimoniaux.

Tab. 2 : Résultats Protocole de mesure de façade par station panoramique et mire de nivellement

		Recouvrement 70% H, 265 img, 977'007 TP e.o. 8.46, m.t. 2 h				Recouvrement 40% H, 41 img, 419'465 TP e.o. 3.14, m.t. 30 min			Recouvrement 10% H, 17 img, 120'609 TP e.o. 2.53, m.t. 2 min		
Base	d_i [m]	l_i [m]	$\Delta l_i - d_i$ [mm]	ppm	l_i [m]	$\Delta l_i - d_i$ [mm]	ppm	l_i [m]	$\Delta l_i - d_i$ [mm]	ppm	
1-2	3.970	3.970	0		3.970	0		3.970	0		
13-14	1.800	1.800	0		1.800	0		1.800	0		
3-4	14.728	14.733	5	339	14.734	6	407	14.723	-5	-340	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	
6-7	6.245	6.247	2	320	6.248	3	480	6.246	1	160	
$\overline{ x }$			4	423		5	493		3	284	

La figure 5 présente un résultat possible pouvant être transmis au commanditaire : une orthophoto métrique, sur laquelle les distances peuvent être mesurées. Le principal avantage de ce produit est de laisser le commanditaire choisir les objets à mesurer.

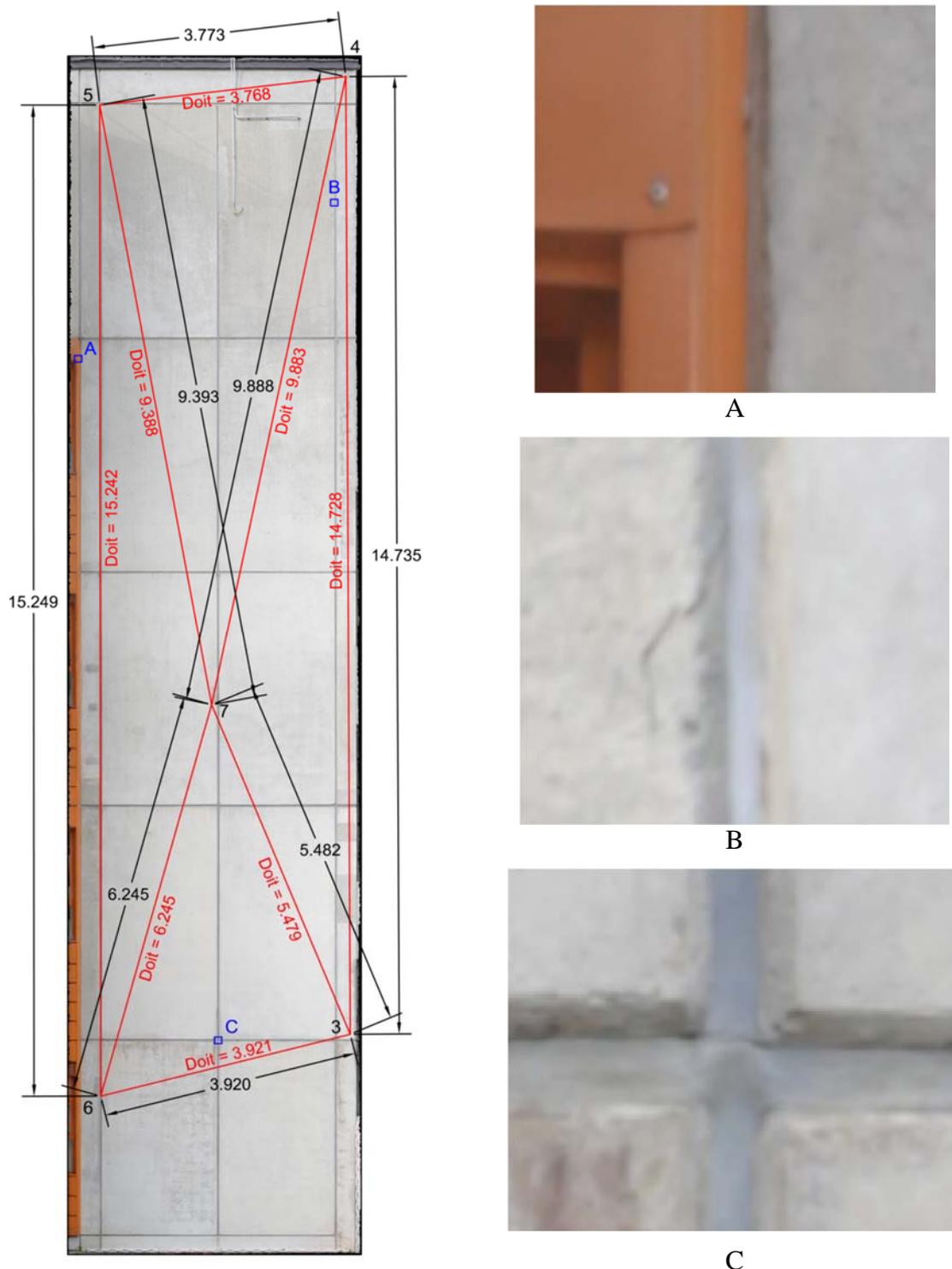


Fig. 5 : mesures DAO sur orthomosaïque 2mm de la façade et échantillons de détails.

3.2 Imagerie panoramique pour du levé de bâtiment et positionnement GNSS

Le principe de cette méthode est de mesurer la position des panoramiques par GNSS. On substitue la tête panoramique par une antenne GNSS comme sur la figure 6. Le principal avantage est de ne pas à avoir à disposer de cibles 3D ou de détails à mesurer, ni de mise à l'échelle. En effet, le géoréférencement est effectué en fixant la position 3D de plusieurs panoramiques. Cela fixe également le facteur d'échelle du bloc. Ensuite, le reste du chantier est

traité par photogrammétrie « standard » : calcul de points homologues et assemblage des images en utilisant la contrainte de position sur les sommets mesurés par GNSS.



Fig. 6 : Antenne GNSS substituée à la tête panoramique

Nous avons testé deux configurations différentes : 7 panoramiques fixées ou 4 panoramiques fixées (K1 et K2 sur la figure 7) afin d'évaluer l'influence du nombre de station GNSS.

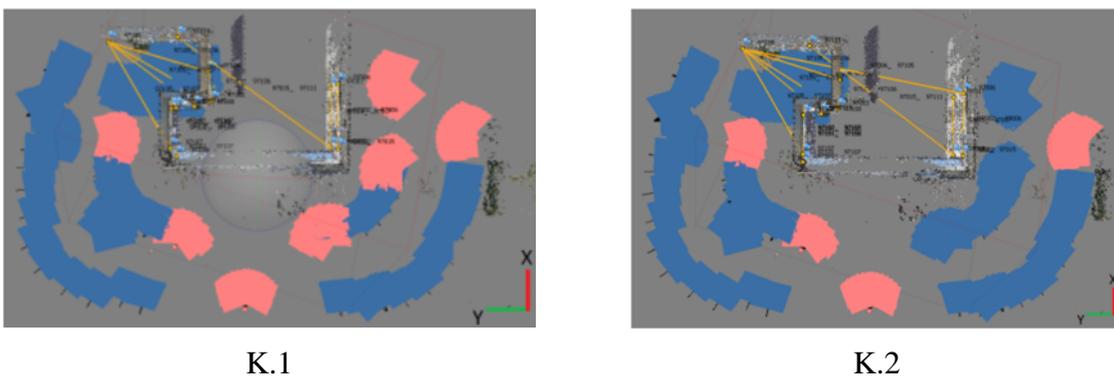


Fig. 7 : Configurations testées : en rose, les panoramiques fixées, en bleu, le reste des images

Le tableau 3 présente les résultats des 2 acquisitions. Sur les deux chantiers présentés, la précision 3D des points est de l'ordre du centimètre : précision du GNSS RTK. On arrive donc à géoréférencer nos images avec la même précision que notre outil de géoréférencement. La configuration à 7 stations GNSS est légèrement meilleure que celle à 4 (de 3mm) : la configuration minimale peut donc être utilisée.

Tab. 3 : Précision sur les coordonnées 3D mesurées par photogrammétrie.

Point	Test K.1. 7 stations RODEON-VR TK M, 586 img, 993'198 TP e.o. 3.9, m.t. 7 h				Test K.2. 4 stations RODEON-VR TK M, 586 img, 993'198 TP e.o. 3.9, m.t. 7 h			
	$\Delta 3D$ [mm]	ΔX [mm]	ΔY [mm]	ΔZ [mm]	$\Delta 3D$ [mm]	ΔX [mm]	ΔY [mm]	ΔZ [mm]
Station 1	16.4	4.3	13.1	8.8	21.2	7.3	18.8	6.5
Station 2	18.3	10.0	13.0	-8.1	27.7	12.7	21.4	-12.1
Station 3	14.1	-9.2	-2.9	-10.3	13.8	-7.7	0.4	-11.4
Station 4	25.4	-12.8	17.5	-13.3	29.4	-14.7	20.6	-15.0
Station 5	12.7	-8.6	-8.8	3.0	3.8	-3.4	-1.7	0.1
Station 6	11.6	8.4	-7.4	3.1	10.9	8.4	-6.2	2.9
Station 7	12.9	3.4	-4.8	11.5	9.0	-0.7	-0.8	8.9
7002	24.8	8.7	16.4	-16.3	31.8	11.0	22.4	-19.7
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
7115	38.5	14.3	26.5	-23.9	43.4	11.4	31.8	-27.4
Moyenne [mm]	32.8				35.3			

Les résultats précédents montrent que nous obtenons une précision centimétrique sur les points, autrement dit de géoréférencement global. Le tableau 4 présente les résultats de cohérence « interne » du processus. Cela permet de renseigner sur la précision des distances entre objet que l'on peut obtenir malgré un géopositionnement à 3 cm. Les résultats montrent que sur 11 distances mesurées (en jaune sur la figure 7), la cohérence des distances est meilleure que le centimètre. Cela signifie que l'on peut mesurer des distances de l'ordre du centimètre malgré des références précises à 3 cm ! En effet, la mise à l'échelle est faite en utilisant toutes les mesures mais il peut y avoir un biais résiduel.

Tab 4 : Précision des distances mesurées par photogrammétrie.

Base	Test K.1. 7 stations RODEON-VR TK M, 586 img, 993'198 TP e.o. 3.9, m.t. 7 h				Test K.2. 4 stations RODEON-VR TK M, 586 img, 993'198 TP e.o. 3.9, m.t. 7 h		
	d_i [m]	l_i [m]	$\Delta l_i - d_i$ [mm]	<i>ppm</i>	l_i [m]	$\Delta l_i - d_i$ [mm]	<i>ppm</i>
7002-7006	5.616	5.617	1	178	5.617	1	178
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
7107-7108	3.605	3.609	4	1108	3.609	4	1108
$\overline{ x }$			9	1027		8	948

3.3 Comparaison avec la lasergrammétrie

Lors de la mise en place de ce chantier, le bâtiment a fait l'objet d'une modélisation lasergrammétrique de référence. Le but de ce test est d'observer les différences entre ces deux différentes méthodes de modélisation. La comparaison du modèle entre le Canon 5D et la lasergrammétrie traditionnelle montrent que 68% des écarts se situent à ± 5 mm et 95% à ± 10 mm. Les 5% restant qui s'écartent du centimètre sont les arrêtes et les encastremets qui restent difficiles à modéliser (avec les 2 techniques). La figure 8 ci-dessous représente deux

Un autre avantage de la technique est de pouvoir compléter les mesures 3D si nécessaire en utilisant les images acquises (sous réserve que l'objet ait été photographié dans au moins 2 images) sans devoir revenir sur le terrain.

Ces techniques sont également peu coûteuses en travail de terrain et en équipement sur place et peuvent être réalisées même si le bâtiment est difficile d'accès ou dangereux car les images se font à distance et sans ajout de cibles directement sur l'objet.

6 Références

AGISOFT, 2016: <http://www.agisoft.com/>, accessible le 28/04/2016.

ALLAZ, G. & KASSER, M., 2016: Auscultation de l'état de surface de revêtements routiers par photogrammétrie automatisée. *Revue XYZ*, numero 146, 1er trimestre 2016, 52-58.

CANNELLE, B., 2015: Les outils du paparazzi au service de la surveillance d'ouvrage ?. *Geomatique Suisse* (9), 376-379.

CANON, 2016: http://fr.canon.ch/for_home/product_finder/cameras/digital_slr/eos_5d_mark_iii/, accessible le 28/04/2016.

CLAUSS, 2016: <http://dr-clauss.de/en/foto-studiotechnik-3/rodeon-pix-series-eng-3>, accessible le 28/04/2016.

MCGLONE, J., MIKHAIL, E., BETHEL, J. & MULLEN, R., 2004: *Manual of photogrammetry* (5th Edition). American Society for Photogrammetry and Remote Sensing, 1151 p.

PIERROT-DESEILLIGNY, M. & CLERY, I., 2011: APERO, an Open Source Bundle Adjustment Software for Automatic Calibration and Orientation of a Set of Images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **38** (5/W16), 269-276.

PIX4D, 2016: <https://pix4d.com/>, accessible le 28/04/2016.

WIKIPEDIAEOS, 2016: https://commons.wikimedia.org/wiki/File:Canon_EOS_5D_Mark_II_with_50mm_1.4_edit1.jpg, accessible le 28/04/2016.

Recognition of Road Network in Aerial Images using CRF and Marked Point Process

XUANZI YONG¹ & UWE SÖRGEL²

Abstract: In this paper, we propose a method to combine a Marked Point Process with Conditional Random Fields for the purpose of road network extraction, in which stochastic process and probabilistic graphical model classification are coupled together, aiming at extracting road network faster and more efficiently. In our method, superpixel segmentation is conducted first and the resulting segments are regarded as the basic units or primitives for classification. The features are mainly derived from the attributes of an ellipse circumscribing the superpixel, which are fed into a conditional random field operating on the superpixel graph. On the next step, each superpixel segment is modeled as a directed point in space. Instead of usual uniform sampling as prior, we take advantage of the labeling results of the previous CRF classification, which provides for each segment a belief of being as a part of road. The spatial interactions between point segments are defined by means of the direction of the superpixel ellipse. Compared with the traditional marked point process, much less input points are considered, the sampling procedure is reduced leading to faster convergence of the optimizing process.

1 Introduction

The knowledge of road network offers plenty useful information for human daily life such as routing and navigation, as well as applications in commercial and industry area, for example, hot topics like unmanned aerial vehicles and self-driving cars, in use for purposes such as surveillance and control. To be useful, this kind of information need to be updated correctly on time. In order to minimize or even avoid such manual operations, the problem of automatic road network recognition has been widely investigated for many years.

However, this topic is still a challenging. There are many difficult factors preventing to get a good solution. The appearances of roads are various. We might deal with a countryside alley, which is narrow and curved, or a crowded road on metropolis, which is full of moving mobiles and parked with parking cars on the side. Moreover, the complex environment around road area may cause mistakes, for example, buildings with concrete roofs or the traditional grey-green roof tiles. These road-like background objects should be eliminated. On the other hand, a core purpose of road networks is to provide good connectivity, which is one of the kern point especially for routing and navigation. But in practice, the shadows of buildings and even the crowns of trees may divide the road into parts.

In this work, we propose a new point process framework for road network extraction. It contains three steps as image segmentation, CRF pre-classification and point processing. The segments are

¹ Technische Universität Darmstadt, Fachgebiet Fernerkundung und Bildanalyse, Franziska-Braun-Str.7, D-64287 Darmstadt, E-Mail: yong@geod.tu-darmstadt.de

² Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, E-Mail: soergel@ifp.uni-stuttgart.de

modelled as shaped points which inherit the feature of the underlying segment. Also, a CRFs as well as MPP model is constructed for road recognition. While common MPP model the objects as random numbers at some random location, the sampling process is quite time-consuming. Our proposed method uses a simple CRF to derive an initial segmentation and classification, leading to a significant lower number of objects compared to the given pixel grid. Then we turn to MPP sampling, which ensures that relevant points are sampled more frequently, leading to significant speed up the simulation process and convergence to final result.

2 Related works

The problem of automated road recognition has been widely investigated for many years. The proposed resolutions range from image segmentation (SONG & CIVCO 2004) to line detection (JIA et al. 2005), from mathematical morphology (VALERO et al. 2010) to advanced statistic models (WEGNER et al. 2013).

Conditional random fields (CRF) are widely used for image classification and labeling (ZHANG & JIA 2012; ZHONG et al. 2014). HUANG et al. (2014) proposed an object-based CRFs model for road extraction, in which each object is represented by the color, the gradient of histogram and the texture. Object-based methods are robust to noise and can make full use of spatial information. A road network can also be represented in a graph as collection of line segments. WEGNER et al. (2015) introduced a higher-order CRF model, where the clique consists of all pixels along minimum cost paths which enables to overcome to some extent occasional false negative labeling in between longer chains of true positive segments.

Marked Point Process (MPP) model objects within a stochastic framework and are adapted by incorporation of geometrical constraints in the solution in a manner which enables dealing with correlated and geometrical noise (DESCOMBES & ZERUBIA 2002). Models based on marked point process have provided convincing results in various applications like population counting, structure detection and texture representation (LAFARGE et al. 2010). The recognition targets are created as parametric objects, e.g., in (STOICA et al. 2004), roads are treated as sets of line segments with length and orientation. The small line segments are random sampled and moved to construct a road network by concerning the interaction between line segments. In this way, the figure of road can be remain well. CHAI et al. (2013) proposed a new point process taking junction point into consideration, which provides a structurally-coherent solution and performs functional well on different network extraction applications.

Besides aerial imagery, many researches are concerning with other kinds of remote sensing data for road network extraction, e.g. satellite imagery (SINGH & GARG 2013; YUAN et al. 2011), LiDAR (ZHAO & YOU 2012), SAR images (LU et al. 2014; HE et al. 2013; LIU et al. 2013), 3D Point clouds (BOYKO & FUNKHOUSER 2011) and others (WANG et al. 2012).



Fig. 1: Superpixel segmentations with different capacity. First two images are segmented by SEEDS and the last two by SLIC.

3 Problem formulation

3.1 Image segmentation and points description

Since the remote sensing data are always in a large scale, we first generate superpixel segments and treat them as the entities for subsequent processing in this work. In this way, the search space is simplified and the computational load therefore reduced.

Usually a road network is regarded as a set of several long and narrow connected parts. In order to extract these properties from our data, we take use of the circumscribed ellipse to describe such irregular shaped segments. And the following features are taken into account, which characterize an ellipse:

- center coordinate, which refers to the location of the segments,
- rotation angel, which points out potentially, if the segment is a part of road, the orientation of the road,
- axes, which represent the scale of the segment.

Besides, we define a feature *ratio* as the quotient of minor axis to major axis of a circumscribed ellipse, $ratio \in (0,1]$, which describe the shape of the segment. A low ratio value indicates an elongated shape.

These superpixel segments are hereinafter referred to as points. And these points could be treated as “line segments”. Thus the image is represented as a set of marked points $X = \{x_i\}$, $x_i = (w_i, g_i, o_i, r_i)$, where w_i, g_i, o_i, r_i are location, gray value intensity, orientation and ratio of the point respectively. These parametric objects will be used later in marked point process.

3.2 Pre-classification using Conditional Random Fields

Conditional random field is a probabilistic model for labelling. We construct a simple pairwise CRF to pre-classify the image. The graph energy function is defined on unitary as well as pairwise cliques as:

$$E = \sum_{i \in \mathcal{V}} \Psi_i(x_i) + \sum_{(i,j) \in \mathcal{E}} \Psi_{ij}(x_i, x_j), \quad (1)$$

where \mathcal{V} refers to the set of points and \mathcal{E} represents the set of adjacent point pairs.

The unary energy Ψ_i is derived by training a Random Forest Classifier (BREIMAN 2001). Considering the local attributes, the grayscale intensity as well as the ratio of the points are chosen

as feature of a point. The orientations of the points are still ignored here, because it is not useful for unary potential which does not consider context yet.

The binary energy $\Psi_{i,j}$ is formed by a contrast sensitive Potts model as:

$$\Psi_{i,j}(x_i, x_j) = \begin{cases} 0, & \text{if } x_i = x_j, \\ \theta_1 + (1 - \theta_1)e^{\theta_2 \|I_i - I_j\|^2}, & \text{otherwise,} \end{cases} \quad (2)$$

where θ_1 and θ_2 are hyperparameters learned during training. I_i and I_j are the feature vectors of point i and j , respectively. The extracted grey values and the orientation properties are applied as features. The model punishes the labelling inconsistency between adjacent points with similar optical performance and a high likelihood to be connected road parts.

Finally, alpha-expansion graph cut algorithm (BOYKOV et al. 2001) is applied to infer the labelling. Many researchers have been working on the topic of CRFs for road extraction and shown a better result (WEGNER et al. 2013). Since in our framework, CRFs are only used as pre-classification, we take only simply model and no more details are discussed here.

3.3 Marked Point Process

A Point process F describes a random configuration of points in a continuous bounded set K . For $n \in \mathbb{N}$, let Ω_n be the set of all unordered configurations $\omega = \{\omega_1, \dots, \omega_n\}$ that consists n unordered points $\omega_i \in K$. A point process is a measurable map from a probability space into the set of configurations Ω . Marked Point Process is extended from a point process with marks. Adding each point p_i with additional parameters m_i associates the point with an object $x_i = (p_i, m_i)$, thus the geometric features of images can be modelled.

Comparing to (LACOSTE et al. 2005), the problem of geometric constrains in this work are simplified. The spatial interaction behavior can be ignored, while the relationship between two different points are only adjacent or not.

3.3.1 Data energy

A probability density $h(\cdot)$ can define the consistency of data and spatial independence between points. The total energy of the model is defined as:

$$U(x) = \sum_{x_i \in X} D(x_i) + \sum_{x_i, x_j \in X} V(x_i, x_j), \quad (3)$$

where $D(x_i)$ is the unary data energy measuring the coherence of the object x_i with respect to image and $V(x_i, x_j)$ measures the quality of a pairwise interaction between objects.

As shown in Fig. 1, the pixels on the road usually have high responses ratios and similar colors. Considering these two assumptions, the data consistency intensity is defined as: $h_d(x) \propto h_g(x) \cdot h_r(x)$, where h_g and h_r are the grayscale and ratio value distributions learned from samples by means of normalized histograms. And the data consistency energy is represented as: $D(x_i) = -\log h_d(x_i)$.

With respect to the graph connectivity, three different criteria are considered: the orientation similarity, the degree of adjacency and a self-defined distance between two points. The pairwise energy is defined as:

$$V(x_i, x_j) = E_{conn}(x_i, x_j) \cdot E_{adj}(x_i, x_j) \cdot E_{dist}(x_i, x_j). \quad (4)$$

Unlike the complex interaction model in (STOICA et al. 2004), we are interested only in the local neighborhood. As introduced in Sec. 3.2, we penalty a pair of adjacent points with great difference in orientation with

$$E_{conn}(x_i, x_j) = 1 - \sin(\|o_i - o_j\|), \quad (5)$$

where o_i is the orientation of x_i .

In some cases rooftops of buildings look quite similar to adjacent road segments. They feature similar or even parallel orientation like the road segments. In order to filter such objects, we enforce a collinearity constraint along roads by calculating the distance from a point to the other adjacent point located line. For a pair of points, the distance from a point to another point standing line is not symmetric. Hence we take the energy as the average of the bidirectional distance:

$$E_{dist}(x_i, x_j) = \frac{1}{2} \left(dist(x_i, x_j) + dist(x_j, x_i) \right), \quad (6)$$

where $dist(i, j)$ is the function calculating the distance from i to the line decided by j and its orientation.

E_{adj} is a term referring to the sum of lengths of the edges that two points have in common. As can be seen from Figure 1, the dark gray segments are connected on the shorter side and thereby construct a road network. Therefore, we decide the adjacent relevant energy as:

$$E_{adj}(x_i, x_j) = \min \left\{ \frac{c}{N_{i,j}}, 1 \right\}, \quad (7)$$

where $N_{i,j}$ refers to the number of adjacent pixels in x_i and x_j , and c is a factor to ensure the value of E_{adj} in a proper range.

Thus two points in the near with similar orientation can be characterized together.

3.3.2 Simulation

The maximum density is usually estimated by a RJMCMC sampler (GREEN 1995) coupled with a simulated annealing in CHAI et al. (2013) and STOICA et al. (2004). In this way, the global optimization can search over varying dimensions. This sampler simulates a discrete Markov Chain $(X_t)_{t \in \mathbb{N}}$ on the configuration space Ω , converging towards an invariant measure specified by U . The conventional RJMCMC sampler performs successively on objects, but takes long time, especially for large scale problems.

Conventionally, in point process, a Poisson or a uniform distribution are used as reference distribution. For Poisson process, the number of points follows a discrete Poisson distribution and the position of them is uniformly and independently distributed. In image analysis applications, uniform distribution is used, while all the points are treated equally in an image. VERDIE & LAFARGE (2012) proposed a new parallel sampler integrated with data-driven partition, which allow the points non-uniformly distributed in the image. In our model, the reference distribution is determined by the results from previous CRF classification. After CRFs pre-classification, the image parts are labelled as roads or backgrounds. With these characteristics, the points of images are distinguished by sampler and treated as non-uniform distributed. The points classified as roads are signed with high possibility. And the interested points with high possibility should be sampled frequently and has a high acceptancy.

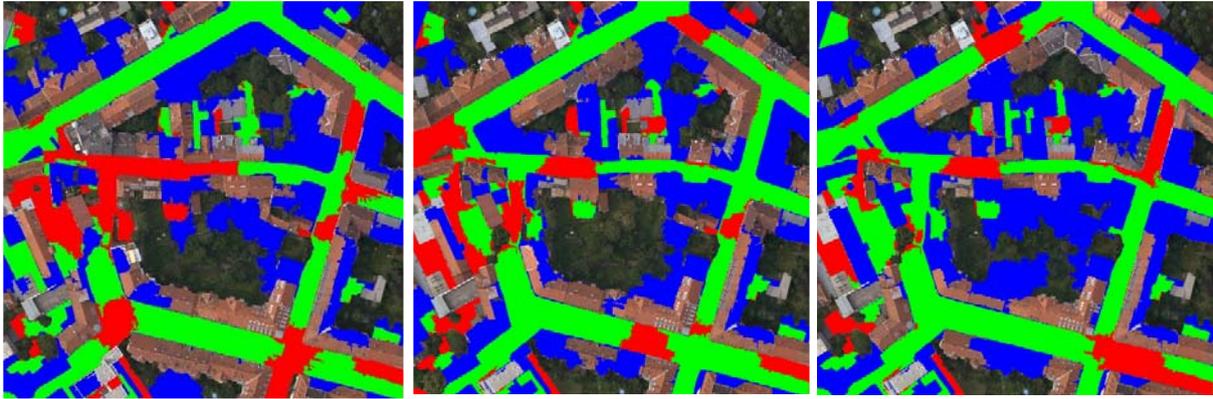


Fig. 2: Proposed road extraction results. From left to right [1]: using simple CRFs; [2]: using MPP with CRFs; [3]: using Poisson point process. True positives are displayed green, false positives blue and false negatives red

4 Experiments

We evaluate our experiments on the dataset of the city Graz, Austria. Each image from the dataset is a tile of the aerial true mosaics with 1000×1000 pixels. The simulation process are dealt with the open source C++ library `librjcmc` (BRÉDIF & TOURNAIRE 2012).

The images are first segmented into superpixels. We apply two algorithms SEEDs (VAN DEN BERGH et al. 2012) and SLIC (ACHANTA et al. 2012) respectively. As shown in Fig. 1, SLIC offers a more regular and compacter superpixels with a low computation. The superpixel generated by SEEDs are usually so irregular, that the major axes of their circumscribed ellipses are long. Hence, the points with low ratios are dominant, which has a bad effect on the quality of results.

For the purpose of studying the effect of the number of points in the model, we segment the image into about 200 superpixels and 600. From Fig.1, more details are separated out such as the parked cars on roads. Since the scale of each superpixel is smaller than before, most of the points are closer to the similar. On this aspect, the shape feature like ratios and orientation cannot describe the attributes of the points reasonable and hence it is not proper to our model. On the other hand, we focus on a fast resolution of the road network extraction problem. Considering the time and space computation, the segmentation leading to a lower number of segments is taken in this work. The quality of results are measured by the following generally used standard criteria for road network extraction. Completeness, defined as: $\text{Completeness} = \text{TP}/(\text{TP}+\text{FN})$, is the fraction of the road parts that are successfully extracted. And correctness estimates the percentage of the road parts extracted in results, as $\text{correctness} = \text{TP}/(\text{TP}+\text{FP})$. Quality is a balance accuracy in the form: $\text{quality} = \text{TP}/(\text{TP}+\text{FN}+\text{FP})$. Noticing that we only deal with the evaluation process pixel-wise.

First of all, comparing to the conventional methods, this proposal accelerate the simulation process in a significant degree. Due to the low data capacity, that 1000×1000 pixels are represented only as about 200 points, the results come out in about 75s on average, and for the same situation, other methods need at least 4 minutes (CHAI et al. 2013; VERDIE & LAFARGE 2012).

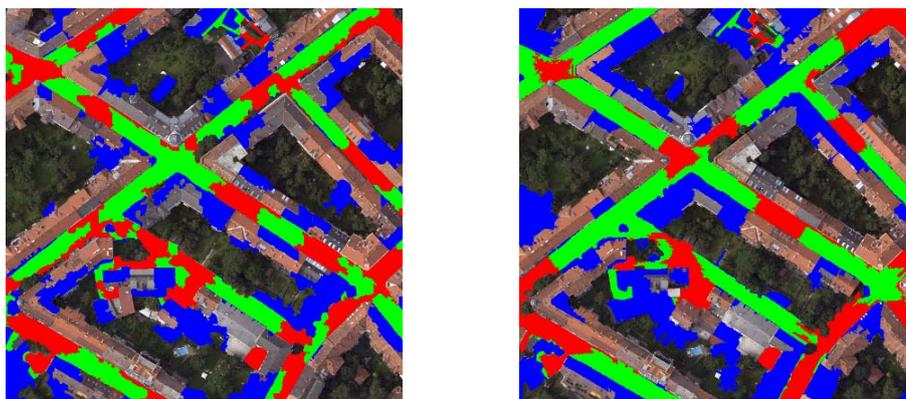


Fig. 3: Proposed road extraction results using SEEDs (left) and SLIC (right). The color indications are the same as in Fig.2.

Tab. 1: Quantitative evaluation of the proposed method. All numbers are percentages.

	Completeness	Correctness	Quality
[1]CRFs	64.80	46.21	36.94
[2]MPP	65.30	52.10	40.80
[3]Poisson	74.45	44.09	38.30

The road network extraction results are displayed in Figure 2 and estimated as Table 1. We first take use of our defined CRFs to get the pre-classification results, which is recorded as [1] CRFs. Then the marked point process is used and the results are marked as [2] MPP. It shows that the proposed MPP improves the result from CRFs on all the considered criteria. Some missing road parts are filled in and new paths are found on the left side of the image. Furthermore, we apply Poisson point process to the model with our proposed parametric points. We can find that with Poisson point process [3] provides a better completeness, by which more road parts are recognized, but the non-road background are false recognized either.

In addition, we evaluate the performance of our framework with different superpixel segmentation methods and the extraction results are shown in Figure 3 using SEEDs and SLIC respectively. Obviously using SLIC superpixels provides a better result than the points generated by SEEDs. The former extracts a network with a better connectivity, while SLIC segmentation usually provides more regular superpixels.

5 Conclusion and future work

We have proposed a new framework for extracting road networks from aerial images. In this work, shape features are extracted from the superpixel segments and then used in a simple pairwise CRFs model for image classification. The algorithms improves the sampling process by adopting the classification results and avoids complex geometric constrains. Comparing to the conventional marked point process relevant methods, the method improves the performance in terms of computing times and data scale.

The current results still have some problems needed to be fix. A better pre-classification results will lead to a more reasonable resolution. Some strategies are provided in many researches, e.g. using higher-order CRF, which can model the total energy of the road graph and ensure more about the connectivity. In the future work, we intend to study on more sophisticated CRFs as well as MPP model based on our point and feature.

6 References

- ACHANTA, R., SHAJI, A., SMITH, K., LUCCHI, A., FUA, P. & SÜSSTRUNK, S., 2012: SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34** (11), 2274-2281.
- BOYKO, A. & FUNKHOUSER, T., 2011: Extracting roads from dense point clouds in large scale urban environment. *ISPRS Journal of Photogrammetry and Remote Sensing* **66** (6), 2-12.
- BOYKOV, Y., VEKSLER, O. & ZABIH, R., 2001: Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23** (11), 1222-1239.
- BRÉDIF, M. & TOURNAIRE, O., 2012: Librjcmc: An open-source generic c++ library for stochastic optimization. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39** (B3).
- BREIMAN, L., 2001: Random Forests. *Machine Learning* **45** (1), 5-32.
- CHAI, D., FÖRSTNER, W. & LAFARGE, F., 2013: Recovering Line-Networks in Images by Junction-Point Processes. *Conference on Computer Vision and Pattern Recognition (CVPR)*, 1894-1901.
- DAL POZ, A.P., GALLIS, R.A., DA SILVA, J. & MARTINS, É., July 2012: Object-Space Road Extraction in Rural Areas Using Stereoscopic Aerial Images. *IEEE Geoscience and Remote Sensing Letters* **9** (4), 654-658.
- DESCOMBES, X. & ZERUBIA, J., 2002: Marked point process in image analysis. *IEEE Signal Processing Magazine* **19** (5), 77-84.
- GREEN, P.J., 1995: Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82** (4), 711-732.
- HE, C., YANG, F., YIN, S., DENG, X. & LIAO, M., 2013: Stereoscopic Road Network Extraction by Decision-Level Fusion of Optical and SAR Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **6** (5), 2221-2228.
- HU, X., TAO, C. V. & HU, Y., 2004: Automatic road extraction from dense urban area by integrated processing of high resolution imagery and LIDAR data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **35** (Part B3), 288-292.
- HUANG, Z., XU, F., LEI, L. & NIE, H., 2014: Object-based conditional random fields for road extraction from remote sensing image. *IOP Conference Series: Earth and Environmental Science* **17** (1).
- JIA, C.-L., JI, K.-F., JIANG, Y.-M. & KUANG, G.-Y., 2005: Road extraction from high-resolution SAR imagery using Hough transform. *IEEE International Geoscience and Remote Sensing Symposium* **1**, 336-339.

- 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* **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* **32** (9), 1597-1609.
- LAFFERTY, J., MCCALLUM, A. & PEREIRA, F. C., 2001: Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data.
- LIU, J., SUI, H., TAO, M., SUN, K. & MEI, X., 2013: Road extraction from SAR imagery based on an improved particle filtering and snake model. *International Journal of Remote Sensing* **22**, 34.
- LU, P., DU, K., WANG, R., DENG, Y. & BALZ, T., 2014: A New Region Growing-Based Method for Road Network Extraction and Its Application on Different Resolution SAR Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **7** (12), 4772-4783.
- MONTOYA-ZEGARRA, J.A., WEGNER, J.D., LADICKY, L. & SCHINDLER, K., 2014: Mind the gap: modelling local and global context in (road) networks. 36th German Conference on Pattern Recognition, GCPR Proceedings, 212-223.
- SHI, W., MIAO, Z. & DEBAYLE, J., 2014: An Integrated Method for Urban Main-Road Centerline Extraction From Optical Remotely Sensed Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 3359 - 3372.
- SINGH, P.P. & GARG, R., 2013: Automatic Road Extraction from High Resolution Satellite Image using Adaptive Global Thresholding and Morphological Operations. *Journal of the Indian Society of Remote Sensing* **41** (3), 631-640.
- SONG, M. & CIVCO, D., 2004: Road Extraction Using SVM and Image Segmentation. *Photogrammetric Engineering & Remote Sensing* **12**, 1365-1371.
- STOICA, R., DESCOMBES, X. & ZERUBIA, J., 2004: A gibbs point process for road extraction from remotely sensed images. *International journal of computer vision* **57** (2), 121-136.
- VALERO, S., CHANUSSOT, J., BENEDIKTSSON, J., TALBOT, H. & WASKE, B., 2010: Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images. *Pattern Recognition Letters* **31** (10), 1120-1127.
- VAN DEN BERGH, M., BOIX, X., ROIG, G., DE CAPITANI, B. & VAN GOOL, L., 2012: SEEDS: Superpixels Extracted via Energy-Driven Sampling. *Computer Vision-ECCV 2012*, 13-26.
- VERDIE, Y. & LAFARGE, F., 2012: Efficient Monte Carlo Sampler for Detecting Parametric Objects in Large Scenes. *Computer Vision – ECCV 2012, Firenze, Italy, Springer*, 539-552.
- WANG, H., CAI, Z., LUO, H. & WANG, C., 2012: Automatic road extraction from mobile laser scanning data. *International Conference on Computer Vision in Remote Sensing (CVRS)*, 136-139.
- WEGNER, J.D., MONTOYA-ZEGARRA, J.A. & SCHINDLER, K., 2013: A Higher-Order CRF Model for Road Network Extraction. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1698-1705.

- WEGNER, J.D., MONTOYA-ZEGARRA, J.A. & SCHINDLER, K., October 2015: Road networks as collections of minimum cost paths. *ISPRS Journal of Photogrammetry and Remote Sensing* **108**, 128-137.
- WIEDEMANN, C., HEIPKE, C., MAYER, H. & JAMET, O., 1998: Empirical evaluation of automatically extracted road axes. *Empirical Evaluation Techniques in Computer Vision*, 172-187.
- YUAN, J., WANG, D., WU, B., YAN, L. & LI, R., 2011: LEGION-Based Automatic Road Extraction From Satellite Imagery. *IEEE Transactions on Geoscience and Remote Sensing* **49** (11), 4528-4538.
- ZHANG, G. & JIA, X., 2012: Simplified Conditional Random Fields With Class Boundary Constraint for Spectral-Spatial Based Remote Sensing Image Classification. *IEEE Geoscience and Remote Sensing Letters* **9** (5), 856-860.
- ZHAO, J. & YOU, S., 2012: Road network extraction from airborne LiDAR data using scene context. Providence: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*.
- ZHONG, Y., ZHAO, J. & ZHANG, L., 2014: A Hybrid Object-Oriented Conditional Random Field Classification Framework for High Spatial Resolution Remote Sensing Imagery. *IEEE Transactions on Geoscience and Remote Sensing* **52** (11), 7023-7037.

Einzelbaumdetektion in bewaldetem Gebiet auf Basis von luftgestützten LiDAR-Daten

ANDREAS B. G. BAUMANN¹ & EMMANUEL BALTSAVIAS¹

Zusammenfassung: Diese Arbeit beschäftigt sich mit der Detektion von Einzelbäumen in bewaldetem Gebiet auf Basis von luftgestützten LiDAR-Daten (ALS). Das Untersuchungsgebiet ist ein naturnaher Hochwald am Südhang der Lägern (Aargau, Schweiz) mit Bäumen in unterschiedlichen Wachstumsphasen sowie mit Totholz. Zur Verfügung standen zwei Referenzdatensätze mit Informationen zur Lage und Art des Baumes, sowie dem jeweiligen Durchmesser (BHD) und zwei ALS-Datensätze (un-/belaubt) mit einer Punktdichte von 23.7 und 19.6 Echos/m². Die Punktwolken werden in den Voxelraum überführt und beginnend bei der untersten Ebene werden die Baumstammkandidaten vertikal verfolgt. Mit einer robusten linearen Regression wird die Orientierung des Baumstammkandidaten bestimmt. Weist die resultierende Gerade einen Zenitwinkel von weniger als 25° auf, wird die Lösung als detektierter Baumstamm deklariert. Mit diesem Algorithmus konnten 28.3% der Nadel- und 47.5% der Laubbäume detektiert werden und dies mit einer Nutzergenauigkeit von 81.3%. Es wurden hauptsächlich Bäume detektiert, die einen langen und freistehenden Baumstammbereich haben und es zeigte sich, dass Informationen über die Baumstämme hauptsächlich im unbelaubten ALS-Datensatz vorhanden sind und eine Kombination mit einem belaubten ALS-Datensatz zu keiner Verbesserung führt.

1 Einleitung

Seit Jahrtausenden nutzt der Mensch den Wald als Rohstoffquelle. Das Holz wurde und wird als Brennstoff und Baumaterial für diverse Güter verwendet. Der Wald hat mit seiner artenreichen Fauna und Flora neben dem finanziellen Aspekt auch eine soziale Funktion als Erholungsraum und eine weitreichende Schutzfunktion, namentlich als Boden- (Erosion), Lawinen- (Verhinderung und Verminderung), Wasser- (Speicherung und Filterung) und Immissionschutz (Luftfilterung). Eine weitere nicht zu vernachlässigende Funktion des Waldes gewinnt im Kontext des sich abzeichnenden Klimawandels (IPCC 2014) immer stärker an Bedeutung: Die Produktion von Sauerstoff und die damit einhergehende Speicherung von Kohlenstoff. In diesem Sinne sind für die Eigentümer der Wälder, ebenso wie für die Wissenschaft, das Wissen über die Zusammensetzung und Eigenschaften der einzelnen Bäume – wie der Zustand, die Größe und die Art – unerlässlich. Sie erlauben es zuverlässige Aussagen zu monetären Werten und der vorhandenen Biomasse zu treffen. Diese Angaben mussten – vor dem Aufkommen diverser Fernerkundungsmethoden – mit Hilfe von zeitintensiven Feldarbeiten in repräsentativen Flächen erfasst und auf die restlichen Gebiete extrapoliert werden. Seit etwa 1995 befassten sich zahlreiche wissenschaftliche Studien mit der flächendeckenden Ableitung solcher Informationen aus luftgestützten LiDAR-, multi- und hyperspektralen Datenakquisitionen (VAUHKONEN et al.

¹ Eidgenössische Technische Hochschule Zürich, Institut für Geodäsie und Photogrammetrie, Stefano-Francini-Platz 5, CH-8093 Zürich, E-Mail: kontakt@andreas-baumann.ch, manos@geod.baug.ethz.ch

2012). Bisher hat sich noch keine Methode durchgesetzt und das Forschungsfeld ist noch immer aktiv.

Ziel dieser Arbeit war es, zu untersuchen, welche Ansätze es ermöglichen in einem dicht bewaldeten Gebiet einzelne Bäume zu detektieren. Dabei soll nicht nur die Position des Baumstamms, sondern auch die Kronenausdehnung und die Baumklasse (Laub- und Nadelbaum) erkannt werden. Hierzu stehen ALS-Datenakquisitionen des Kantons Aargau von März/April (unbelaubt) und Juni/Juli (belaubt) 2014 zur Verfügung. Für die Untersuchung wurden diese beiden Datensätze auch kombiniert prozessiert. Mit einer Flughöhe von 600-700 m.ü. Grund konnten im Untersuchungsgebiet beim unbelaubten Datensatz eine Punktdichte von 23.7 Echos/m² und beim belaubten eine Dichte von 19.6 Echos/m² realisiert werden. Die Erfassung erfolgte durch MILAN Geoservice GmbH mit einem LMS-Q680i von RIEGL. Der verwendete full waveform (FW) Laserscanner ermöglicht es, eine sehr hohe Anzahl an Echos zu registrieren. Eine hohe Pulsrepetitionsrate von bis zu 400kHz ermöglicht eine effektive Messrate von 266,000 Bodenechos pro Sekunde. Das Sichtfeld beträgt maximal $\pm 30^\circ$ und hat eine Winkelauflösung von 0.001° . Es wird im Nahinfrarotbereich (1,550nm) gemessen, wobei die Strahlendivergenz weniger als 0.5mrad beträgt, was einem Footprint von 50cm auf 1,000m Distanz entspricht. Für die Messkampagne wurde eine Lage-Standardabweichung von 2 – 15 cm und für die Höhe von 5 – 7 cm angestrebt und innerhalb der Zielspuren auch weitestgehend erreicht. Für diese Arbeit standen ALS-Daten mit Positions- und Intensitätswerten zur Verfügung. Die FW Informationen (z.B. Echobreite) konnten nicht zur Verfügung gestellt werden. Für Visualisierungen und Interpretationen standen zudem multispektrale optische Daten (Orthophotos) zur Verfügung.

Das Untersuchungsgebiet befindet sich in einem Naturwaldreservat am südlichen Abhang der Lägern (47°28'42''N / 8°21'52''E), auf rund 700 Metern über Meer. Auf einer Fläche von 7.3 ha mit einer Neigung von bis zu 60% wachsen bis zu 50 Meter hohe Rotbuchen (*Fagus sylvatica*), gemeine Eschen (*Fraxinus excelsior*), Bergahorne (*Acer pseudoplatanus*), Weiß- und Rottannen (*Abies alba* und *Picea abies*) und dreizehn weitere Arten. Die Waldstruktur ist dabei sehr komplex mit Bodenbewuchs (Büsche und Sträucher), Jung und Altbäumen, teilweise geneigten und sich überkreuzenden Baumstämmen, sowie mit einer großen Anzahl an stehendem und liegendem Totholz (Abb. 1).



Abb. 1: Foto des naturbelassenen Hochwaldes auf der Lägern.

Für die quantitative Evaluation der Resultate standen zwei Referenzdatensätze zur Verfügung. Die Bauminventuren der «Eidgenössischen Forschungsanstalt für Wald, Schnee und Landschaft (WSL)» (ETZOLD et al. 2011) und des «Remote Sensing Labors der Universität Zürich (RSL)» (LEITERER et. al. 2015). Beide Datensätze beinhalten Informationen zur Art und Lage der Bäume. Das Inventar des WSL umfasst 659 Bäume, wobei alle Bäume mit einem Brusthöhendurchmesser (BHD) von mind. 4 cm erfasst wurden. Die Lage der Bäume wurde von Vermessungsfachleuten bestimmt und mit einer Lagegenauigkeit von 5cm ausgewiesen. Das Inventar des RSL umfasst 1,365 Bäume und es wurden nur Bäume mit einem BHD von mind. 20 cm erfasst. Das RSL-Untersuchungsgebiet wurde mit einem terrestrischen Laserscanner gescannt und die Lagekoordinaten der Bäume aus deren Punktwolke abgeleitet. Die Positionsangaben beziehen sich auf einen Punkt auf dem Baumstamm und weisen eine Genauigkeit von einem halben Meter aus. Die Datensätze überlagern sich teilweise, weswegen ein Vergleich der Referenzdaten möglich ist. In der folgenden Abb. 2 ist ein Orthophoto und zwei Ausschnitte der Referenzdatensätze erkennbar. In der östlich gelegenen Überschneidung (WSL-1 / RSL) ist eine Translation in Nord-Süd-Richtung von ungefähr zwei Metern erkennbar. Im westlichen Gebiet ist eine Rotation der Datensätze von 10° erkennbar, was zu einem Lageversatz von bis zu sieben Metern führt. Betrachtet man den Hintergrund und die darin sichtbaren Bäume, ist erkennbar, dass der Datensatz des WSL korrekt ist und der Datensatz des RSL in diesem Abschnitt eine Rotation aufweist.

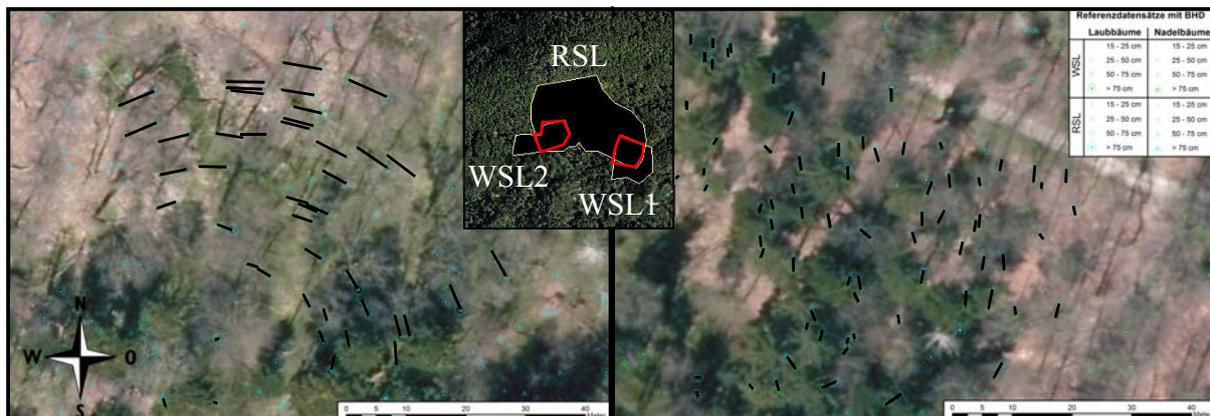


Abb. 2: Vergleich der Referenzdatensätze WSL-1/2 und RSL mit den entspr. Verschiebungsvektoren.

2 Prozessierungsmethode

Um einzelne Bäume in einem bewaldeten Gebiet in einer Punktwolke zu differenzieren, kann man sich grundsätzlich auf zwei Bereiche fokussieren: Die teilweise überlappenden Baumkronen können segmentiert oder die Baumstämme extrahiert werden. Bei Baumstämmen wird von der Annahme ausgegangen, dass sie in einer Punktwolke, mindestens teilweise, als lineare Objekte erkennbar sind. Bäume haben ein dichtes Blätterdach und infolgedessen ist insbesondere in der Nähe des Stammes nur ein niedriger Bodenbewuchs vorhanden. Eine Filterung nach freistehenden Punkten innerhalb einer horizontalen Ebene sollte es ermöglichen, Baumstämme zuverlässig zu detektieren.

2.1 Vorprozessierung

Um die Prozessierung zu beschleunigen und die zu verarbeitende Datenmenge zu reduzieren, erfolgt eine Vorprozessierung der Punktwolken. Wegen des dichten Bodenbewuchses und dem liegenden Totholz werden nur Echos, die mindestens zwei Meter über Grund liegen, berücksichtigt. Dieser Höhenschwellwert wurde verwendet, weil er in zahlreichen Publikationen (VAUHKONEN et al. 2012) angewandt und empfohlen wurde. Die nicht ausgeschlossenen Punkte werden zuerst in den Voxel- und anschließend in den Bildraum überführt. Der Voxelraum bildet sich aus regelmäßig angeordneten und nicht überlappenden Voxeln mit einer Kantenlänge von je einem Meter. Die Überführung erfolgt, indem die Echos dem jeweils umschließenden Voxel zugeteilt werden (Abb. 3a). Danach wird jeder Höhenstufe im Voxelraum ein Bild und jedem Voxel innerhalb der Höhenstufe ein entsprechendes Pixel im Bild (Abb. 3b) zugeordnet. Im Bildraum können Informationen als Grauwerte gespeichert werden. Für die Baumstammverfolgung wurde eine Schwarz-Weiß-Abbildung gewählt. Sofern einem Voxel ein oder mehrere Punkte zugeordnet sind, wird dem entsprechenden Pixel der Wert «1», ansonsten der Wert «0» zugewiesen. Ein Parameter minimale Intensität wird hier als Selektionskriterium eingeführt. Bei genauer Betrachtung der Punktwolken hat sich herausgestellt, dass schwache Intensitäten oftmals auf dünne Äste/Blätter oder Multipath zurückzuführen sind. Mit einem empirisch festgelegten Schwellwert von 15dB können diese herausgefiltert werden. Dadurch bleiben nur Baumstämme und andere großflächige Objekte erhalten.

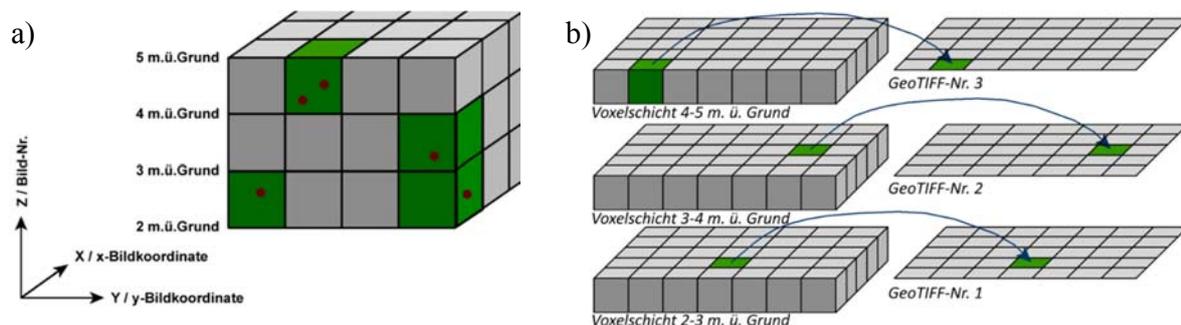


Abb. 3: (a) Zuordnung der ALS-Echos im Voxelraum und (b) Übertragung in den Bildraum.

2.2 Baumstammextraktion mittels Linienverfolgung

2.2.1 Dreistufige Filterung

Die dreistufige Filterung soll die Daten prozessieren, damit nur noch die Echos erhalten bleiben, die mit großer Wahrscheinlichkeit einem Baumstamm zugeordnet werden können. Die erste Filterstufe befasst sich mit der *maximalen Pixelgruppengröße*. Ein visueller Vergleich zwischen den generierten Bildern und den Referenzdaten zeigt, dass Baumstämme oftmals als Einzelpixel oder Zweierpixelgruppen auftreten (Abb. 4a). Größere Pixelgruppen sind meist auf Bodenbewuchs oder den Baumkronenbereich zurückzuführen. Als mögliche Kandidaten werden deshalb nur Pixelgruppen von bis zu zwei Pixeln (Parameter: Pixelgruppengröße) innerhalb einer 8er-Nachbarschaft der jeweiligen Höhenschicht berücksichtigt.

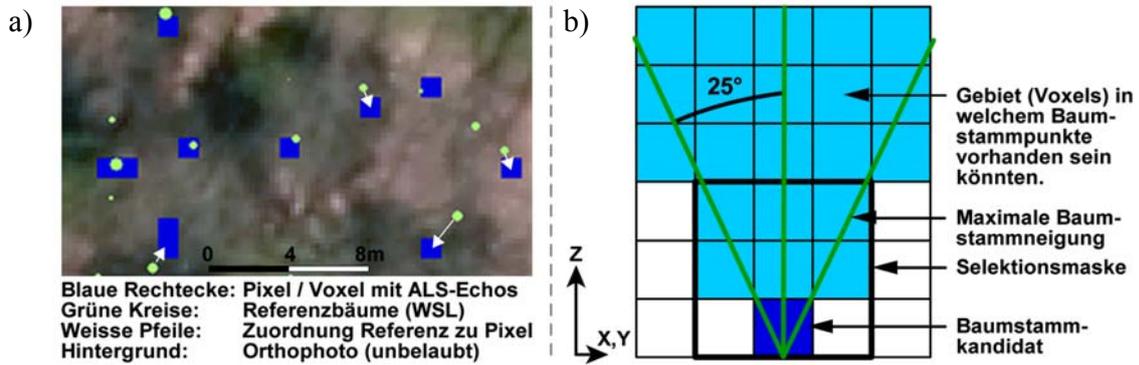


Abb. 4: (a) Ausschnitt aus Untersuchungsgebiet 1 (b) Querschnitt durch Voxelraum.

Die zweite Filterstufe widmet sich der Linienverfolgung. Die gefilterten Pixel einer jeden Höhenstufe sollen miteinander verknüpft werden (Clustering). Es wird die Annahme getroffen, dass ein Baum grundsätzlich vertikal wächst, aber infolge von Umwelteinflüssen (Instabilität des Bodens durch die Hangneigung, ständige, starke, gleichgerichtete Winde infolge der Exposition, etc.) bis zu 25° Grad geneigt sein kann (Abb. 4b). Die Linienverfolgung beginnt in der untersten Höhenstufe und jede darin enthaltene Pixelgruppe wird angesteuert. Über jede Pixelgruppe wird eine dreidimensionale Selektionsmaske gelegt, welche in x/y-Richtung um ein Voxel und in die positive z-Richtung um zwei Höhenstufen erweitert wird (Parameter: Selektionsmaskenhöhe). Alle Voxel, die sich innerhalb dieses Selektionsfilters befinden, werden einem Cluster zugeordnet. Sind Voxel bereits einem anderen Cluster zugeordnet, werden die Cluster vereint. Dies wird für alle Stufen und Voxel durchgeführt und am Ende werden alle Cluster denen mindestens drei oder mehr Voxel zugeordnet sind als potenzielle Baumstämme deklariert.

Abhängig von den gewählten Parametern können nahe beieinander stehende Bäume zu einem Cluster vereint werden. Um solche Baumgruppencluster aufzulösen, wurde ein empirischer Schwellenwert festgelegt: Sobald einem Baumcluster 40 oder mehr Echos zugeordnet sind, wird dieser aufgesplittet. Hierzu wird in der xy-Ebene eine Fläche aufgespannt. An jeder Position werden die in der z-Ebene übereinanderliegenden Voxel aufsummiert. Daraus resultiert ein Histogramm (Abb. 5a-b) an welches anschließend eine Segmentierung der Cluster via Watershed-Algorithmus (Abb. 5c) angewandt wird.

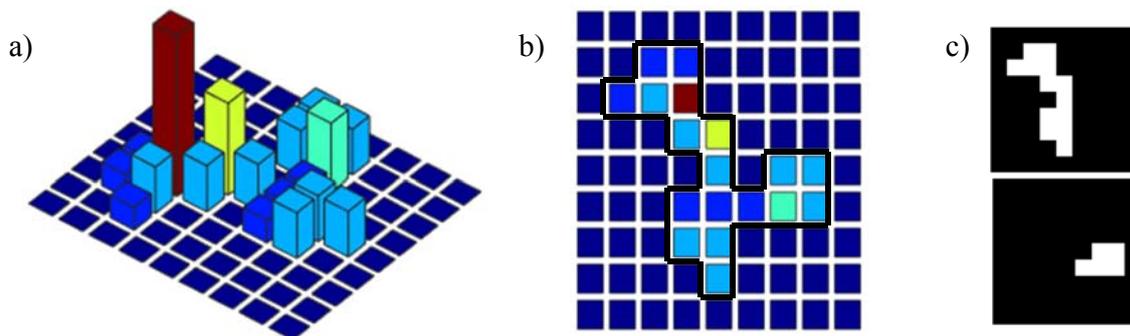


Abb. 5: Voxelhistogramm eines Baumstammkandidaten in der Seitenansicht (a) und Aufsicht (b), sowie den resultierenden Segmenten (c).

Die Cluster der potenziellen Baumstämme enthalten zahlreiche ALS-Punkte, die nicht nur Teil des Stammes sondern auch Bodenbewuchs, Nebenäste und Baumkronen wiedergeben. Um diese zu reduzieren wird in der dritten Filterstufe der freistehende Baumstammbereich detektiert. Dieser ist definiert als die Zone zwischen der Oberkante Bodenvegetation (OB) und der Unterkante Baumkrone (UB). Im Baumkronen- und Bodenvegetationsbereich gibt es im Vergleich zum Baumstammbereich zahlreiche Echos (Abb. 6). Die Vegetationskanten werden durch Segmentierung des Höhen-Histogramms der ALS-Punktswolke auf Basis der Dichteänderung bestimmt und beinhaltet folgende Schritte:

- a) Approximation der Baumstammposition. Das zuvor berechnete Voxelhistogramm wird hinzugezogen. Die Cluster-Voxel die der höchsten Säule zugeordnet sind werden selektiert und die enthaltenen Echos ausgelesen. Die Baumstammposition wird dann aus dem Mittelwert der X- und Y-Koordinaten dieser Echos geschätzt.
- b) Im Umkreis von fünf Metern um die approximierten Baumstammposition werden alle Echos des unbelaubten Datensatzes (Abb. 6, rote Kreuze), die nicht der Klasse Boden zugeordnet sind, selektiert und für die Berechnung des relativen Höhenhistogramm (Schichtdicke: ein Meter) verwendet.
- c) In Abb. 6 ist gut zu erkennen, dass die Histogramme lokale Maxima aufweisen, welche mit einem gleitenden Mittelwert (3er Nachbarschaft in der vertikalen Richtung) reduziert werden können.
- d) Ausgehend von der höchsten Säule im Histogramm (höchste Dichte in %) wird die UB gesucht. Die UB wird jeweils um einen Meter Richtung Boden verschoben, solange bis die folgenden beiden, empirisch festgelegten Bedingungen erfüllt sind:
 - Schwellenwert A = $1/7 \cdot \text{Säulenhöhe}_{\text{höchste Dichte}} \quad \text{Säulenhöhe}_i \leq \text{Schwellenwert A}$
 - Schwellenwert B = $1/70 \cdot \text{Säulenhöhe}_{\text{höchste Dichte}} \quad \text{Nachbardifferenz}_i \leq \text{Schwellenwert B}$
- e) Für die Oberkante der Bodenvegetation wird anschließend die höchste Säule die unterhalb der UB liegt ausgelesen. Hat diese eine höhere Dichte als der zuvor bestimmte Schwellenwert A, erfolgt die Detektion der OB nach dem gleichen Verfahren wie zur Bestimmung von UB, doch in umgekehrter Richtung. Hat die Säule eine geringere Dichte, wird die OB auf zwei Meter über Boden festgelegt.

2.2.2 Lineare Regression und RANSAC

Als nächstes wird überprüft, ob sich mindestens zwei dem Baumcluster zugeordnete Echos zwischen OB und UB befinden (Abb. 6, gelbe Ebenen). Trifft dies zu, könnte es sich um einen Baumstamm handeln. Für alle Echos, die dem Baumcluster zugeordnet sind und sich zwischen OB und UB befinden (Abb. 6, blaue Kreise), wird die Streuungsellipse (95% Konfidenzintervall) der projizierten Lage berechnet. Empirisch wurde eine minimale Halbachsenlänge von 75cm festgelegt, um eine zu starke Vorselektion zu verhindern. Diese Streuungsellipse wird zwischen OB und UB aufgespannt (Abb. 6, schwarze Ellipsen). Alle unbelaubten Echos die sich innerhalb des Zylinders befinden, werden für die Berechnung der Baumstammorientierung hinzugezogen. Dieser Schritt wird durchgeführt, weil Baumstammunkte – infolge fehlender Echos – zu weit auseinander liegen könnten, um miteinander verknüpft zu werden oder weil einfallende Äste zu große Pixelgruppen bilden. Auf die selektierten Punkte wird eine robuste lineare Regression

durchgeführt, um eine Gerade zu bestimmen, wobei das robuste Fitting auf dem RANSAC-Algorithmus (FISCHLER & BOLLES 1981) basiert. Von den Lösungen wird jene bevorzugt, welche die geringste Anzahl an groben Fehlern (Abstand von Gerade $\geq 30\text{cm}$) und den kleinsten Zenitwinkel aufweist. Ist der Zenitwinkel kleiner als 25° , wird die Lösung als «detektierter Baumstamm» deklariert und abgespeichert.

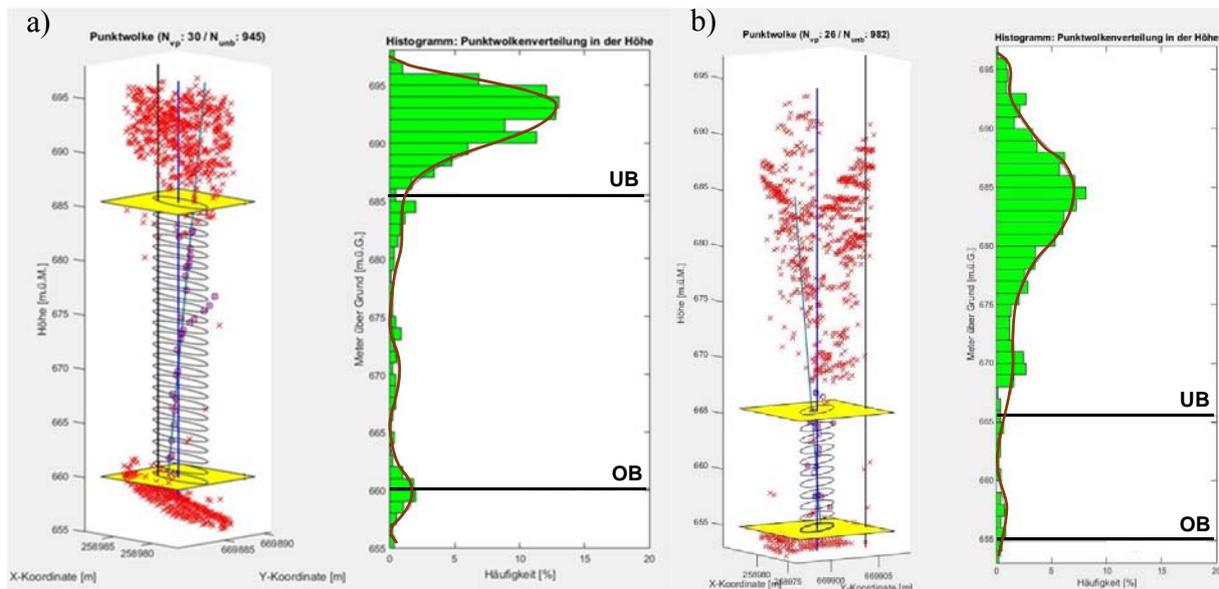


Abb. 6: Punktwolke und Höhenhistogramm eines (a) korrekt und (b) falsch detektierter Baumstammes. Punktwolke: Vertikale schwarze Linie: Baumstammposition nach Referenzdaten; Vertikale blaue Linie: approx. Baumstammposition; Nahezu vertikale, grüne Linie: Detektierter Baumstamm. Rote Kreuze: unb. Echos im Umkreis von 5 Meter um die approx. Baumstammposition; Blaue Kreise: Dem Baumcluster zugeordnete Echos; Gelbe Ebene: Vegetationskanten; Schwarze Ellipsen: Streuungsellipse (95%); Höhenhistogramm: Grüne Balken: Relative Punktdichte je Höhenschicht; Braune Polylinie: Gleitendes Mittel der relativen Punktdichte. Überschriften: NVP: Anzahl Echos die dem Baumcluster zugeordnet sind; Nunb: Gesamtzahl an unbelaubten Echos innerhalb des Umkreises von 5 Metern.

2.3 Baumstammextraktion mittels Lückendetektion

Die Baumstammextraktion mittels Lückendetektion ist eine zweite Methode, die getestet wurde. Bei dieser Methode wird nach etwas gesucht, das nicht vorhanden ist, also eine Lücke aufweist. Dabei kann davon ausgegangen werden, dass – wenn man den Boden und die niedrige Vegetationsschicht betrachtet – Punkte überall auftauchen, ausser dort, wo massive Objekte sind, die über die niedrige Vegetationsschicht hinausragen (Abb. 7). Wenn diese Lücken betrachtet werden, sind die ovalen Objekte mit grosser Wahrscheinlichkeit Baumstämme.

Für die Evaluation des Algorithmus wurden zwei Datensätze verwendet: Die kombinierten ALS-Punktwolken (un-/belaubt) und Aufnahmen eines terrestrischen Laserscanners, welche im Rahmen der Bachelorarbeit «Erfassung archäologischer Steinreihen (Reusseggmauer) mittels terrestrischem Laserscanning» durchgeführt wurden.

Die Detektion läuft in drei Hauptschritten ab. Der erste Schritt beinhaltet die Selektion der relevanten Punkte. Als Selektionskriterium wird ein Höhenwert (Meter über Grund) festgelegt

und alle Echos die darunter liegen werden selektiert. Im zweiten Schritt werden diese Echos auf eine Ebene projiziert und als Rasterdatei abgespeichert. Der letzte Schritt beinhaltet die eigentliche Objektdetektion auf Basis der Rasterdatei und wurde im Rahmen dieser Arbeit nur manuell mittels visueller Interpretation durchgeführt. In der Rasterdatei wurde nach kreisförmigen Lücken gesucht, welche von Baumstämmen hinterlassen wurden. Es zeigte sich, dass dieser Ansatz für die Baumstammedetektion erst ab einer Punktdichte von 25 Echos/m² eindeutige Ergebnisse liefern würde und deshalb wurde dieser Ansatz nicht weiter verfolgt.

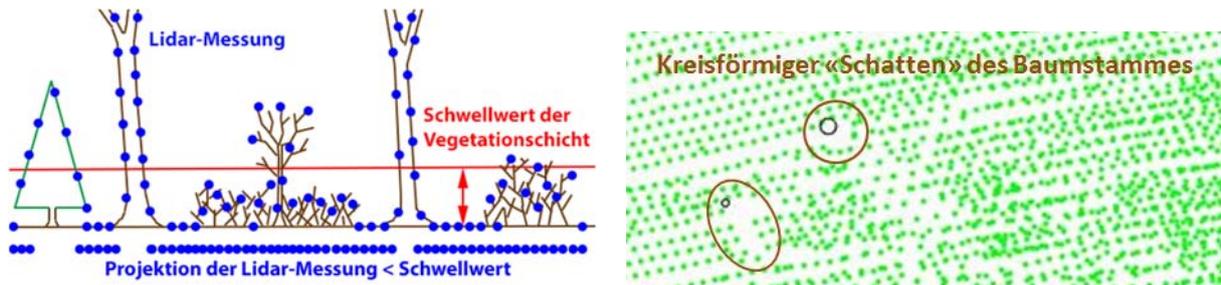


Abb. 7: (a) Querschnitt eines skizzierten Waldstücks und (b) Visualisierung aller Echos die bis zu einer Höhe von 2m über Grund liegen (grüne Punkte) und die Referenzdaten (schwarze Kreise)

3 Resultate und Diskussion

Der Algorithmus wurde mit unterschiedlichen Parametern (Pixelgruppengröße und Selektionsmaskenhöhe) getestet. Der Einfluss der Echofilterung auf Grund einer minimalen Intensität und des Splittings der Voxelpuppen wurde ebenso analysiert. Für die Überprüfung der Resultate wurden die detektierten Bäume zusammen mit den Referenzdaten (WSL-1/-2) visualisiert und anschließend manuell zugeordnet. In Abb. 8 sind die Resultate dieser Auswertung für das erste, Gebiet (WSL-1) ersichtlich. Es kann festgestellt werden, dass sich – mit Berücksichtigung einer minimalen Intensität von 15dB – die Erkennungsrate von Nadelbäumen nur geringfügig verändert (31.9% → 29.7%), hingegen hat sie bei Laubbäumen einen signifikanten Einfluss. Sie steigt um 7.4% von 43.6% auf 51%. Mit zunehmender Selektionsmaskenhöhe (3-5 Meter) nimmt die Anzahl Detektionen ab. Weiterhin ist mit zunehmender Pixelgruppengröße (2-3 Pixel) eine Abnahme der Detektionsrate von 5-10% feststellbar. In den Resultaten zeigte sich ebenfalls, dass das Splitting der Voxelpuppen essentiell ist. Abhängig von der Wahl der Parameter haben die «Splitting»-Bäume einen prozentualen Anteil zwischen 4 und 51% aller detektierten Bäume. Würde kein Splitting durchgeführt, könnte pro Voxelpuppe nur jeweils ein Baum detektiert werden, was zu einer stark verringerten Detektionsrate führen würde.

Durchschnittlich konnten $41.5 \pm 2.9\%$ aller Bäume im Gebiet detektiert werden, was 122 von insgesamt 293 Exemplaren ($BHD \geq 15\text{cm}$) entspricht. Dabei ist ein Unterschied zwischen Nadel- und Laubbäumen feststellbar. Nur $28.3 \pm 2.7\%$ der Nadelbäume konnten detektiert werden (26 von 91 Exemplaren), demgegenüber steht eine Erkennungsrate von $47.5 \pm 3.5\%$ bei Laubbäumen (96 von 202 Exemplaren). Nicht nur wie viele Bäume detektiert werden, ist wichtig um einen Algorithmus einschätzen zu können, sondern auch, wie zuverlässig das Resultat ist. Im Gebiet wurden bis zu 10% der Bäume mehrfach detektiert, was darauf zurückzuführen ist, dass die Cluster aufgrund ihrer großen Höhendifferenz, nicht miteinander vereint werden konnten. Des

Weiteren wurden 28 Objekte fälschlicherweise als Bäume detektiert, was einer Nutzergenauigkeit von $81.3 \pm 1.3\%$ entspricht.

Bei näherer Untersuchung der falsch detektierten Bäume hat sich gezeigt, dass sie sehr oft bei kleinen Lichtungen auftreten, beispielsweise dort wo Tannen sehr nahe beieinander stehen. Durch das dichte Baumkronendach dringen nur wenige Signale hindurch und dort wo die Kronen sich überschneiden, entsteht ein Trichter, an dem die Pulse reflektiert werden und die registrierten Echos eine Linie bilden. Die Punktwolken (Abb. 6, rot) und dazugehörigen Höhenhistogramme (Abb. 6, grün) zeigen in diesem Fall ein ähnliches Bild wie bei einem «echten» Baumstamm.

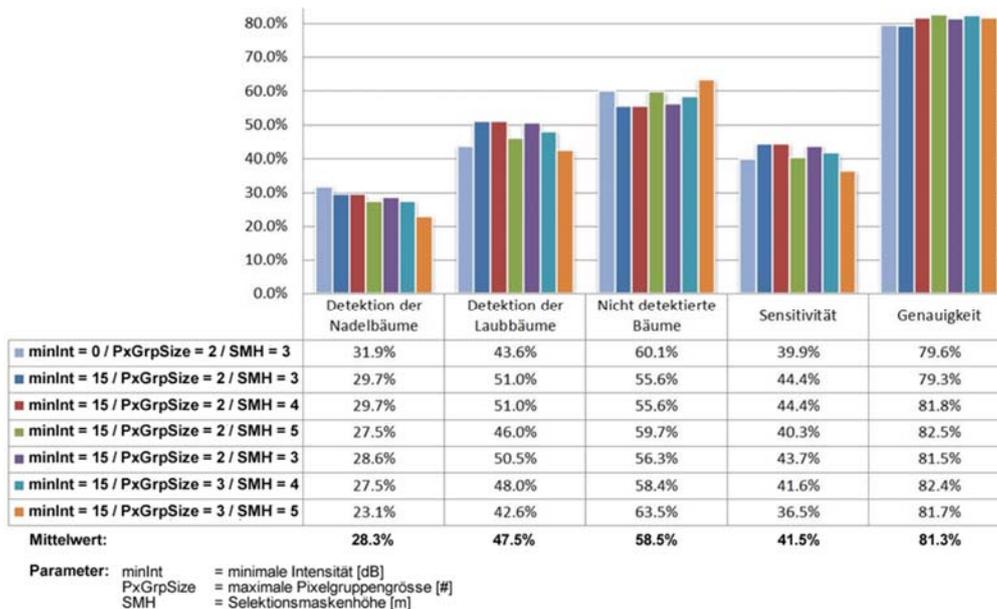


Abb. 8: Genauigkeit der Baumstammdetektion für Bäume mit BHD ≥ 15 cm im Referenzgebiet 1 (WSL-1).

4 Fazit

Es zeigte sich, dass für die Extraktion mittels Linienverfolgung nicht alle Echos verwendet werden sollten. Auf die belaubten ALS-Daten kann verzichtet werden, weil Messungen am Baumstamm fast ausschließlich in den unbelaubten ALS-Daten vorhanden sind. Des Weiteren sollten Punkte mit einer niedrigen Intensität vorgängig gefiltert werden, dadurch können 5-10% mehr Baumstämme detektiert werden. Ebenfalls ist das Splitting der Baumstammkandidaten wichtig, weil je nach gewählten Parameterwerten Cluster mit mehreren Bäumen gebildet werden. Die Erkennung der Kronenunterkante und der Bodenvegetationsoberkante funktioniert zuverlässig, weswegen der Baumstammbereich gut extrahiert werden kann. Es wurden nur die dem Cluster zugewiesenen Echos berücksichtigt, welche zwischen den Vegetationskanten liegen. Gab es Ausreißer bei diesen Echos oder handelte es sich bei diesen um Bodenvegetations- oder Kronenpunkte, wurde dies nicht überprüft und direkt versucht eine Linie an diese Punkte anzupassen. Dies kann unter Umständen zu falsch detektierten Bäumen führen.

Betrachtet man die Anzahl an detektierten Bäume, stellt man fest, dass nur gut 28.3% der Nadelbäume und 47.5% der Laubbäume detektiert werden können. Es konnten fast nur hohe Bäume, mit einem entsprechend langen, freistehenden Baumstamm detektiert werden. Jungbäume können, in Folge mangelnder Baumstammpunkte, nicht detektiert werden. Im Rahmen einer zukünftigen Arbeit sollte untersucht werden, ob die Fullwave-Informationen, insbesondere die Echobreite, durch eine verbesserte Echo-Filterung zu einer besseren Erkennungsrate führt. Ebenfalls könnte untersucht werden, ob die ALS-Punktwolke – basierend auf der Dichteänderung des Höhenhistogramms – separiert werden sollten, um die jeweiligen Segmente mit unterschiedlichen Algorithmen zu prozessieren. So könnten die hohen Bäume mit der Baumstammextraktion und die niedrigen Bäume mit einem Watershed-Algorithmus detektiert werden.

Zusammenfassend kann gesagt werden, dass es alleine unter Zuhilfenahme der Baumstammextraktion nicht möglich ist, alle Bäume zu erfassen. Es ist unerlässlich, dass andere Methoden (bspw. Watershed auf Canopy Height Models) und weitere Daten (bspw. multi-/hyperspektrale Daten) hinzugezogen werden, um eine vollständigere Erfassung des Baumbestandes zu ermöglichen.

5 Literaturverzeichnis

- ETZOLD, S., RUEHR, N. K., ZWEIFEL, R., DOBBERTIN, M., ZINGG, A., PLUESS, P., HÄSLER, R., EUGSTER, W. & BUCHMANN, N., 2011: The carbon balance of two contrasting mountain forest ecosystems in Switzerland: Similar annual trends, but seasonal differences. *Ecosystems* **14** (8), 1289-1309.
- FISCHLER, M. & BOLLES, R., 1981: Random Sample Consensus: Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM* **24** (6), 381-395.
- IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Core Writing Team, Pachauri, R.K & Meyer, L.A. (eds.), IPCC, Switzerland, 151 p.
- LEITERER, R., FURRER, R., SCHAEPMAN, M. E., & MORSORF, F., 2015: Forest canopy-structure characterization: A data-driven approach. *Forest Ecology and Management* **358**, 48-61.
- VAUHKONEN, J., ENE, L., GUPTA, S., HEINZEL, J., HOLMGREN, J., PITKÄNEN, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T. & Maltamo, M., 2012: Comparative testing of single-tree detection algorithms under different types of forest. *Forestry* **85** (1), 27-40.

Klassifikation von Stereobildern aus Mobile Mapping Daten mittels Conditional Random Fields

MAX COENEN¹ & FRANZ ROTTENSTEINER¹

Zusammenfassung: In dieser Arbeit wird ein neues Verfahren zur kontextbasierten Klassifikation von Stereobildern mittels Conditional Random Fields (CRF) vorgestellt. Die Klassifikation setzt auf Segmenten als Knoten für das CRF auf. Die Segmentierung erfolgt im Bildraum und wird mittels einer 3D-Rekonstruktion der Szene auf die 3D-Punktwolke übertragen, was die Extraktion von 3D-Merkmalen zusätzlich zu den Bildmerkmalen sowie die Definition von realistischen Nachbarschaftsbeziehungen zwischen den Segmenten im Objektraum ermöglicht. Die Evaluierung der Methode erfolgt anhand von im urbanen Raum aufgenommenen Stereosequenzen eines Benchmark Datensatzes und liefert Ergebnisse mit einer Gesamtgenauigkeit von bis zu über 90%. Außerdem wird gezeigt, dass die Berücksichtigung von Kontext in der Klassifikation zu einer Erhöhung der Gesamtgenauigkeiten führt.

1 Einleitung

Die Nutzung von mobilen Systemen zur Erfassung von 3D-Geodaten über berührungslos aufnehmende Sensoren gewinnt v.a. im urbanen Raum zunehmend an Bedeutung. Zum Einsatz kommen dabei in der Regel hohe Punktdichten erzeugende Sensoren wie 3D-Laserscanner oder Stereokameras, wobei letztere zusätzlich zu Tiefeninformationen auch Farbinformationen liefern und deutlich kostengünstiger sind. Für viele Anwendungen, wie z.B. zur Generierung von 3D-Stadtmodellen oder in der Fahrzeug- und Roboternavigation, ist eine Klassifikation der 3D-Daten, also eine Zuweisung von semantischen Objektklassen an Teile der Szene, essentiell. So nutzen ANGULOV et al. (2005) zum Beispiel geometrische 3D-Punkt-Merkmale für die punktweise Klassifikation einer Laserscan-Punktwolke, während MATTI & NEBIKER (2014) zusätzlich zu geometrischen Merkmalen auch Farbinformation zur Klassifikation einer texturierten Laserscan-Punktwolke verwenden. SENGUPTA et al. (2013) führen eine pixelweise Klassifikation von Stereobildern aus Mobile Mapping Daten durch, wofür sie ausschließlich Bildinformationen als Merkmale nutzen und übertragen das Klassifikationsergebnis auf die aus den Stereobildern rekonstruierte 3D-Punktwolke. Um die Rechenkomplexität von punkt- bzw. pixelweise klassifizierenden Methoden zu verringern, kann eine generalisierende Datenreduktion, z.B. in Form von einer Segmentierung der Daten, sinnvoll sein, wonach die Segmente anstelle der einzelnen Punkte bzw. Pixel klassifiziert werden. So segmentieren LIM & SUTER (2009) beispielsweise terrestrisch aufgenommene Punktwolken in sogenannte Supervoxel, welche sie anhand von für jedes Supervoxel berechneten Merkmalen klassifizieren.

Insbesondere im urbanen Raum ist die Klassifikation aufgrund von komplexen und vielfältigen Strukturen und Objektklassen eine Herausforderung. Doch gerade in von Menschen gemachten Umgebungen weisen Objekte spezifische räumliche Beziehungen zueinander auf, welche als

¹ Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Str. 1, D-30167 Hannover, E-Mail: [coenen, rottensteiner]@ipi.uni-hannover.de

Kontextinformation mithilfe sogenannter Graphischer Modelle in der Klassifikation berücksichtigt werden können. Ziel dieser Arbeit ist es, einen solchen kontextbasierten Ansatz für die Klassifikation von aus einem bewegten Fahrzeug aus aufgenommenen Stereobildern bzw. daraus abgeleiteten 3D Punktwolken auf Basis von Conditional Random Fields (CRF) (KUMAR & HEBERT 2006), in welchen die statistischen Abhängigkeiten zwischen benachbarten Objekten explizit modelliert werden, zu entwickeln und zu testen.

2 Methodik

Das Ziel der Klassifikation besteht darin, jedem 3D-Punkt einer aus Stereobildern abgeleiteten 3D-Punktwolke ein Klassenlabel aus einer Menge à priori definierter Objektklassen zuzuweisen. Zu diesem Zweck wurde ein segmentbasiertes Klassifikationskonzept entwickelt, welches aus Stereobildsequenzen generierte 3D-Punktwolkensegmente klassifiziert. Hierbei werden die Stereobildpaare einzeln betrachtet und jedes Stereomodell separat klassifiziert. Der Ablauf der Klassifikation ist in Abb. 1 dargestellt und gliedert sich in mehrere Schritte, auf welche im Folgenden näher eingegangen wird.

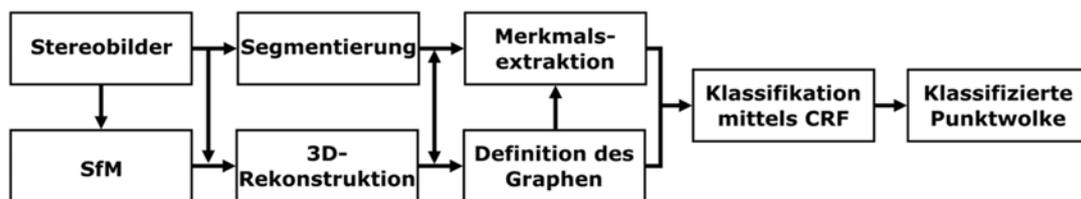


Abb. 1: Übersicht über das Konzept der Klassifikationsstrategie

2.1 3D-Rekonstruktion

Das Ziel der 3D-Rekonstruktion ist die Generierung von 3D-Punktwolken aus den Stereobildern in einem für die gesamte Stereosequenz einheitlichen Koordinatensystem, welche die Berechnung von zusätzlichen und für die Klassifikation sehr wertvollen geometrischen 3D-Merkmalen ermöglichen. Ausgehend von bekannten Parametern der inneren Orientierung und der relativen Orientierung der beiden Kameras wird zunächst mittels eines *Structure from Motion* (SfM) Verfahrens ähnlich zu (REICH et al. 2013) die äußere Orientierung aller Bilder der Stereosequenz ermittelt. Anschließend wird mithilfe des *Efficient Large-Scale Stereo Matching* (ELAS-) Verfahrens (GEIGER et al. 2011) ein dichtes Parallaxenfeld bestimmt, aus dem für jedes Pixel des linken Stereopartners mittels Triangulation unter Berücksichtigung der äußeren Orientierung ein 3D-Punkt im Objektkoordinatensystem rekonstruiert wird. Man erhält so für jedes Stereopaar eine dichte 3D-Punktwolke, die die Grundlage für die weitere Verarbeitung darstellt.

2.2 Datenreduktion durch Segmentierung

Die große Anzahl an Punkten, welche durch die dichte Bildzuordnung generiert wird, würde sich negativ auf die Rechenkomplexität der kontextbasierten Klassifikation auswirken. Um dem entgegenzuwirken und weil Segmente stabilere Merkmale als einzelne Punkte oder Pixel liefern können, wird eine Segmentierung der Daten durchgeführt. In den darauffolgenden Schritten

werden nicht die Einzelpunkte, sondern die Segmente klassifiziert. Aufgrund der geringeren Rechenkomplexität im Vergleich zu Verfahren zur Segmentierung von Punktwolken erfolgt die Segmentierung zunächst in einem der Bilder (Referenzbild, in dieser Arbeit der linke Stereopartner). Ihr Ergebnis wird auf die 3D Punktwolke übertragen, indem jeder 3D-Punkt dem Segment jenes Pixels im Referenzbild zugeordnet wird, das zur Bestimmung seiner Koordinaten beitrug. Zur Segmentierung wird das *Simple Linear Iterative Clustering* (SLIC-) Verfahren (ACHANTA et al. 2012) verwendet, welches die Pixel des Bildes in Regionen, sogenannte Superpixel, gruppiert, wobei sich die Segmentgröße als Anzahl der Pixel pro Segment (Px/Seg) vorgeben lässt. Abb. 2 zeigt das Beispiel einer SLIC-Segmentierung mit unterschiedlichen Segmentgrößen. Da während der Klassifikation jedem Segment nur eine Objektklasse zugewiesen werden kann, ist die Einhaltung von Objektgrenzen durch die Segmentierung wichtig. Wie man Abb. 2 entnehmen kann, passen sich die Segmentgrenzen den Objektgrenzen größtenteils sehr gut an (z.B. Gebäudedach, Vegetation, Stromkasten), mit zunehmender Größe der Segmente (z.B. 450 Px/Seg in Abb. 2) werden allerdings kleinere Objekte, wie z.B. die Stange des Straßenschilds oder auch der Bürgersteig, mit benachbarten Segmenten verschmolzen. Wählt man kleinere Segmentgrößen (z.B. 250 Px/Seg), passt sich das Segmentierungsergebnis den Objektgrenzen besser an, die daraus resultierende größere Zahl an Segmenten lässt allerdings eine etwas längere Rechenzeit erwarten.

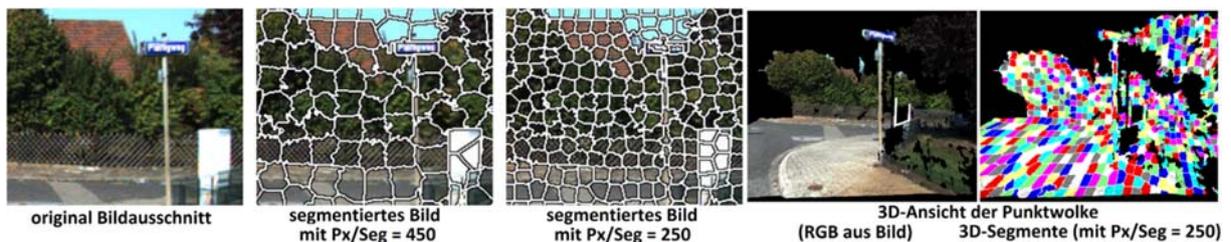


Abb. 2: Ergebnisse der SLIC Segmentierung mit unterschiedlicher Segmentgröße

2.3 Conditional Random Fields (CRF)

CRFs gehören zu den probabilistischen, diskriminativen und Kontext berücksichtigenden Klassifikationsverfahren. Die Berücksichtigung von Kontext erfolgt mithilfe eines statistischen Ansatzes, der die Abhängigkeiten zwischen benachbarten Primitiven modelliert. Dadurch beruht die Klassifizierung der Primitive nicht mehr allein auf ihren Merkmalen, sondern auch auf ihren Nachbarprimitiven. In einem CRF wird die a posteriori Verteilung $P(\mathbf{C}|\mathbf{x})$ der Klassenlabels aller Primitive (gesammelt in einem Vektor \mathbf{C}) bei gegebenen Daten \mathbf{x} direkt modelliert und kann durch folgende Gleichung ausgedrückt werden (KUMAR & HEBERT 2006):

$$P(\mathbf{C}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \exp \left(\sum_{i \in \mathbf{n}} A(\mathbf{x}, C_i) + \sum_{e_{ij} \in \mathbf{e}} I(\mathbf{x}, C_i, C_j) \right). \quad (1)$$

Dabei wird $A(\mathbf{x}, C_i)$ als *Assoziationspotenzial* und $I(\mathbf{x}, C_i, C_j)$ als *Interaktionspotenzial* bezeichnet; diese Potenziale werden in Abschnitt 2.3.2 näher beschrieben. $Z(\mathbf{x})$ dient als Normalisierungskonstante. Das CRF wird durch einen Graphen $G = (\mathbf{n}, \mathbf{e})$ repräsentiert, der aus einer Menge von Knoten \mathbf{n} und einer Menge von Kanten \mathbf{e} besteht (BISHOP 2006). Die Knoten ent-

sprechen den zu klassifizierenden Punktwolkensegmenten, der Index i bezeichnet einen spezifischen Knoten in \mathbf{n} . Der Vektor \mathbf{C} enthält die Labels C_i für jeden Knoten i und hat demnach dieselbe Anzahl an Elementen wie \mathbf{n} . Eine Kante e_{ij} aus \mathbf{e} verbindet einen Knoten i mit seinem Nachbarn j ; diese Kanten repräsentieren die statistischen Abhängigkeiten zwischen benachbarten Knoten. Ziel der Klassifikation ist es, jene Konfiguration \mathbf{C} von Klassenlabels zu finden, für die $P(\mathbf{C}|\mathbf{x})$ maximal wird. Dabei werden die Parameter von $P(\mathbf{C}|\mathbf{x})$, d.h. die Parameter der Potentiale in Gleichung 1, aus Trainingsdaten gelernt. Die sich aus der Klassifikation der Segmente ergebenden Klassen werden anschließend auf die zugehörigen 3D-Punkte übertragen.

2.3.1 Definition des Graphen und der Merkmale

Die Knoten des CRF entsprechen in dieser Arbeit den 3D-Punktwolkensegmenten. Die Definition der Kanten des CRF erfolgt auf Grundlage von Nachbarschaften im Objektraum, weil diese nicht wie Nachbarschaftsbeziehungen im Bild durch die Abbildung der 3D-Szene auf die Bildebene beeinflusst sind. Zu diesem Zweck werden für jeden Punkt seine k -nächsten Nachbarn in der Punktwolke bestimmt. Für die k -nächste Nachbarn-Suche wird ein kd -Baum als Datenstruktur verwendet. Für Punkte an den Segmentgrenzen enthält die Menge dieser Nachbarn Punkte aus zwei oder mehr Segmenten. Die Häufigkeit, mit der Punkte aus einem Segment j in der Nachbarschaft von Punkten eines anderen Segmentes i anzutreffen sind, dient als Grundlage für die Definition der Kanten: liegt diese Häufigkeit über einem festzulegenden Schwellwert, wird zwischen den Knoten i und j eine Kante erzeugt. Da die den Knoten entsprechenden Segmente sowohl im Bild als auch in der Punktwolke vorliegen, kann zur Berechnung der Knotenmerkmalsvektoren \mathbf{f}_i , welche die Grundlage für das Assoziationspotenzial jedes Knoten i bilden, sowohl 2D-Bild- als auch 3D-Information genutzt werden. Das gleiche gilt für die Kantenmerkmalsvektoren \mathbf{m}_{ij} , welche im Interaktionspotenzial für jede Kante e_{ij} berücksichtigt werden und die aus den Merkmalen der jeweiligen, durch die Kante verbundenen Knoten bestimmt werden. Für jedes Segment wird zunächst im 3D-Raum aus den zugehörigen 3D-Punkten \mathbf{p}_n eine ausgleichende Ebene Ω der Form $\Omega: ax + by + cz + d = 0$ bestimmt, deren Parameter über eine Hauptkomponentenanalyse der Matrix \mathbf{M} der Summen der Quadrate und Produkte der schwerpunktreduzierten Segmentpunkte $\mathbf{p}'_n = \mathbf{p}_n - \bar{\mathbf{p}} = [x'_n, y'_n, z'_n]^T$ ermittelt werden, mit $\bar{\mathbf{p}}$ als jeweiligem Segmentenschwerpunkt und

$$\mathbf{M} = \begin{bmatrix} \Sigma x_n'^2 & \Sigma x_n' y_n' & \Sigma x_n' z_n' \\ \cdot & \Sigma y_n'^2 & \Sigma y_n' z_n' \\ \cdot & \cdot & \Sigma z_n'^2 \end{bmatrix}.$$

Der Normalvektor $\mathbf{n} = [a \ b \ c]^T$ entspricht dabei dem Eigenvektor zum kleinsten Eigenwert von \mathbf{M} und d ergibt sich über $d = -\mathbf{n}^T \bar{\mathbf{p}}$. Aus den 3D-Punkten und den so ermittelten Ebenen kann zusätzlich zu der reinen Bildinformation eine Reihe von Merkmalen abgeleitet werden. Alle hier verwendeten radiometrischen und geometrischen Knoten- und Kantenmerkmale sind in Tab. 1 zusammengefasst.

Tab. 1: Übersicht über verwendete Knoten- und Kantenmerkmale

Knotenmerkmale f_i		Kantenmerkmale m_{ij}
2D-Merkmale	3D-Merkmale	
<ul style="list-style-type: none"> - RGB-Werte - Intensität, Sättigung, Farbton - Haralick Features (Energie, Kontrast, Homogenität, Entropie) (HARALICK ET AL., 1973) 	<ul style="list-style-type: none"> - Eigenwerte und Eigenvektoren von \mathbf{M}, Normalvektor der 3D-Ebenen - Segmentschwerpunkt - Anisotropie und Planarität (CHEHATA ET AL., 2009) - Höhe über Grund 	<ul style="list-style-type: none"> - Winkel α_{ij} zw. Normalvektoren d. Seg. - Höhenunterschied Δh_{ij} und Distanz der Schwerpunkte der Segmente - Differenz der mittleren Intensität der Segmente

2.3.2 Definition der Potenziale

Für die Berechnung der Potenziale können beliebige diskriminative Klassifikatoren mit probabilistischem Output verwendet werden (KUMAR & HEBERT 2006). Das Assoziationspotenzial $A(\mathbf{x}, C_i)$ beschreibt die Wahrscheinlichkeit $P(C_i|\mathbf{x}) = P(C_i|f_i)$ für das Auftreten einer Klasse bei gegebenem Knotenmerkmalsvektor f_i , wobei als Klassifikator *Random Forests* (RF) (BREIMAN 2001) eingesetzt werden. Zur Berechnung des Interaktionspotenzials werden in dieser Arbeit zwei Modelle genutzt und evaluiert: Ein *binäres* Modell, welches eine Wahrscheinlichkeit dafür beschreibt, dass benachbarte Segmente i und j bei gegebenem Merkmalsvektor $\mathbf{g}_{ij} = [f_i, f_j, \mathbf{m}_{ij}]^T$ zur gleichen Klasse gehören, wobei auch hier ein RF als Klassifikator verwendet wird, sowie ein als *segmentbasiert* bezeichnetes Modell, in welchem die Stärke des Glättungseffekts mit

$$I(\mathbf{x}, C_i, C_j) = \begin{cases} w \cdot (\lambda_1 + \lambda_2 \cdot \cos \alpha_{ij} + \lambda_3 \cdot \exp(-\frac{(\Delta h_{ij})^2}{2\sigma^2})), & \text{für } C_i = C_j \\ 0, & \text{sonst} \end{cases} \quad (2)$$

von dem Winkel zwischen den Normalvektoren α_{ij} und dem Höhenunterschied Δh_{ij} der benachbarten Segmente i und j abhängt. Die Parameter $\lambda_{1,2,3} \geq 0$ mit $\lambda_1 + \lambda_2 + \lambda_3 = 1$ beschreiben die Gewichte der einzelnen Komponenten, wobei λ_1 den Grad der datenunabhängigen Glättung bestimmt. w ist ein Gewichtungsfaktor des gesamten Interaktionspotenzials relativ zum Assoziationspotenzial, während σ^2 die Varianz der Höhendifferenz bezeichnet. Die Glättung durch dieses Modell ist somit dann am stärksten, wenn benachbarte Segmente ähnliche Normalvektoren haben, sodass $\cos \alpha_{ij} \approx 1$ und der Höhenunterschied zwischen den Segmentschwerpunkten mit $\Delta h_{ij} \approx 0$ gering ist.

2.3.3 Training und Inferenz

Im Zusammenhang mit graphischen Modellen bedeutet *Inferenz* die Bestimmung der optimalen Labelkonfiguration \mathbf{C} durch die Maximierung von $P(\mathbf{C}|\mathbf{x})$ aus Gleichung (1). Für Graphen mit Zyklen ist die Inferenz im Mehrklassenfall nicht exakt lösbar. Daher werden approximative Lösungen verwendet, wobei in dieser Arbeit die *Loopy Belief Propagation* (LBP) (FREY & MACKAY 1998) zum Einsatz kommt. Das überwachte Training der beiden Potenziale, für welches manuell gelabelte Datensätze benötigt werden, erfolgt getrennt voneinander. Bei der Verwendung von RF als Klassifikatoren müssen die im RF verwendeten Entscheidungsbäume anhand der Trainingsdaten angelernt werden. Für das segmentbasierte Modell wird während des

Trainings der Parameter σ bestimmt, die Festlegung der Glättungskoeffizienten sowie des Gewichtsfaktors erfolgte empirisch.

3 Ergebnisse und Diskussion

Die Evaluierung der entwickelten Klassifikationsmethodik erfolgte anhand von 44 Stereobildern aus drei Stereosequenzen der „KITTI Vision Benchmark Suite“ (GEIGER et al. 2012) durch Kreuzvalidierung. Wie bei SENGUPTA et al. (2013) werden die zehn Objektklassen *Himmel*, *Gebäude*, *Vegetation*, *Straße*, *Gehweg*, *Zaun*, *Fahrzeug*, *Schild*, *Stange/Mast* und *Person* unterschieden, wobei die von SENGUPTA et al. (2013) übernommenen Referenzdaten durch manuelle Erfassung zusätzlicher Daten erweitert wurden. Die Gesamtgenauigkeit der Klassifikation mittels der unterschiedlichen Modelle ist in Tab. 2 für verschiedene Segmentgrößen dargestellt. Die Tabelle zeigt ebenso wie Abb. 3, dass, während sich für das Modell ohne Kontext bereits gute Ergebnisse ergeben, die Berücksichtigung von Kontext mittels des *binären* Modells und insbesondere bei Verwendung des *segmentbasierten* Modells zu einer Verbesserung sowie einer Glättung des Klassifikationsergebnisses führt.

Tab. 2: Punktweise ermittelte Gesamtgenauigkeiten unter Verwendung verschiedener Modelle für das Interaktionspotenzial.

Modell	100 [Px/Seg]	250 [Px/Seg]	450 [Px/Seg]
Ohne Kontext	87.9%	89.0%	89.1%
Binär	88.1%	89.6%	90.0%
segmentbasiert	91.4%	90.9%	91.0%

↓ Genauigkeit

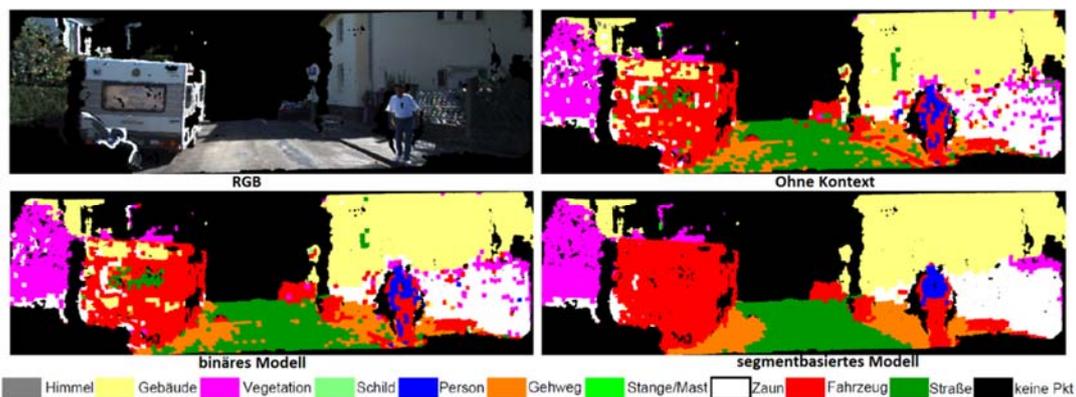


Abb. 3: Klassifikationsergebnis eines Stereomodells unter Verwendung unterschiedlicher Modelle für die Interaktionspotenziale (100 Px/Seg).

Betrachtet man die Klassifikationsqualität einzelner Klassen für das segmentbasierte Modell in Tab. 3, mit $Qualität[\%] = 100 \cdot \#TP / (\#TP + \#FN + \#FP)$, wobei

- $\#TP$: Anzahl der richtig zugeordneten Primitive,
- $\#FN$: Anzahl der irrtümlich einer anderen Klasse zugeordneten Primitive und
- $\#FP$: Anzahl der irrtümlich dieser Klasse zugeordneten Primitive sind,

sowie das anschauliche Beispiel der Klassifikation einer Person in Abb. 4, so fällt auf, dass Klassen mit kleinen Objekten (*Schilder, Stange/Mast, Personen*) problematisch für die Klassifikation sind und mit zunehmender Segmentgröße schlechter erkannt werden. Ein Grund hierfür ist die bereits in Abb. 2 gezeigte Problematik der Segmentierung, mit zunehmender Segmentgröße die Objektgrenzen kleinerer Objekte zu erhalten, wodurch in die Merkmalsberechnung dieser Segmente Pixel und Punkte unterschiedlicher Objekte und/oder unterschiedlicher Objektklassen einbezogen werden, was zu instabileren Segmentmerkmalen führt. Ebenso können durch die Nichteinhaltung von Objektgrenzen durch die Segmente Nachbarschaftsbeziehungen zwischen Objekten im Vordergrund und Objekten im Hintergrund entstehen, welche durch die hier verwendete Definition der Nachbarschaften im 3D-Objektraum anstatt in der Bildebene eigentlich vermieden werden sollten. Auch die im Vergleich zu anderen Objekten relativ geringe Menge an Trainingsdaten für diese Klassen kann sich negativ auf das Ergebnis auswirken.

Tab. 3: Punktweise ermittelte Qualität einzelner Klassen bei verschiedenen Segmentgrößen

Segmentgröße	Straße	Gebäude	Vegetation	Fahrzeug	Schild	Stange/Mast	Person
100 [Px/Seg]	90.1%	86.1%	87.1%	84.7%	28.3%	21.2%	46.6%
450 [Px/Seg]	90.2%	85.8%	86.5%	82.0%	8.8%	6.1%	26.1%

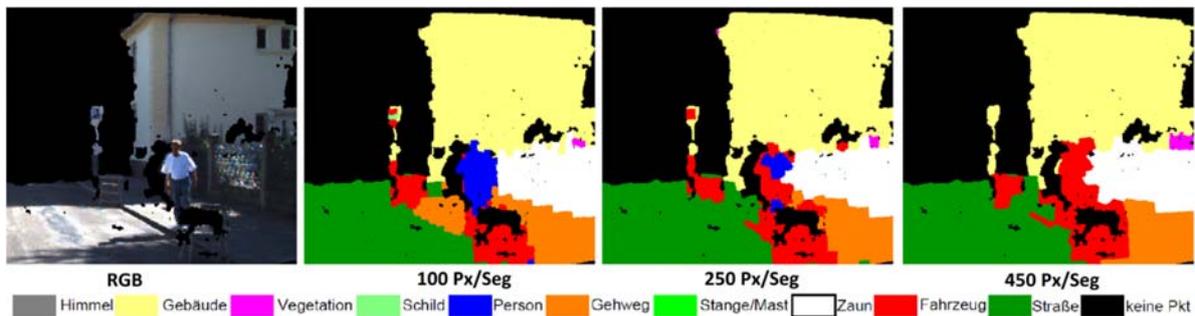


Abb. 4: Ausschnitt der Klassifikation einer Szene unter Verwendung unterschiedlicher Segmentgrößen

4 Fazit und Ausblick

Zusammenfassend lässt sich festhalten, dass sich mittels der in dieser Arbeit entwickelten Klassifikationsmethode für Stereobilder bereits ohne die Verwendung von Kontextinformation sehr gute Gesamtgenauigkeiten erreichen lassen, welche durch die Berücksichtigung von Kontext mithilfe verschiedener Modelle für das Interaktionspotenzial, insbesondere unter der Verwendung des *segmentbasierten* Modells, weiter verbessert werden können. Ferner wird damit eine deutlich glattere Labelkonfiguration erreicht. Die segmentweise Bearbeitung führt zu einer Reduktion der Rechenkomplexität, allerdings bei zu groß gewählten Segmenten auch zu Problemen für die Klassifikation kleinerer Objekte.

Weiterführende Untersuchungen der in dieser Arbeit entwickelten Methodik sollten insbesondere dahingehen, die bisher separat und unabhängig voneinander klassifizierte Stereobilder einer Sequenz gemeinsam innerhalb eines multitemporalen CRF-Modells, unter Hinzunahme einer zusätzlichen zeitlichen Komponente, zu klassifizieren.

5 Literaturverzeichnis

- ACHANTA, R., SHAJI, A., SMITH, K., LUCCHI, A., FUA, P. & SUSSTRUNK, S., 2012: SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34** (11), 2274-2282.
- ANGUELOV, D., TASKARF, B., CHATALBASHEV, V., KOLLER, D., GUPTA, D., HEITZ, G. & NG, A., 2005: Discriminative learning of Markov random fields for segmentation of 3D scan data. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* **2**, 169-176.
- BISHOP, C., 2006: *Pattern Recognition and Machine Learning*. New York, USA: Springer.
- BREIMAN, L., 2001: Random Forests. *Machine Learning* **45** (1), 5-32.
- CHEHATA, N., GUO, L. & MALLET, C., 2009: Airborne Lidar Feature Selection for urban Classification using Random Forests. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **38** (3/W8), 207-212.
- FREY, B. & MACKAY, D., 1998: A revolution: Belief propagation in graphs with cycles. *Advances in Neural Information Processing Systems* **10**, 479-485.
- GEIGER, A., LENZ, P. & URTASUN, R., 2012: Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite. *IEEE Conference on Computer Vision and Pattern Recognition*, 3354-3361.
- GEIGER, A., ROSER, M. & URTASUN, R., 2011: Efficient Large-Scale Stereo Matching. *Computer Vision - ACCV 2010*, **6492**, Springer Berlin Heidelberg, 25-38.
- HARALICK, R.M., SHANMUGAM, K. & DINSTEN, I., 1973: Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics* **3** (6), 610-621.
- KUMAR, S. & HEBERT, M., 2006: Discriminative Random Fields. *International Journal of Computer Vision* **68** (2), 179-201.
- LIM, E.H. & SUTER, D., 2009: 3D terrestrial LIDAR classifications with super-voxels and multi-scale Conditional Random Fields. *Computer-Aided Design 2009* **41** (10), 701-710.
- MATTI, E.K. & NEBIKER, S., 2014: Geometry and Colour Based Classification of Urban Point Cloud Scenes Using a Supervised Self-Organizing Map. *Photogrammetrie Fernerkundung Geoinformation*, (3), 161-173.
- REICH, M., UNGER, J., ROTTENSTEINER, F. & HEIPKE, C., 2013: On-Line compatible Orientation of a Micro-UAV based on Image Triplets. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2** (3), 37-42.
- SENGUPTA, S., GREVERSON, E., SHAHROKNI, A. & TORR, P., 2013: Urban 3D semantic modelling using stereo vision. *IEEE International Conference on Robotics and Automation*, 580-585.

Simulating an Autonomously Operating Low-Cost Static Terrestrial LiDAR for Multitemporal Maize Crop Height Measurements

SOPHIE CROMMELINCK¹

Zusammenfassung: Im Zuge der Optimierung landwirtschaftlicher Prozesse in ökonomischer und ökologischer Sicht (Präzisionslandwirtschaft), sind Informationen zu pflanzenspezifischen Parametern (z.B. Wachstumsdynamik) unverzichtbar. Solche Informationen können mit einem Laserscanner erfasst werden. Allerdings ist deren Anwendung momentan durch hohe Beschaffungskosten und mangelnde standardisierte Methoden der automatischen Datenauswertung nur eingeschränkt sinnvoll. In der im Folgenden dargestellten Masterarbeit wird daher das Potenzial eines kostengünstigen, automatisch funktionierenden Laserscanners im Bereich Präzisionslandwirtschaft untersucht. Dazu werden Daten, die die multitemporale Wachstumsentwicklung von Maispflanzen beschreiben, aufgenommen und ausgewertet. Die Punktdichte der aufgenommenen Daten wird künstlich reduziert um einen kostengünstigen Laserscanner zu simulieren. Anhand dieser Daten werden rasterbasierte Höhenmodelle des Pflanzenbestandes berechnet. Die Ergebnisse zeigen eine hohe Genauigkeit im Vergleich zu Referenzpflanzen für Modelle, die basierend auf unreduzierten Daten berechnet sind (mittlere Abweichung = 0.02 m, Std. Abw. = 0.15 m, RMSE = 0.16 m). Wenn die Daten auf 2% ihres ursprünglichen Umfangs reduziert werden, um ein kostengünstiges Scan-System zu simulieren, verschlechtert sich die Genauigkeit (mittlere Abweichung = 0.12 m, Std. Abw. = 0.19 m, RMSE = 0.22 m). Die Ergebnisse zeigen, dass die Anwendung eines kostengünstigen Scan-Systems mit einer Auflösung von bis zu 8 mrad (i.e., Punktabstand von 80 mm in 10 m Entfernung) in der Präzisionslandwirtschaft realisierbar ist. Die Masterarbeit liefert (a) umfassende Hinweise zum Scanner Aufbau sowie (b) eine Methode, die multitemporale Höhenmodelle des Pflanzenbestandes automatisch erstellt und auswertet.

1 Introduction

Light detection and ranging (LiDAR) has emerged as a powerful active remote sensing method for direct measurement of 3D plant structure in precision agriculture (LEE et al. 2010). The importance of 3D geodata in agriculture lies in its usability for plant modeling and plant-specific parameterization to monitor and improve crop management strategies (ROSELL & SANZ 2012). Applications include fertilization, irrigation management, yield estimation and optimization of harvesting processes (ZHANG et al. 2002).

Complexities caused by the outdoor agricultural environment, such as variable natural lighting, occlusion and obscuration of plant features by foliage from neighboring plants make an automatic observation of in-field variations challenging (MCCARTHY et al. 2010). Terrestrial laser scanning (TLS) provides a tool for generating a unique and comprehensive quantitative description of 3D plant and crop structure (PAULUS et al. 2014). TLS can be used to capture crop height, which is an important parameter for the local assessment and monitoring of crop types such as maize and wheat

¹ Heidelberg University, Institute of Geography, GIScience Research Group, Berliner Straße 48, D-69120 Heidelberg, E-Mail: crommelinck@stud.uni-heidelberg.de

(EHLERT et al. 2010; HÖFLE 2014). Crop height responds to nutrients, water, and temperature and can be used as an indicator for external conditions of soil, weather, irrigation and fertilization (OMASA et al. 2007).

A TLS system can acquire multitemporal data such as crop height development when operating constantly and autonomously (ATLS). Such scanning systems can be emulated as a simplified low-cost version by mounting a laser rangefinder on a pan tilt unit, which automatically carries out horizontal and vertical scanning by rotating within predefined angles (CULVENOR et al. 2014; EITEL et al. 2013). Such autonomously operating scanning systems capture multitemporal vegetation development, which adds time information to the 3D point cloud.

This study aims to identify specific requirements for a robust sensing system with automatically manageable data to improve precision agriculture applications. This is done by simulating a low-cost stationary TLS, which collects multitemporal crop height data. A workflow for processing and assessing ATLS data is developed. The workflow extracts the crop height development of maize by deriving crop height models (CHM), which represent the spatial crop growth pattern on a field level (HOFFMEISTER et al. 2009). The specific outcomes of this study are (a) clear requirements for a static ATLS system commendable for its successful use in precision agriculture and (b) a point cloud processing workflow for multitemporal monitoring of maize's crop height.

2 Materials and Methods

2.1 Study Area and Measurement Set-Up

A maize field in Heidelberg, Germany (49.43430° N, 8.65466° E, WGS84) served as the study area. The maize data were captured eight times over a period of 75 days starting on 28 May 2015 and ending on 10 August 2015. The measurements were conducted 40, 66, 75, 95, 100, 107 and 114 days after seeding, which took place on 18 April 2015.

For scanning maize crops a time-of-flight scanner VZ-400 (Riegl, Horn, Austria) with full-waveform echo detection was used. The system applies a near-infrared (1550 nm) laser beam with a beam divergence of 0.3 mrad and a range accuracy of 5 mm (one sigma) at 100 m. The system was installed on a water barrel to reduce occlusion effects, which was located at a field's edge.



Fig. 1: Scanner set up on stable barrel for elevated scanning at a height of 3.6 m, located at one edge of the monitored maize field.

This single scan position was used to simulate and determine the most cost-effective solution, since installation of each further scan position would raise the cost of the entire system. The scanner was registered to a global coordinate system in each scan campaign, since it could not be precluded that the simulated ATLS system would not change in position between measurements, as it had to be reinstalled for each measurement. To accurately register the ATLS position, a real time kinematic (RTK) GNSS base with a coordinate quality, on average, of 0.01 m was installed before each measurement on a known unchanging position. A RTK GNSS rover (Leica GS15) then captured the accurate position of the scanner.

Furthermore, two positions at the side of the field were chosen to set up tie point reflectors. They were fastened to metal poles that were secured about 30 cm into the ground. This allowed the re-establishment of the reflectors at the same position for each campaign. To verify this, the accurate position of the reflectors was measured with an RTK GNSS rover. In addition, the height and position of 140 plants along different transects in the field were measured manually with a measuring tape and a RTK GNSS rover each time. An overview of the measurement set-up is provided in Fig. 2.

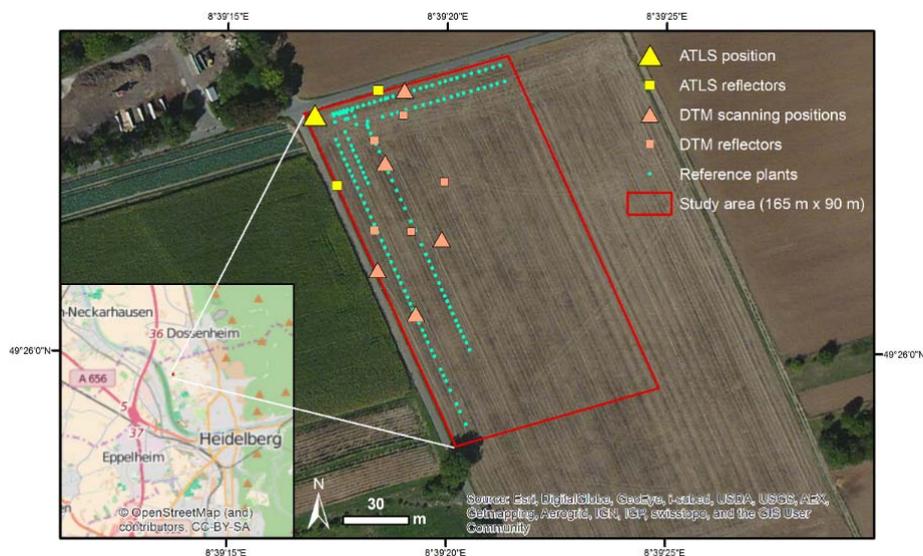


Fig. 2: Overview of measurement set-up located in Heidelberg (Germany) for simulation of autonomously operating terrestrial laser scanner (ATLS). Since the first campaign (t_1) aimed to gain data for a complete high-resolution digital terrain model (DTM), multiple scanning positions and reflectors were used. For the subsequent set-up of the simulated ATLS system ($t_2 - t_8$), only one scan position and two reflectors were used.

2.2 Generation of Crop Height Models

A crop surface model (CSM) represents the upper boundary of a crop field as an elevation value for deriving spatial crop growth patterns on a field level (HOFFMEISTER et al. 2009). A crop height model (CHM) represents the crop height by subtracting the DTM value from the maximum CSM elevation value per cell. The workflow of CHM generation (Fig. 3) aims to extract multitemporal CHMs with a cell size of 0.25 m x 0.25 m based on TLS measurements of different point densities. From the TLS data of the first campaign (t_1), a DTM for the study area was extracted. TLS data for each campaign ($t_1 - t_8$) were captured and registered to a common coordinate system. Then, point density was systematically decimated by keeping every n^{th} emitted laser beam according to

its timestamp with increasing step width n , i.e., $n = 0$ to $n = 50$ at increments of 2 (HÄMMERLE & HÖFLE 2014). This allows the simulation of ATLS systems with different scanning capabilities in terms of point density. From these datasets, CHMs were derived by calculating the difference between CSM per campaign ($t_1 - t_8$) and the DTM from the first campaign (t_1). The workflow is based on modules of the software Orientation and Processing of Airborne Laser Scanning data (OPALS) (PFEIFER et al. 2014), which are merged in an automatically executable Python script (*Link to the script that contains crop height modelling and assessing steps*).

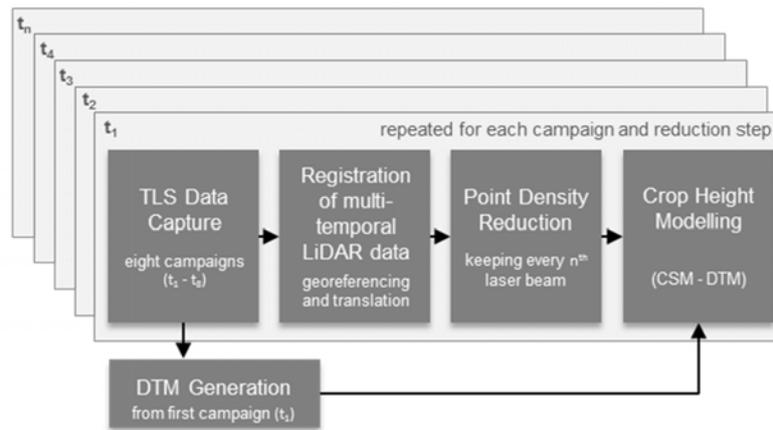


Fig. 3: Workflow of crop height model (CHM) generation from TLS data capture to CHMs calculation from reduced point density data sets.

2.3 Accuracy Assessment of Crop Height Models

As for CHM accuracy assessment, the manually measured height at 140 unchanging reference plants was compared to corresponding CHM cell values. If a reference plant was within a CHM cell, the difference between CHM height and the manually measured height of the reference plant was captured as Δh . Factors influencing Δh were statistically investigated by calculating Pearson's coefficient of determination (R^2) between Δh and distance to the scanner as well as between Δh and point density. This information helps to decide on parameters such as distance to scanner and point density when constructing an ATLS set-up for crop height measurements.

3 Results and Discussion

3.1 Generation of Crop Height Models

After registering multitemporal TLS data from eight point clouds ($t_1 - t_8$), the difference of coordinate pairs in all scans varies 0.001 m (x), -0.001 m (y) and 0.024 m (z), on average. This accuracy is similarly low as obtained by Tilly et al. (0.06 m and 0.01 m) for aligning multitemporal scans of rice crops (TILLY et al. 2015) and 0.04 m for scans of barley crops (TILLY et al. 2015). Whilst the first reduction step ($n = 2$) leads to a reduction of about 50% (7,388,193 points) of the original point cloud, the final reduction step ($n = 50$) leads to a reduction of 98% (333,773 points). All point clouds resulting from step widths $n > 10$ result in a reduction between 2% and 10%. The reference scans, which were captured with a lower scan resolution, show a similar point count development. The artificial reduction of point density leads to similar results considering the point count of reference scans in (HÄMMERLE & HÖFLE 2014).

3.2 Accuracy Assessment of Crop Height Models

Calculated heights in CHMs are close to those manually measured regardless of point density reduction until approximately 95 days after seeding and a crop height of 2.3 m (Fig. 4). The unreduced dataset ($n = 0$) leads to the smallest difference between manually measured reference plants and CHM based crop heights.

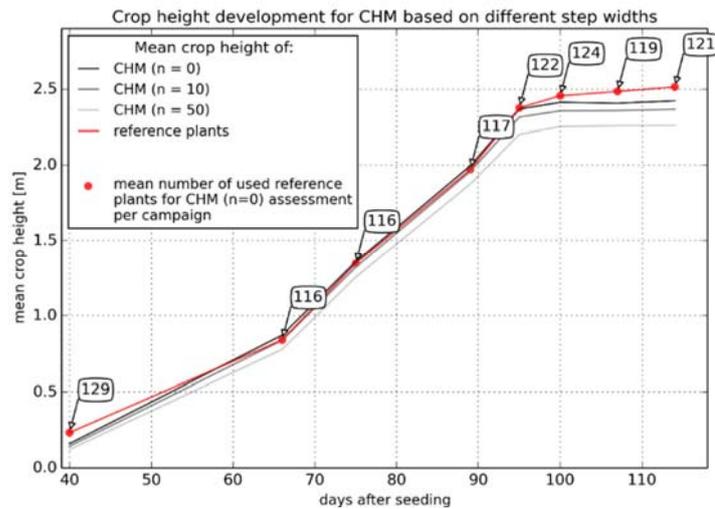


Fig. 4: Crop height development based on CHMs for step width $n = 0$ (i.e., no artificial reduction), $n = 10$ (10% of data remaining) and $n = 50$ (2% of data remaining) and on manually measured reference plants. These are only used for calculation of Δh if a CHM cell contains a value at their location. Connecting lines do not indicate linear crop growth.

The development of Δh for the CHMs shows that a reduction of point density leads to a decrease in accuracy, with Δh augmenting with increasing crop height (Fig. 5).

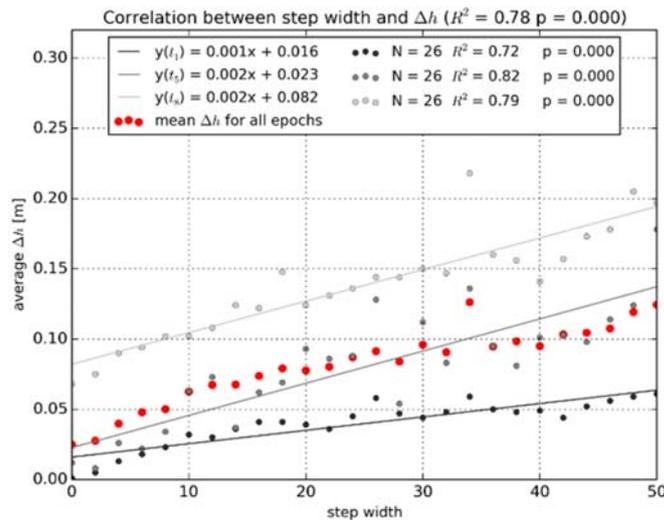


Fig. 5: Accuracy of CHM at 40 (t_1), 95 (t_5) and 114 (t_3) days after seeding based on point clouds reduced with different step widths ($n = 0$, i.e., no artificial reduction to $n = 50$, i.e., 2% of data remaining). The points belonging to a specific campaign, show Δh averaged per step width and are summarized in a line of regression in the same color (y). Each campaign encompasses 26 Δh values, i.e., one for each step width from $n = 0$ to $n = 50$ at increments of 2. Pearson's coefficient of determination (R^2) and p value (p) are additionally added.

Overall, the unreduced data ($n = 0$) leads to high accuracy values for Δh (mean deviation = 0.02 m, std. deviation = 0.15 m, RMSE = 0.16 m). This accuracy decreases when reducing the point cloud with step width $n = 50$, i.e., 2% of data kept (mean deviation = 0.12 m, std. deviation = 0.19 m, RMSE = 0.22 m). These results for $n = 50$ can be expected for a maximum distance of 100 m (150 m for $n = 0$, i.e., no artificial reduction) between scanner and crops.

These findings permit some general statements about the minimum ATLS point density needed for the monitoring of maize and structurally similar crops. As stated in (HÄMMERLE & HÖFLE 2014), the scan resolution applied needs to be set in accordance to crop stand features such as crop type, fertilization stage, growth stage and required accuracy. Additionally, it has to be mentioned that the artificial reduction of point density simulates a low-cost ATLS system only to a certain extent. High-resolution scans used as a basis for reducing point density have a high penetration rate. Subsequently, conducted point clouds of reduced point density encompass points of laser beams that would not have been captured in the case of a real low-resolution scan as applied in (EITEL et al. 2013; GLENNIE & LICHTI 2010).

4 Conclusion and Outlook

This study shows that LiDAR technology in general and multitemporal 3D geodata in particular are feasible for application in precision agriculture. No reduction of the original points cloud ($n = 0$) results in an accuracy of 0.02 m, a reduction of 90% ($n = 10$) results in an accuracy of 0.06 m and a reduction of 98% ($n = 50$) results in an accuracy of 0.12 m. The last reduction step corresponds to a scan resolution of 8 mrad, i.e., point spacing of 80 mm at 10 m distance.

It can be concluded that models and applications that use RMSE of Δh are robust against reducing scanning resolution. CHM's accuracy measured via Δh decreases for later measurements when crops are higher than 2.3 m, because occlusion effects appear more often. The scanner height should be chosen according to maximum crop height in order to optimize the accuracy. The workflow developed, which automatically derives and assesses CHMs is reusable in a multitude of further studies and application scenarios. Analogously, it could be refined to process multitemporal 3D geodata for monitoring environmental changes (e.g. landslides, snow cover or glaciers).

Further research could investigate the transferability of the approach by using a real low-cost laser scanning system as in (EITEL et al. 2013) based on the specific outcomes of this study. The transfer of the results to additional crops (e.g. wheat, rye or rice), crop parameters (e.g. plant volume or nutrient content) and different measurement set-ups (e.g. mobile vs. static) could be investigated to complement and enhance knowledge about the potential of a TLS in PA.

5 References

- CULVENOR, D.S., NEWNHAM, G.J., MELLOR, A., SIMS, N.C. & HAYWOOD, A., 2014: Automated in-situ laser scanner for monitoring forest Leaf Area Index. *Sensors* **14** (8), 14994-15008.
- EHLERT, D., HEISIG, M. & ADAMEK, R., 2010: Suitability of a laser rangefinder to characterize winter wheat. *Precision Agriculture* **11** (6), 650-663.
- EITEL, J. U. H., VIERLING, L. A. & MAGNEY, T. S., 2013: A lightweight, low cost autonomously operating terrestrial laser scanner for quantifying and monitoring ecosystem structural dynamics. *Agricultural and Forest Meteorology* **180**, 86-96.

- GLENNIE, C. & LICHTI, D. D., 2010: Static calibration and analysis of the Velodyne HDL-64E S2 for high accuracy mobile scanning. *Remote Sensing* **2** (6), 1610-1624.
- HÄMMERLE, M. & HÖFLE, B., 2014: Effects of reduced terrestrial LiDAR point density on high-resolution grain crop surface models in precision agriculture. *Sensors* **14** (12), 24212-24230.
- HOFFMEISTER, D., BOLTEN, A., CURDT, C., WALDHOFF, G. & BARETH, G., 2009: High-resolution Crop Surface Models (CSM) and Crop Volume Models (CVM) on field level by terrestrial laser scanning. 6th International Symposium on Digital Earth (78400E-78400E-78406). International Society for Optics and Photonics.
- HÖFLE, B., 2014: Radiometric correction of terrestrial LiDAR point cloud data for individual maize plant detection. *Geoscience and Remote Sensing Letters, IEEE* **11** (1), 94-98.
- LEE, W.S., ALCHANATIS, V., YANG, C., HIRAFUJI, M., MOSHOU, D. & LI, C., 2010: Sensing technologies for precision specialty crop production. *Computers and Electronics in Agriculture* **74** (1), 2-33.
- OMASA, K., HOSOI, F. & KONISHI, A., 2007: 3D lidar imaging for detecting and understanding plant responses and canopy structure. *Journal of Experimental Botany* **58** (4), 881-898.
- PAULUS, S., SCHUMANN, H., KUHLMANN, H. & LÉON, J., 2014: High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants. *Biosystems Engineering* **121**, 1-11.
- 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.
- ROSELL, J.R. & SANZ, R., 2012: A review of methods and applications of the geometric characterization of tree crops in agricultural activities. *Computers and Electronics in Agriculture* **81**, 124-141.
- TILLY, N., AASEN, H. & BARETH, G., 2015: Fusion of Plant Height and Vegetation Indices for the Estimation of Barley Biomass. *Remote Sensing* **7** (9), 11449-11480.
- TILLY, N., HOFFMEISTER, D., CAO, Q., LENZ-WIEDEMANN, V., MIAO, Y. & BARETH, G., 2015: Transferability of Models for Estimating Paddy Rice Biomass from Spatial Plant Height Data. *Agriculture* **5** (3), 538-560.
- ZHANG, N., WANG, M. & WANG, N., 2002: Precision agriculture - a worldwide overview. *Computers and Electronics in Agriculture* **36** (2), 113-132.

Cloud basiertes Rendering von komplexen 3D-Stadtmodellen

MARKUS JUNG¹

Zusammenfassung: Dreidimensionale Stadtmodelle zeichnen sich meist durch eine sehr große Komplexität aus. Insbesondere aufgrund ihrer Texturen summierte sich ein solches Modell zu sehr großen Datenmengen. Für die Darstellung solch komplexer und datenintensiven Stadtmodellen werden entsprechend viel Ladezeit, sowie Speicherkapazität und Rechenpower benötigt. Im Rahmen meiner Masterthesis wurde ein Verfahren entwickelt um komplexe Modelle effizient über einen Webbrowser auf beliebigen Endgeräten darzustellen. Dazu wurde ein Prototyp implementiert. Das Rendering mittels Raytracing Verfahren, welches nicht auf dem Client, sondern auf dem Server ausgeführt wird, ermöglicht die Generierung sehr realitätsgetreuer Bilder. Zudem lässt sich dieses Verfahren in einer Cloudumgebung sehr gut parallelisieren und ist somit sehr gut skalierbar. Beim Image Based Provisioning wird nicht nur ein Farbbild generiert, sondern weitere Layer mit Informationen zur Tiefe, Normalen oder Objektzugehörigkeit jedes Pixels erstellt. Das Modell wird dabei in mehreren Bildpyramidenstufen gerendert und in Kacheln fixierter Größe unterteilt. Dadurch müssen auf den Client nur Bilder übertragen werden, welche konstant und damit kalkulierbar in ihrer Datenmenge sind. Mit diesem Verfahren lässt sich die Komplexität des 3D Modells von der Darstellung entkoppeln und der Betrachter kann trotzdem im Modell navigieren (Zoom, Pan) und Koordinatenmessungen oder Attributabfragen durchführen.

1 Einleitung und Motivation

In den letzten Jahren wurden verschiedene webbasierte 3D-Geodienste, insbesondere auch virtuelle Globen entwickelt (CHRISTEN et al. 2012; COZZI & BAGNELL 2013). Mit diesen Geobrowsern konnte unter anderem gezeigt werden, dass es möglich ist im Webbrowser 3D-Stadtmodelle darzustellen. Aufgrund der teilweise sehr hohen Detailtreue weisen solche Modelle eine sehr große Datenmenge auf. So können 3D Modelle aus mehreren Tausenden oder Millionen Objekten und Texturen bestehen. Trotz Ansätzen mit Level of Detail oder dem dynamischen Nachladen räumlich aufgeteilter Segmente wie Kacheln, können komplexere 3D-Stadtmodelle zu erheblich längeren Ladezeiten und schlechterer Performance in der Darstellung führen (CHRISTEN et al. 2014; KLIMKE et al. 2014).

Auch Hardwareressourcen auf Smartphone und Tablets sind begrenzt nutzbar für eine performante interaktive 3D-Visualisierung. Für einen möglichst vielfältigen Nutzen der Geodaten ist es jedoch gerade wichtig, dass die Verteilung der Daten über verbreitete Hardware, Software und Netzwerke erfolgen kann. Die meisten Personen verfügen heute über mehrere mobile Geräte wie Notebooks, Tablets oder Smartphones und tragen diese ständig mit sich. Die Verteilung über das World Wide Web bietet einen zeit- und ortsunabhängigen flexiblen Zugang zu den Daten. Ein weiterer Vorteil der Darstellung über den Webbrowser ist, dass ein einfacher Zugang von den meisten Endgeräten ohne Installation weiterer Software ermöglicht wird.

¹ Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: markus.jung@thurweb.ch

Aus diesen Gründen sind vermehrt neue, alternative Methoden zur 3D-Visualisierung im Webbrowser gefragt.

2 Technologien

In dieser Masterthesis wurde für die Darstellung von 3D Stadtmodellen über den Webbrowser ein bildbasierter Ansatz verwendet. Die Bilder werden in der Cloud prozessiert und zwischengespeichert. Für die Darstellung im Browser wurden die gespeicherten Bilder geladen und dargestellt. Das cloudbasierte Rendering erfolgt dabei mit dem Raytracing Verfahren. Verschiedene Interaktionen werden durch Image Based Provisioning ermöglicht.

2.1 Image Based Provisioning

SAITO & TAKAHASHI (1990) präsentierten das Prinzip von geometrischen Puffern (G-Buffer). Diese bestehen aus einem Set von Ansichten, welche jede davon eine geometrische Eigenschaft beinhaltet. Es wird teilweise auch von Multi Layer Images gesprochen (Abb. 1). Diese aus mehreren Layern bestehenden Bilder werden für das Image Based Provisioning verwendet. Die 3D Modelle werden dazu auf einem Server zu Multi Layer Images gerendert. Diese Bilder werden auf den Client gesendet und können dadurch auch auf Geräten mit geringerer Hardwareleistung visualisiert werden. Obwohl es sich um Bilder handelt, ist dank den verschiedenen Layern trotzdem eine Interaktion möglich. Die Tiefenkarte ermöglicht ein Messen im dreidimensionalen Raum. Anhand des Layer zur Objektidentifikation lassen sich Objekte anwählen und weitere thematische Daten dazu abrufen. Das Open Geospatial Consortium (OGC) definiert zurzeit dazu den 3D Portrayal Service, ein Standard zur Verbreitung von 3D-Geodaten über das Internet (OPEN GEOSPATIAL CONSORTIUM 2015).

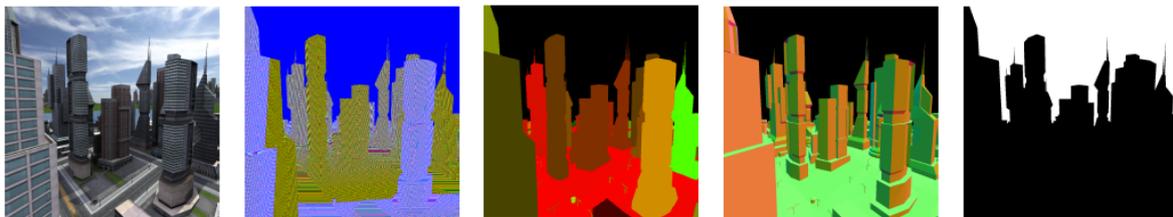


Abb. 1: Multi Layer Images: (v.l.n.r) RGB, Tiefenkarte, Objektidentifikation, Normalen, Maske (HAGEDORN 2009)

2.2 Raytracing

Beim Rendering wird aus einer dreidimensionalen Szene mit definiertem Blickpunkt ein Bild gerendert. Raytracing ist ein solches Renderingverfahren das von APPEL (1968) vorgestellt und von WHITTED (1980) verbessert wurde. Dabei wird für einen definierten Kamerastandpunkt für jeden Pixel mittels einer Strahlverfolgung die Farbe berechnet. Entlang des Strahles wird das erste sichtbare Objekt ermittelt und die Farbe berechnet. Für die Farbberechnung werden einerseits die Textur und Oberflächenbeschaffenheit des Objektes verwendet. Zudem werden weitere Strahlen vom Objekt ausgesendet um Beleuchtung, Schattierung, Reflexionen und Transparenz zu simulieren. Dadurch können sehr realistische Bilder aus Modellen generiert werden, weshalb Raytracing oft in Filmen (z.B. Ice Age, Robots, Ratatouille) verwendet wird (Abb. 2). Ein großer

Vorteil ist auch die Möglichkeit die Berechnung zu parallelisieren, wodurch sie die Berechnungsdauer verkürzen lässt.



Abb. 2: Bild aus dem Film Ratatouille (2007), berechnet mittels Raytracing

3 Rendering

Das Rendering beim erstellten Prototypen erfolgte nach dem RenderMan Standard mit der Software PRMan. Beides wurde von den Pixar Animation Studios entwickelt (PIXAR 2016). Weiter wurden einige Untersuchungen mit Aqsis vorgenommen, welches eine Open Source Alternative zu PRMan ist.

3.1 Datenaufbereitung

Als Ausgangsdaten für den Prototyp wurden mehrere 3D Modelle aus CityEngine verwendet. Diese wurden in das offene Dateiformat Wavefront OBJ exportiert, ein verbreitetes Format für die Speicherung und den Austausch von 3D Daten. Anschliessend wurden diese Daten in RenderMan spezifische Formate konvertiert, welche Geometrien und Texturen des Modells beinhalten. Zusätzlich wurden noch verschiedene Shader definiert. In den Shadern werden die beim Rendering benötigten Berechnungen der Farbwerte, Lichtverhältnisse, etc. programmiert.

3.2 Transformation

Für das Rendering der Stadtmodelle wurde eine Orthogonalprojektion verwendet. Dazu wurden vordefinierte Schrägansichten auf das Modell gerendert (Abb. 3). Für jedes Bild wurden die Kameraausrichtung und der Bildausschnitt berechnet. Die Berechnung beinhalteten das Parsing der Koordinaten und Ermittlung der Bounding Box. Daraus konnten dann Transformationsmatrix für die Kamerapositionierung und der Gesamtausschnitt des Modells berechnet werden. Um das Modell auch auf mobilen Geräten effizient darstellen und darin navigieren zu können, wurden diese mittels Kachelung und Bildpyramiden in mehrere tausend Einzelbilder unterteilt. Für jedes Einzelbild wurde eine RIB Datei gemäss RenderMan Standard generiert und anschliessen mit PRMan gerendert. Nebst den Farbwerten wurden für jedes Bild weitere Layer generiert (Tiefenkarte, Normalen, Objektidentifikation). Anhand dieser zusätzlichen Layer lassen sich auch in den Bildern Koordinatenmessungen durchführen, Beleuchtungen simulieren oder Attribute zu Objekten anzeigen.

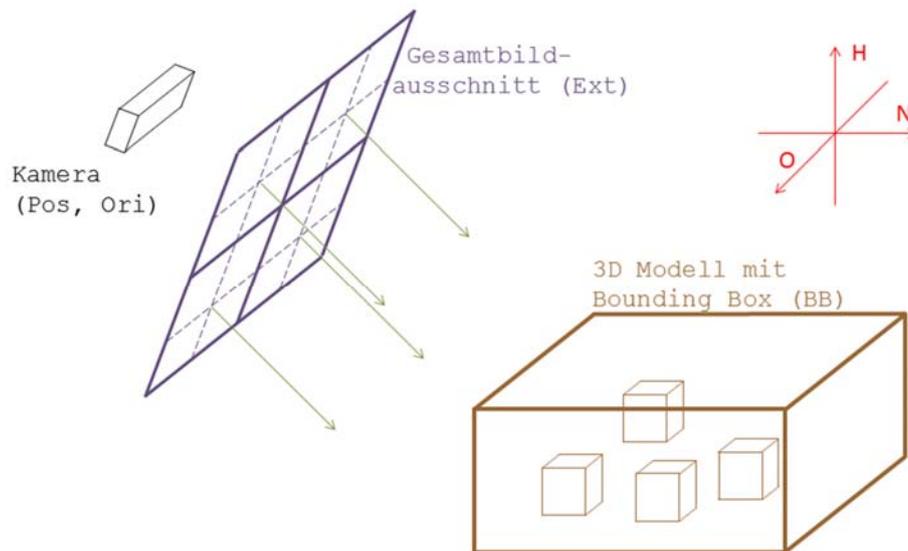


Abb. 3: Positionierung der Kamera und des Bildausschnitts zum 3D Modell

4 Webanwendung

Anhand des im letzten Kapitel vorgestellten serverseitigen Renderings lassen sich sehr komplexe Stadtmodelle über herkömmliche Webbrowser auf jedem beliebigen Gerät darstellen. Dies wurde mit einer einfachen Webanwendung aufgezeigt. Dazu wurde ein node.js Webserver aufgesetzt. Für die Darstellung des Modells wurde die Javascript Bibliothek Leaflet verwendet. Darin kann sich ein Nutzer verschiedene Ansichten auf ein Modell anzeigen lassen und darin navigieren (Zoom, Pan). Die angezeigten Kacheln werden dazu laufend vom Server geladen. Sofern vorhanden werden die Kacheln aus dem Zwischenspeicher (Cache) geladen oder ansonsten gerendert. Mittels zwei weiteren Service lassen sich im Webbrowser räumliche und attributive Eigenschaften zu den Objekten abrufen. Diese werden serverseitig aus den verschiedenen Layern ermittelt. Eine vereinfachte Version der Webanwendung (Anzeige von zwischengespeicherten Kacheln ohne Messfunktionalität und Attributanzeige) kann über mth2014.3dmaps.ch aufgerufen werden.

5 Cloud-Computing

Beim entwickelten Prototypen werden die Kacheln vorgerendert und für die Darstellung aus dem Cache geladen. Für das Rendering in Echtzeit werden leistungsfähige Server benötigt. Abb. 4 zeigt eine Cloudarchitektur, die mehrere Vorteile vereint. Das Rendering eines einzelnen Bildes kann auf mehreren Servern parallel prozessiert werden. Ein Webserver verarbeitet die verschiedenen Nutzeranfragen. Sofern vorhanden können die benötigten Daten aus dem Cache geladen werden. Ansonsten wird das Bild auf mehreren Servern gerendert und für künftige Anfragen im Cache abgelegt. Ein Monitoring überwacht die Antwortzeit und die Auslastung der Server fürs Rendering. Die einzelnen Server lassen sich in der Cloud nach Bedarf hinzu- oder abschalten, wodurch die Anwendung sehr kosteneffizient und skalierbar betrieben werden kann.

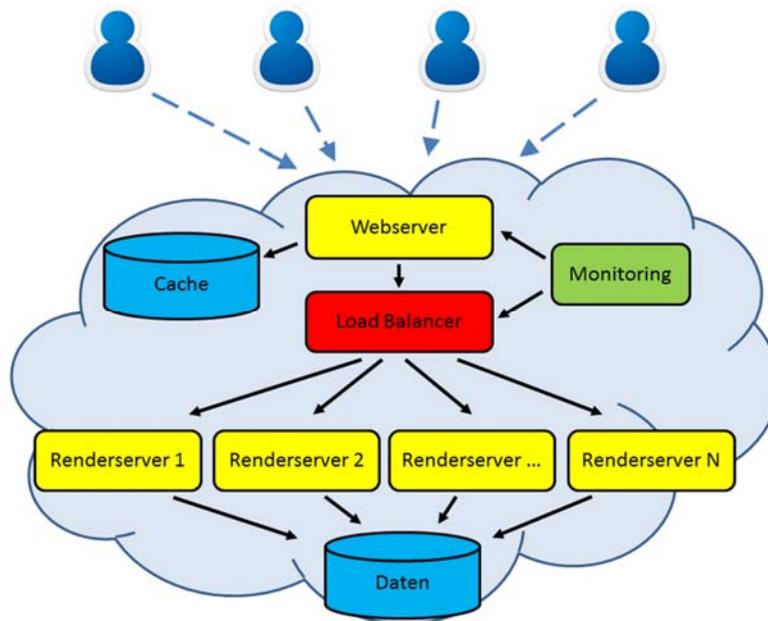


Abb. 4: Parallelisierung des Rendering in der Cloud

6 Resultate

Durch das serverseitige Rendern kann die Komplexität eines 3D Modells von seiner Darstellung entkoppelt werden. Somit lassen sich auch sehr grosse Stadtmodelle auf beliebigen Endgeräten ohne grosse Hardwareleistung darstellen. Nachfolgend zwei gerenderte Kacheln mit der Software PRMan (Abb. 5) und Aqsis (Abb. 6). Mit PRMan konnten sehr realitätsgetreue Bilder generiert werden. Grund für den auffälligen Qualitätsunterschied ist die bessere Berechnung der Lichtverhältnisse mittels Ambient Occlusion in PRMan.



Abb. 5: PRMan mit Ambient Occlusion



Abb. 6: Aqsis ohne Ambient Occlusion

Das Rendering von mehreren Layern (Farbwerte, Tiefenkarte, Objektzugehörigkeit) ermöglicht auch in den Bildern verschiedene Interaktionen und Abfragen. Abb. 7 und Abb. 8 zeigen Koordinatenmessungen sowie Attributabfragen in der Webanwendung.

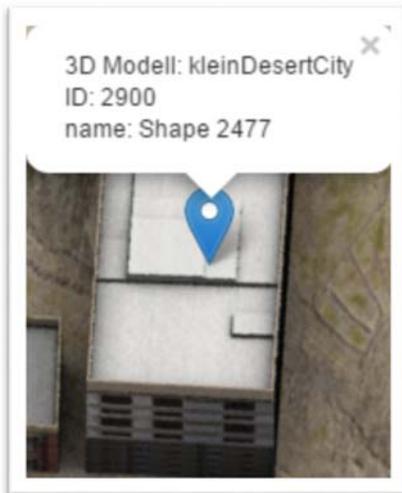


Abb. 7: Abfrage von Attributen

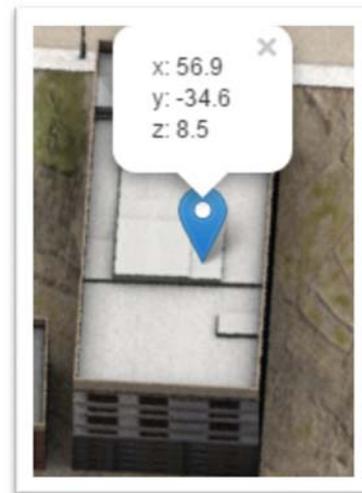


Abb. 8: Koordinatenmessung



Abb. 9: 3D-Stadtmodell Rotterdam mit Orthophoto: Highlight mittels Objektidentifikationslayer (CHRISTEN 2016)

7 Fazit und Ausblick

Mit dem entwickelten Verfahren und dem Prototyp konnte gezeigt werden, dass sich auch sehr komplexe 3D-Stadtmodelle effizient auf verschiedenen Endgeräten darstellen lassen. Image Based Provisioning und Raytracing ermöglichen die Komplexität eines Modells von der Übertragung und der Darstellung zu entkoppeln. In der prototypischen Webanwendung werden zwischengespeicherte vordefinierte Ansichten angezeigt. Optimierungen für das Rendering in

Realtime wurden mit der Cloudarchitektur vorgestellt. Diese Rasterdaten liessen sich im Client mit weiteren Raster oder Vektordaten erweitern. Dazu können verschiedenen Szenarien simuliert und dargestellt werden. Somit lassen sich diverse Anwendungen dreidimensional darstellen und verbreiten, beispielweise Verkehrssimulationen, Überschwemmungen, Stadtentwicklungen, Überbauungen und nur einige zu nennen.

Die Arbeit wurde inzwischen weiterentwickelt, um 3D-Stadtmodelle aus City-GML und Orthophotos zu unterstützen (CHRISTEN 2016), was in der Abb. 9 zu sehen ist. Geplant sind unter anderem auch die Visualisierung von Punktwolken, sowie Animationen bei denen Echtzeit-3D mit der Schrägansicht kombiniert werden kann.

8 Literaturverzeichnis

- APPEL, A., 1968: Some techniques for shading machine renderings of solids. AFIPS Conference Proc. **32**, 37-45
- 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 e.V., Band 23, E. Seyfert, E. Gülch, C. Heipke, J. Schiewe, M. Sester, (Hrsg.), 34. Wissenschaftlich-Technische Jahrestagung der DGPF, Beitrag 250.
- CHRISTEN, M., 2016: Visualisierung von 3D-Geodaten im (mobilen) Webbrowser. Geoinformationssysteme 2016, Kolbe, T.H., Bill, R. & Donaubaue, A. (Hrsg.), Wichmann, ISBN 978-3-87907-610-9.
- COZZI, P. & BAGNELL, D., 2013: A webgl globe rendering pipeline. GPU Pro 4: Advanced Rendering Techniques **4**, 39-48.
- HAGEDORN, B., 2009: Service-Based, Interactive Portrayal of 3D Geovirtual Environments. Proceedings of the 4th Ph.D. Retreat of the HPI Research School on Service-oriented Systems Engineering, Meinel, C., Plattner, H., Döllner, J., Weske, M., Polze, A., Hirschfeld, R., Naumann, F. & Giese, H. (Hrsg.), Potsdam, Universitätsverlag, 13:1-13:9.
- KLIMKE, J., HAGEDORN, B. & DÖLLNER, J., 2014: Scalable Multi-Platform Distribution of Spatial 3D Contents. International Journal of 3-D Information Modeling **3**, 3, 35-49.
- OPEN GEOSPATIAL CONSORTIUM (OGC), 2015: OGC seeks public comment on candidate 3D Portrayal Service Standard, OGC Press Release, <http://www.opengeospatial.org/pressroom/pressreleases/2165>, letzter Zugriff 25.02.2016.
- PIXAR, 2016: Renderman. <https://renderman.pixar.com/view/renderman>, letzter Zugriff 02.03.2016.
- SAITO, T. & TAKAHASHI, T., 1990: Comprehensible Rendering of 3-D Shapes. Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '90. New York, NY, USA: ACM, 197-206
- WHITTED, T., 1980: An Improved Illumination Model for Shaded Display. Commun. ACM, **23** (6), 343-349

Entwicklung einer großen Invardraht-Maßverkörperung zur Anwendung in der Industriephotogrammetrie

OLIVER KAHMEN¹ & THOMAS LUHMANN¹

Zusammenfassung: In diesem Beitrag wird eine Arbeit vorgestellt, die es als Ziel hatte, eine Maßverkörperung zu entwickeln, welche speziell bei Projekten großer Volumina höchste Genauigkeiten photogrammetrischer Vermessungen ermöglicht. Sie soll die Genauigkeit herkömmlicher Referenzmaßverkörperungen übertreffen und möglichst den Einsatz eines kostspieligen Lasertrackers zur Erzeugung eines Referenzpunktfeldes überflüssig machen. Die Entwicklung der 21 m langen Invardraht-Maßverkörperung unter Berücksichtigung physikalischer Materialeigenschaften bildet den Kern der Arbeit. Verschiedene Maßverkörperungen werden an einem praktischen Beispiel hinsichtlich ihrer Qualität untersucht, um einen Überblick des tatsächlichen Nutzens der neu entwickelten Maßverkörperung gegenüber herkömmlichen Systemen zu gewinnen.

1 Einleitung

Die industrielle Messtechnik bildet in vielen Sektoren der modernen Fertigung und Qualitätssicherung einen wichtigen Bestandteil. In interdisziplinärer Anwendung hat die optische 3D-Messtechnik höchste Genauigkeitsanforderungen zu erfüllen und bedarf daher individueller Lösungen auf hohem Niveau. Eine individuelle Lösung für großvolumige Objekte wurde in dieser Arbeit im Hause der sigma3D GmbH in Gescher erarbeitet und an der Jade Hochschule Oldenburg zum Abschluss des Studiengangs *Angewandte Geodäsie (B.Sc.)* vorgelegt.

In der Photogrammetrie stellt die Aufgabe der absoluten Skalierung einen wesentlichen Qualitätsfaktor dar. Die Skalierung maßstabsloser 3D-Punktwolken wird in der Praxis durch standardisierte CFK-Maßverkörperungen namhafter Hersteller realisiert. Diese werden mit sehr hoher Güte hergestellt und üblicherweise in Längen von 0,5-2 m vertrieben. Da sich Kalibrier- und Messgenauigkeiten einer Maßverkörperung fortpflanzend in dem Gesamtergebnis widerspiegeln, sind kleine Maßverkörperungen für Messobjekte großer Ausdehnung nur bedingt geeignet. Eine große Maßverkörperung kann daher im Anwendungsfall von Projekten großer Volumina Genauigkeitsvorteile mit sich bringen. Die Entwicklung einer Invardraht-Maßverkörperung soll neben einem Genauigkeitserfolg zusätzlich den Kostenfaktor gering halten, um mit herkömmlichen Referenzkörpern bzw. Instrumenten zur Erzeugung von Passpunktfeldern in Wettbewerb treten zu können.

2 Theoretische Genauigkeitsbetrachtung

Die mathematische Modellierung der Bündelblockausgleichung auf Grundlage der Kollinearitätsgleichungen stellt sich wie folgt dar:

¹ Jade Hochschule Oldenburg, Institut für Angewandte Photogrammetrie und Geoinformatik (IAPG), Ofener Straße 16/19, D-26121 Oldenburg, E-Mail: oliver.kahmen@posteo.de, luhmann@jade-hs.de

$$\begin{aligned}
 x' &= x'_0 + z' \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta x' \\
 y' &= y'_0 + z' \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta y'
 \end{aligned}
 \tag{1}$$

Durch die Einführung einer Bedingungsgleichung kann eine Streckenbeobachtung in der Ausgleichsrechnung berücksichtigt werden. Die Beobachtungsgleichung für Strecken im Raum lautet:

$$s = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2} \tag{2}$$

Liegen Referenzpunkte vor, können die sieben Parameter (3 Translationen, 3 Rotationen, 1 Maßstab) des Datums über die Informationen der Referenzpunkte gelöst werden. Liegen keine Passpunkte in einem übergeordneten System vor, wird die Ausgleichung nach dem Prinzip der freien Netzausgleichung gerechnet.

Die bekannte Streckeninformation fließt als beobachtete Größe mit einem ihrer Genauigkeit entsprechenden Gewicht in den Ausgleichungsprozess ein und legt den Maßstab als Datumparameter fest. Die Strecke wird als Maßverkörperung physisch definiert und wirkt wie folgt nach der folgenden schematischen Darstellung auf die Bündelblockausgleichung ein.

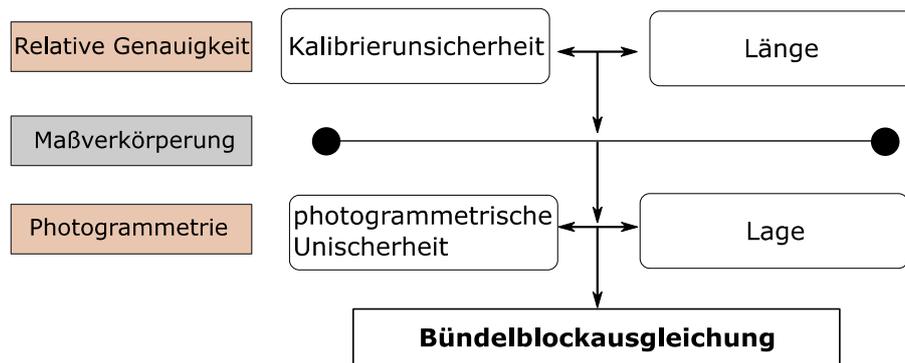


Abb. 1: Definition und Wirkung einer Maßverkörperung (KAHMEN 2015: 20)

Nach Abb. 1 wird die Maßverkörperung mit einer relativen Genauigkeit definiert, die von Kalibrierunsicherheit und Länge abhängt. Die Maßverkörperung wird dann mit einer von der Lage im Objektraum und photogrammetrischen Messunsicherheit abhängigen Güte in den Ausgleichungsprozess als übergeordnete Streckenbeobachtung eingeführt. Bei diesem Ansatz der Verwendung der Bündelblockausgleichung sind also vier Parameter bzgl. der Maßverkörperung von Bedeutung, die das Resultat beeinflussen:

1. Kalibrierunsicherheit
2. Länge
3. Photogrammetrische Unsicherheit
4. Lage im Objektraum

Die Unsicherheit der gemessenen Maßstabspunkte bei kurzen Maßverkörperungen wirkt sich auf extrapolierende Weise auf die Objektkoordinaten aus, wohingegen große Maßverkörperungen

keine extrapolierende Fehlerfortpflanzung aufweisen, da diese das gesamte Messvolumen im Idealfall umschließen.

Da die photogrammetrische Unsicherheit mit zunehmender Objektdimension und Aufnahmedistanz größer wird, wirkt sich die Kalibrierunsicherheit nur noch sekundär auf die Unsicherheit der Maßstabspunkte aus. Die Unsicherheit durch die Photogrammetrie selbst liefert also einen höheren Anteil am Unsicherheitsbudget als die Kalibrierunsicherheit der Maßverkörperung. Es kann daher folgende Hypothese aufgestellt werden:

„Große Maßverkörperungen können unter Umständen, auch mit vergleichsweise hoher Kalibrierunsicherheit, bessere Ergebnisse erzielen, als kleine Maßverkörperungen mit geringer Kalibrierunsicherheit.“

3 Entwicklung der Invardraht-Maßverkörperung

Seit vielen Jahrzehnten sind Drähte in der Geodäsie als zuverlässiges Mittel zur Bestimmung von Strecken und Streckenänderungen bekannt, wie es KAHMEN (2005: 148ff) für den Anwendungsfall der hochgenauen Streckenmessung zeigt. Nach dem Vorbild des Distometer ISETH (KERN 2016) wird eine Konstruktion entwickelt, die eine physische Länge von 21 m realisiert. Dazu wird ein Draht an einem Punkt fixiert und an einem Gegenpunkt linear verfahrbar gelagert. Bei dem Aufbau der Konstruktion muss der Draht zunächst am Fixpunkt an einem 360° drehbaren Kopf fixiert werden. Durch das Anbringen einer definierten Zugkraft am verfahrbaren Gegenpunkt wird dieser dann auf Spannung gebracht. Der Draht ist in eine Drahtkupplung geklemmt, welche wiederum an den drehbaren Kopf arretiert werden kann. Durch den drehbaren Zylinder kann sich das System selbstständig zur Zugkraftfrichtung ausrichten.

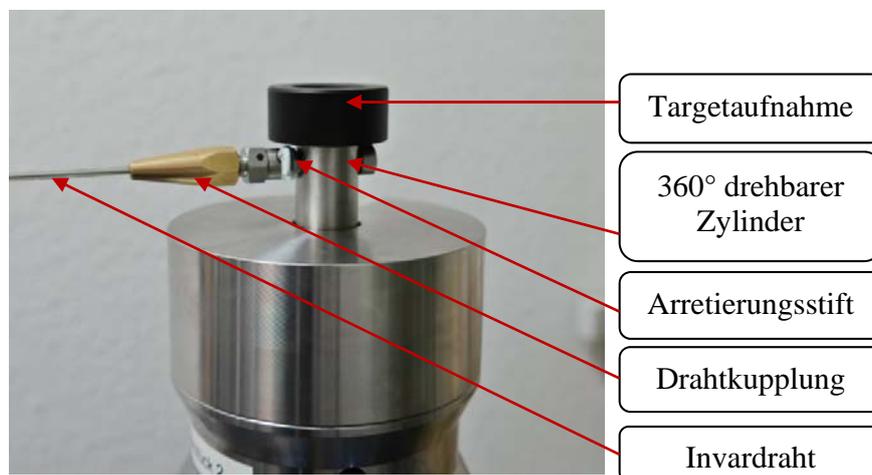


Abb. 2: Fixiertes Drahtende

Die Targetaufnahme ermöglicht die Messung des Punktes sowohl durch ein 1,5“ SMR (Spherically Mounted Reflector) mittels Lasertracker, wie auch durch photogrammetrische Methoden mit Hilfe einer 1,5“ Photogrammetrie-Halbkugel.

Die Lineareinheit auf der Gegenseite ermöglicht die Realisierung eines definierten Abstandes. Durch die angebrachte Zugkraft kann sich der 360° kugelgelagerte Linearzylinder vor und zurück bewegen, bis sich die Länge des Drahtes je nach herrschender Zugspannung stabil eingependelt

hat. Das Stativ, auf dem die Lineareinheit angebracht ist, muss dabei lediglich auf ± 5 cm genau, im Abstand von 21 m aufgestellt werden.

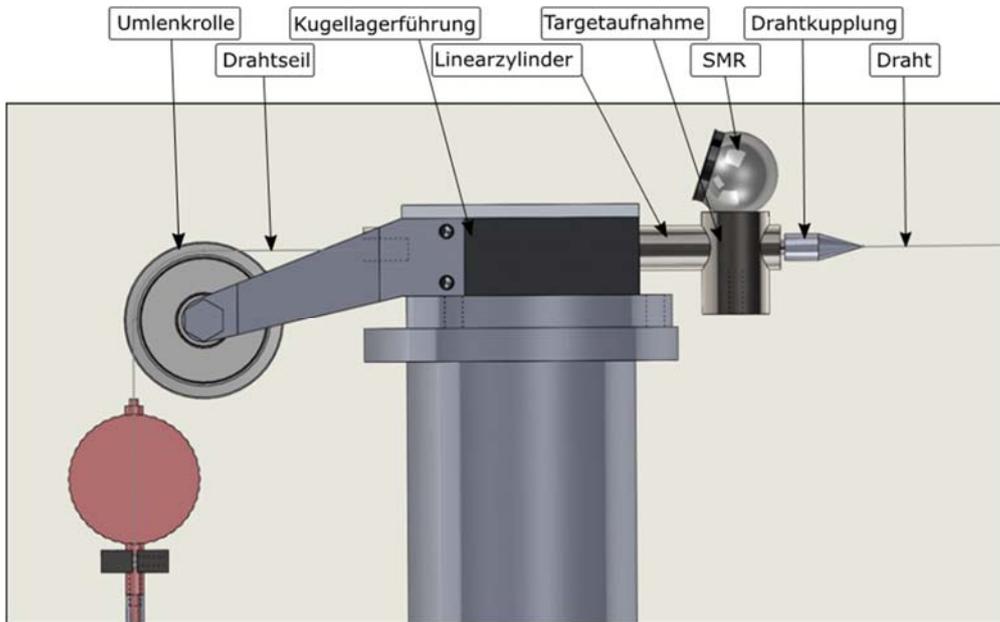


Abb. 3: Seitenansicht des linear verfahrenen Drahtendes

Eine Umlenkrolle, welche mit dem Linearzylinder verbunden ist, ermöglicht die Kraftanbringung über ein Spannsystem.

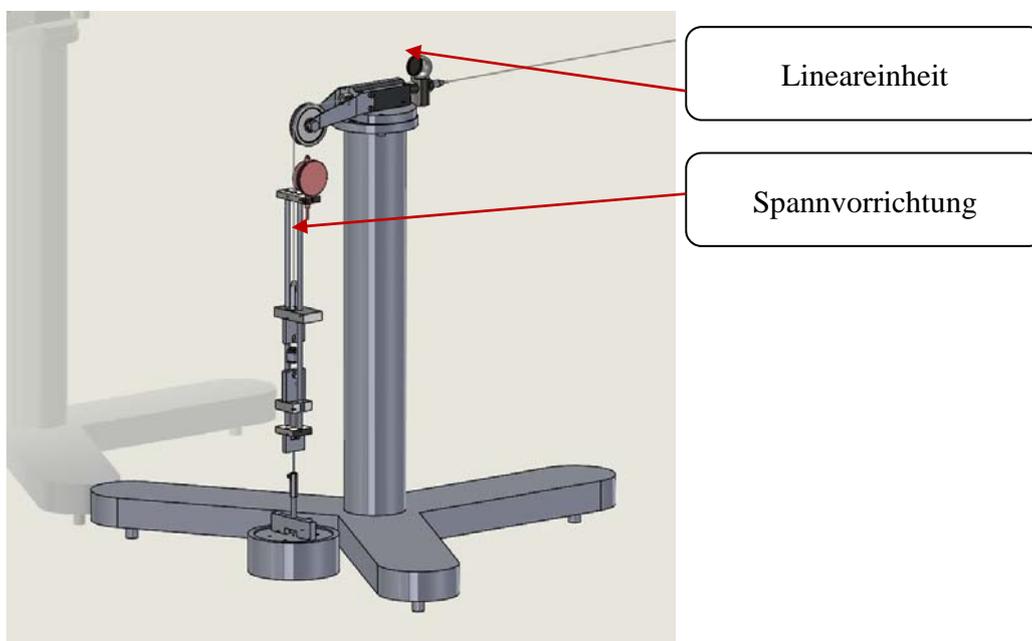


Abb. 4: 3D-Ansicht des Statives mit Lineareinheit und Spannvorrichtung

Die eigens konstruierte Spannvorrichtung ermöglicht die Erzeugung einer konstanten Zugkraft. Ein reibungsarmer PTFE-Block gleitet je nach Spannkraft auf einem Rundstahl auf und ab. Eine handelsübliche Messuhr greift dann den Stand des Blocks ab, woraufhin sich die geometrische Messgröße aufgrund bekannter Zugfederparameter in eine physikalische Zugkraft umrechnen lässt. Die Spannkraft, die größtmögliche Stabilität des gesamten Systems schafft, ohne dabei die Elastizitätsgrenze des Invardrahtes mit 1 mm Durchmesser zu überschreiten, beträgt 80 N. Die absolute Kraft ist dabei weniger interessant als ihre Reproduzierbarkeit.

Die absolute Drahtlänge kann im Labor mit hoher Güte durch interferometrische Lasertracker-Messungen (IFM) nach dem Abbe'schen Prinzip kalibriert werden. In der eigentlichen Anwendung wird der Draht dann vor Ort gespannt und kann aufgrund seiner bekannten kalibrierten Länge als Maßstab eingeführt werden.

4 Praktische Genauigkeitsuntersuchung

Zunächst wird die Invardraht-Maßverkörperung im Labor unter konstanten Temperaturbedingungen kalibriert, und damit in ihrer Länge definiert. Die Kalibrierung erfolgt mit einem übergeordnet genauen Messsystem. Das qualitativ höhere Messinstrument ist ein Lasertracker der Fa. Faro (ION). Mit einer Streckenmessgenauigkeit von $4\ \mu\text{m} + 0,8\ \mu\text{m}/\text{m}$ (Maximal Permissible Error) im IFM-Modus (FARO 2012), bestimmt der Lasertracker eine 21 m lange Strecke theoretisch mit einer Genauigkeit von 0,021 mm. Die Messung, bestehend aus 20 Einzelmessungen, brachte folgendes Ergebnis für die Maßverkörperung hervor:

Länge: 20834,374 mm Standardabweichung: 0,026 mm

Die erreichte Standardabweichung liegt mit 0,026 mm geringfügig über dem theoretischen Genauigkeitsniveau des Kalibriersystems, lässt aber dennoch auf einen stabilen und reproduzierbaren Aufbau der entwickelten Maßverkörperung schließen.

In einem Bildverband soll nun die Wirkung der großen Invardraht-Maßverkörperung hinsichtlich der absoluten Genauigkeit der Objektkoordinaten untersucht werden. Dazu wird ein Testfeld von 10x15 m geschaffen. Das Testfeld wird im Stil einer Luftbildphotogrammetrie hoch redundant photogrammetrisch vermessen. Verschiedene Maßstabsreferenzen sollen nun autonom in den Ausgleichungsprozess integriert werden und dadurch unterschiedliche praktisch denkbare Konfigurationen darstellen. Die verschiedenen Maßstabsreferenzen welche jeweils als Referenzinformation dienen sind die folgenden:

1. Invardraht-Maßverkörperung diagonal durch das Testfeld
2. 2 kurze CFK-Verkörperungen im Zentrum des Testfeldes
3. 5 kurze CFK-Verkörperungen räumlich verteilt im Testfeld
4. Passpunkt-Referenz aus Lasertracker-Messungen als 4x4 Gitter im Testfeld

Die Passpunkte aus den Lasertracker-Messungen dienen als Qualitätsmerkmal für die Auswertung. Die Längenmessabweichungen (LME) aller 120 Streckenkombinationen der 4x4 Passpunkte werden in jeder der 4 Variationen analysiert. Das Ergebnis zeigt deutlich, dass die große Maßverkörperung, genau wie die Variante der Passpunkt-Referenz bessere Ergebnisse liefert, als die herkömmlichen kurzen Maßverkörperungen aus CFK (vgl. Tab. 1).

Tab. 1: Statistik der LME verschiedener Maßstabskonfigurationen

PP Referenz			Invardraht		
Min	-0,117	mm	Min	-0,339	mm
Max	0,175	mm	Max	0,453	mm
Spanne	0,292	mm	Spanne	0,791	mm
Mittel	0,031	mm	Mittel	0,077	mm
Stabw.	0,064	mm	Stabw.	0,160	mm

2xCFK			5xCFK		
Min	-0,876	mm	Min	-0,832	mm
Max	0,882	mm	Max	0,827	mm
Spanne	1,757	mm	Spanne	1,659	mm
Mittel	-0,223	mm	Mittel	-0,195	mm
Stabw.	0,390	mm	Stabw.	0,363	mm

Trotz der vergleichsweise geringen Kalibrierqualität der Invardraht-Maßverkörperung ist die äußere Genauigkeit der Objektkoordinaten höher als die der hochgenau kalibrierten CFK-Maßstäbe. Die erzielten Ergebnisse stützen die Hypothese aus Abschnitt 2 und zeigen, dass die Investition in Länge signifikante Genauigkeitssteigerungen bei photogrammetrischen Projekten großer Volumina mit sich bringen kann und die Kalibriergüte hingegen an Relevanz verliert.

5 Fazit und Ausblick

Aus theoretischen Überlegungen entstand zunächst eine Hypothese, die die Idee der Entwicklung einer großen Maßverkörperung mit sich brachte. Die Entwicklung einer großen 21 m langen Invardraht-Maßverkörperung kann als gelungen betrachtet werden. Auch die Validierung hinsichtlich der physischen Stabilität des Systems ist durch die Kalibriermessung belegt. Der Effekt und statistische Größen wurden in einem ersten Versuch untersucht und unterstützen die vorangestellte Hypothese. Dennoch sind dahingehend weitere Untersuchungen anzustellen, um die Ergebnisse weiterführend zu evaluieren und variierende Versuchsaufbauten zu betrachten.

Die Arbeit zeigt, dass es nicht trivial ist, die Maßstabsreferenz bzgl. den Parametern Länge, Lage, Qualität und Anzahl festzulegen. Für große Messobjekte kann eine Invardraht-Maßverkörperung als eine gute Alternative zur Referenzpunkt-Erstellung mittels Lasertracker dienen, ohne dabei einen hohen Genauigkeitsverlust zu verursachen. Insbesondere in wirtschaftlicher Hinsicht könnte eine günstige Draht-Konstruktion den kostspieligen Lasertracker in einigen Anwendungen ersetzen.

Es sind Folgeuntersuchungen anzustellen, die den Effekt variierender Maßverkörperungsparameter im Bildverband validieren. Hier sind Simulationen wie z.B. Monte-Carlo-Simulationen denkbar, um die Statistik frei von praktisch nicht kontrollierbaren Einflüssen zu berechnen.

Das Institut für Angewandte Photogrammetrie und Geoinformatik (IAPG), Oldenburg stellt weiterführende Untersuchungen zu dem hier vorgestellten Thema an. Erste Ergebnisse aktueller Forschungsarbeiten können aus REZNICEK et al. (2016: 286ff) entnommen werden.

6 Danksagung

Ein besonderer Dank geht an das Unternehmen sigma3D GmbH, welches durch kompetente Unterstützung wesentlich zum Gelingen dieser Arbeit beigetragen hat. Insbesondere Herrn Frank Mönsters danken die Autoren für die tolle Betreuung seitens des Betriebs. Dem Photogrammetrie-Team des IAPG danken die Autoren für das Engagement und die gute Zusammenarbeit seitens der Jade Hochschule.

7 Literaturverzeichnis

- KAHMEN, H., 2006: Angewandte Geodäsie. Vermessungskunde. 20., völlig neu bearbeitete Auflage, Walter de Gruyter (De Gruyter Lehrbuch), Berlin.
- KAHMEN, O., 2015: Entwicklung einer großen Invardraht-Maßverkörperung zur Anwendung in der Industriephotogrammetrie. Bachelorarbeit (unveröffentlicht), Jade Hochschule Oldenburg.
- KERN, 2016: Distometer ISETH: www.kern-aarau.ch/fileadmin/user_upload/Aldo/Prospekte/Distometer_ISETH.pdf, letzter Zugriff 05.02.2016.
- FARO, 2012: Technisches Datenblatt FARO ION. www.faro.com, letzter Zugriff 05.02.2016.
- REZNICEK, J., EKKEL, T., HAESTADT, H., LUHMANN, T., KAHMEN, O. & JEPPIING, C., 2016: Zum Einfluss von Maßstäben in photogrammetrischen Projekten großer Volumina. Photogrammetrie - Laserscanning - Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2016, Luhmann, T. & Schumacher, C., (Hrsg.), Wichmann, 286-295.

Entwicklung und Anwendung einer Methode zur Erstellung von Wechselbilanzen der Flächennutzung auf Grundlage des ATKIS Basis-DLM

MARTIN SCHORCHT¹

Zusammenfassung: Die Umsetzung einer nachhaltigen Flächenpolitik erfordert ein Monitoring, welches nicht nur den jeweils aktuellen Stand der Flächennutzung ausweist, sondern auch den Wandel der Nutzungsarten beschreiben kann. Dieses ist bei laufenden Veränderungen der flächennutzungsbeschreibenden Modelle und laufenden Geometriekorrekturen eine große Herausforderung. Die jährliche statistische Flächenerhebung nach Art der tatsächlichen Nutzung (Grundlage: ALB/ALKIS) weist bis zur Ebene der Gemeinden die Flächenanteile nach Nutzungsarten aus. Aussagen zur Vor- bzw. Nachnutzung sind momentan nicht möglich. Besonders die Bilanzierung dieses Flächennutzungswandels ist zum Prozessverständnis von Nöten. Ziel dieser Arbeit ist es, eine Methode zu erarbeiten, mit der Wechselbilanzen auf Grundlage des ATKIS Basis-DLM erstellt werden können und bei der insbesondere geometrische Verschiebungen bzw. andere Veränderungen, die keine reale Flächennutzungsänderung darstellen, beachtet werden.

1 Einleitung

In Deutschland ist in vielen Regionen ein Flächennutzungswandel festzustellen, welcher von einer Zunahme von Siedlungs- und Verkehrsflächen geprägt ist (UBA 2014). Aus Sicht einer nachhaltigen Raumplanung ist diese Entwicklung negativ, da der Boden ein knappes Schutzgut ist und kein Bevölkerungszuwachs verzeichnet wird, der eine Flächenneuanspruchnahme rechtfertigen könnte. Für diese Flächenneuanspruchnahme ist die Entwicklung der Siedlungs- und Verkehrsfläche (SuV) ein wichtiger Indikator. Die erste Erhebung erfolgte 1979 und seit 1992 existieren gesamtdeutsche Daten (BBSR 2015), bei denen eine stetige Zunahme zu erkennen ist. Derzeitig beträgt der SuV-Zuwachs täglich 73 Hektar (UBA 2014). Um dieser Entwicklung entgegenzusteuern, verfolgt die Bundesregierung im Rahmen der nationalen Nachhaltigkeitsstrategie das Ziel, bis zum Jahr 2020 die Neuanspruchnahme von Flächen für Siedlung und Verkehr auf 30 Hektar pro Tag zu reduzieren. Zur Erreichung dieser Vorgaben ist die genaue Beobachtung, Quantifizierung und das Monitoring der Flächennutzung nötig (MEINEL 2013).

2 Datengrundlage

Grundlage des im Folgenden vorgestellten Verfahrens zur Bilanzierung des Flächennutzungswandels ist das ATKIS Basis-DLM. Für ganz Deutschland ist das ATKIS Basis-DLM der aktuellste und genaueste topographische flächendeckend verfügbare Datensatz (RÖBER et al. 2009). Mit diesem liegt ein amtlicher, deutschlandweiter, weitestgehend homogener und in

¹ Technische Universität Dresden, Fakultät Umweltwissenschaften: Fachrichtung Geowissenschaften, Institut für Kartographie, Helmholtzstr. 10, D-01062 Dresden, E-Mail: m.schorcht@ioer.de

seinem Aufbau genau dokumentierter Geobasisdatensatz vor, welcher die wichtigsten Voraussetzungen für ein Flächennutzungsmonitoring hinsichtlich der Datenkonsistenz und Aktualität erfüllt (KRÜGER 2011). Darin sind topographische Erscheinungsformen und Sachverhalte objektstrukturiert und attribuiert im Vektorformat (Punkte, Linien und Polygone) modelliert. Grundlegende Eigenschaften zur Bildung aller Objekte sind in der GeoInfoDok 6.0 (ADV 2008a) geregelt, welche das von der AdV und den Ländern entwickelte AAA-Anwendungsschema beschreibt. Das Basis-DLM wird jeweils von den einzelnen Bundesländern nach diesen einheitlichen Grundsätzen erarbeitet und bereitgestellt, wodurch diese Daten für räumliche Vergleiche in ganz Deutschland geeignet sind. Sämtliche Definitionen und Deklarationen zum Inhalt der Daten sind im Objektartenkatalog (OK) des Basis-DLM festgelegt (ADV 2008b). Eine der Hauptobjektkategorien ist die „Tatsächliche Nutzung“ (TN), welche die Nutzungsarten der Erdoberfläche flächendeckend und überlagerungsfrei beschreibt. Aus diesen Flächennutzungsdaten ist am Leibniz-Institut für ökologische Raumentwicklung in Dresden eine für Nutzungsartenanalyse übersichtlichere Nomenklatur abgewandelt worden.

Die in ATKIS als Linien modellierten Objekte des Verkehrs- und Gewässernetzes werden im Laufe eines Preprocessings anhand von Objektbreitenangaben gepuffert, wodurch Flächenobjekte gebildet werden. Die dadurch entstehenden Überlagerungen zu benachbarten bereits vorhandenen Flächenobjekten werden entfernt, damit wiederum eine überlagerungsfreie Abdeckung erreicht wird (MEINEL & KRÜGER 2014). Diese vorprozessierten Datensätze sind die Datengrundlage, welche für diese Arbeit verwendet worden ist.

Im IÖR-Flächenschema (Abb. 1) wird zwischen 35 einzelnen Nutzungsarten und 13 Oberklassen unterschieden. Diese Nomenklatur ist sowohl für das alte als auch das neue ATKIS-Modell (GeoInfoDok 6.0) gültig, sodass sich Flächennutzungsdaten unabhängig vom Migrationszeitpunkt vergleichen lassen. Die Migration in das neue ATKIS-Modell erfolgte in den Bundesländern zu unterschiedlichen Zeitpunkten zwischen 2008 und 2013 (ADV 2014). Dieser Zeitraum ist von besonderem Interesse, um neben dem Flächenwandel selbst auch den Einfluss der ATKIS-Migration zu untersuchen.

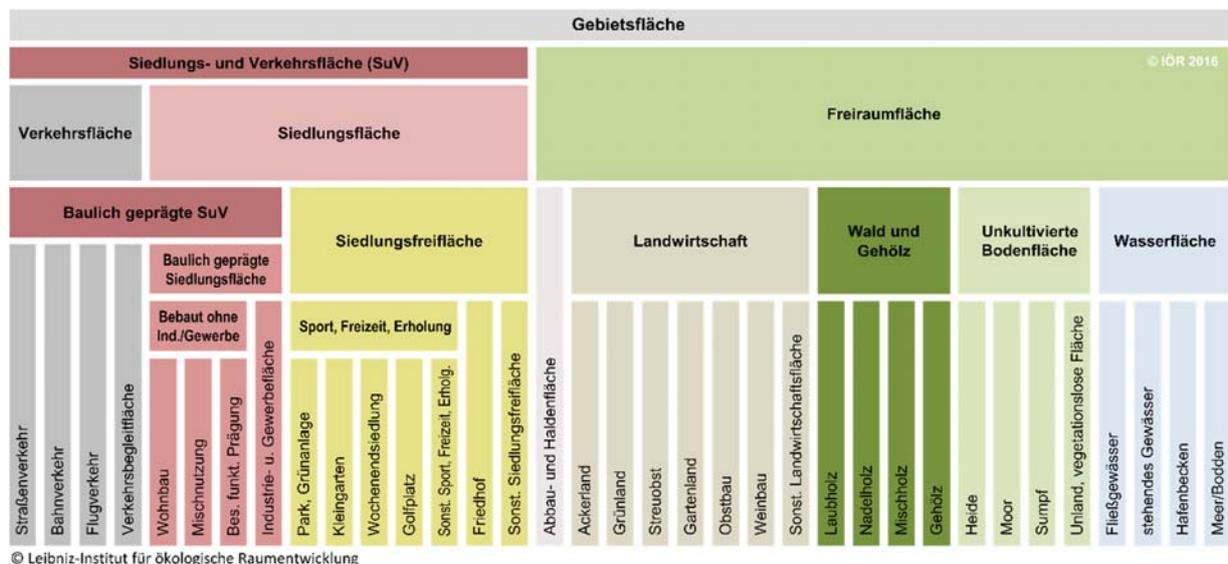


Abb. 1: Nomenklatur des IÖR Monitor (IÖR-MONITOR 2014)

3 Methode zur Ermittlung von Flächennutzungsänderungen

Das Verfahren zur Bilanzierung von Flächennutzungsänderungen erkennt im ersten Schritt durch Verschneidung der Datensätze verschiedener Zeitpunkte die Änderungen von Flächenobjekten. Diese ermittelten Änderungen werden anschließend bezüglich geringer geometrischer Verschiebungen der Polygonränder bereinigt, sodass diese geringen Verschiebungen sich nicht mehr auf die Bilanzierung auswirken. In einem weiteren Schritt werden migrationsbedingte Objektneuordnungen, denen keine reale Nutzungsänderung zugrunde liegt, berücksichtigt.

3.1 Verschneiden der Vektordaten

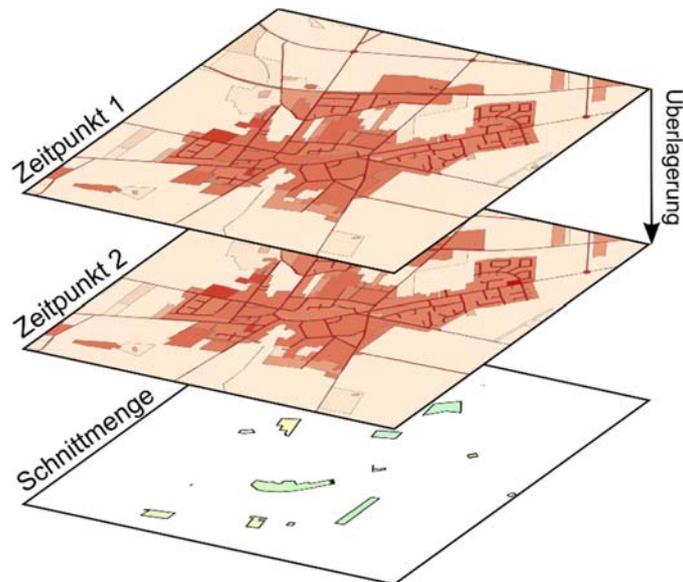


Abb. 2: Flächennutzungsänderungen zwischen zwei Zeitschnitten mittels Schnittmengenbildung (Quelle: eigene Darstellung)

Die Polygone der Datensätze zweier Zeitschnitte werden übereinandergelegt und die sich ändernden Schnittmengen erfasst (Abb. 2). Der Flächenzugang einer Nutzungsart führt notwendigerweise zu einem Flächenabgang einer anderen Nutzungsart. Dementsprechend bildet sich die Nettobilanz einer jeden Nutzungsart aus der Summe des ermittelten Flächenzugangs und -abgangs. Ist der Flächenzugang gleich des -abgangs, führt dies zu einer ausgewogenen Nettobilanz (= 0).

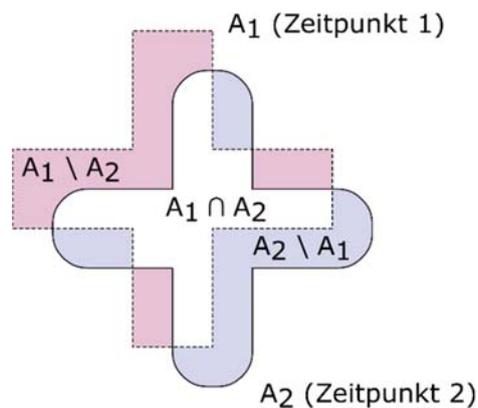


Abb. 3: Überlagerung eines sich ändernden Objektes; $A_1 \setminus A_2$: abgehende Fläche (A1 ohne A2); $A_1 \cap A_2$: sich nicht ändernde Fläche (A1 gleich A2); $A_2 \setminus A_1$: hinzukommende Fläche (A2 ohne A1) (abgewandelte Abbildung nach HAGEDORN & VELTKAMP (2001))

In Abb. 3 wird eine Überlagerung eines sich ändernden Objektes derselben Nutzungsart zweier unterschiedlicher Zeitpunkte dargestellt. Dabei entstehen sich überlagernde Bereiche, welche für eine unveränderte Nutzungsart stehen ($A_1 \cap A_2$), sowie sich nicht überlagernde Bereiche, welche für einen Rückgang ($A_1 \setminus A_2$), bzw. für einen Zugang ($A_2 \setminus A_1$) stehen. Die sich ändernden Objektbereiche gehen entweder in eine andere Nutzungsart (B) über ($\cong A_1 \cap B_2$), oder kommen aus einer anderen Nutzungsart her ($\cong B_1 \cap A_2$), welche der Vor- und Nachnutzung von A entsprechen. Die Differenz aus hinzukommenden und rückläufigen Flächen entspricht der Nettobilanz einer Nutzungsartenänderung von A:

$$\text{Netto}(A) = \text{area}(A_2 \setminus A_1) - \text{area}(A_1 \setminus A_2) \cong \text{area}(A_2 \cap B_1) - \text{area}(A_1 \cap B_2)$$

Allein bei einer Kombination der 35 Nutzungsarten des Flächenschemas (IÖR-MONITOR 2014) ergeben sich insgesamt 1190 (= $35 \cdot 34$) mögliche Nutzungsänderungen. Zudem können je nach Fragestellung diese Kategorien zu Obergruppen aggregiert werden (z. B. SuV), wodurch sich noch weitere Kombinationsmöglichkeiten ergeben. Die Selektion und Analyse der Nutzungsänderung wurde durch eine automatisierte Datenbankabfrage realisiert.

3.2 Detektieren geometrischer Verschiebungen

Untersuchungen haben gezeigt, dass zwischen den Zeitpunkten der Datensätze häufig geringe geometrische Lageverschiebungen von Objektgrenzen bzw. Linienobjekten auftreten. Mögliche Gründe dafür können Lagekorrekturen, Änderungen von Erfassungsunterlagen, sich ändernde Generalisierungsmethoden, Koordinatentransformationen oder Stützpunktausdünnung sein. Besonders bei Linienobjekten können diese geringen Verschiebungen, auch wenn sie im Rahmen der Lagegenauigkeit geschehen (± 3 m), zu erheblichen Abweichungen der Wechselbilanzen führen. Aufgrund unterschiedlicher geometrischer Gegebenheiten sind für Linien- und Flächenobjekte jeweils ein Verfahren zur Beachtung von Lageverschiebungen entwickelt worden.

3.2.1 Behandlung von Flächenobjektverschiebungen



Abb. 4: Beispiel einer Objektänderung inklusive minimaler Verschiebung beteiligter Grenzen: (a) Situation Zeitpunkt 1; (b) Situation Zeitpunkt 2 (graue Linien stellen Situation des Zeitpunkt 1 dar); (c) Fläche der sich ändernden Nutzungsart (eigene Darstellung)

Geringe Verschiebungen von zwei benachbarten Objekten unterschiedlicher Objektarten führen zu Splitterflächen. In Abb. 4 sind die Auswirkungen geringer Verschiebungen anhand von zwei Objektarten beispielhaft dargestellt. Datensätze zweier unterschiedlicher Zeitpunkte werden überlagert und die Schnittmenge der sich ändernden Nutzungsart gebildet (Abb. 4c). Es ergibt sich generell der Grundsatz, dass nach einer Schnittmengenbildung die Änderungen kompletter Objekte in ihrem ganzen Ausmaß erfasst werden (große Objekte). Geringe Verschiebungen von Grenzlagen zeigen sich dagegen durch sehr schmale Objekte (Splitterpolygone; engl.: Sliver Polygons). Diese beiden Situationen treten sehr häufig in Kombination auf, da bei Objektänderungen oft auch festgestellte Abweichungen berichtigt werden. Abb. 4c zeigt deutlich, dass sich bei der Überlagerung der beiden Datensätze ein einzelnes Objekt bildet, welches der Änderung von grün zu beige entspricht. Dieses Objekt besteht aus größeren und unerwünschten schmalen Bereichen, welche voneinander getrennt betrachtet werden sollen.

Diese Splitterpolygone werden mit Hilfe von morphologischen Operatoren (Opening) nach GONZALEZ & WOODS (2008) eliminiert. Ein Problem dabei ist jedoch, dass sich die Ecken aller Objekte abrunden, was zu einer generellen Unterschätzung führen würde. Um diesem Problem entgegenzutreten, ist ein Korrekturschritt entwickelt worden:

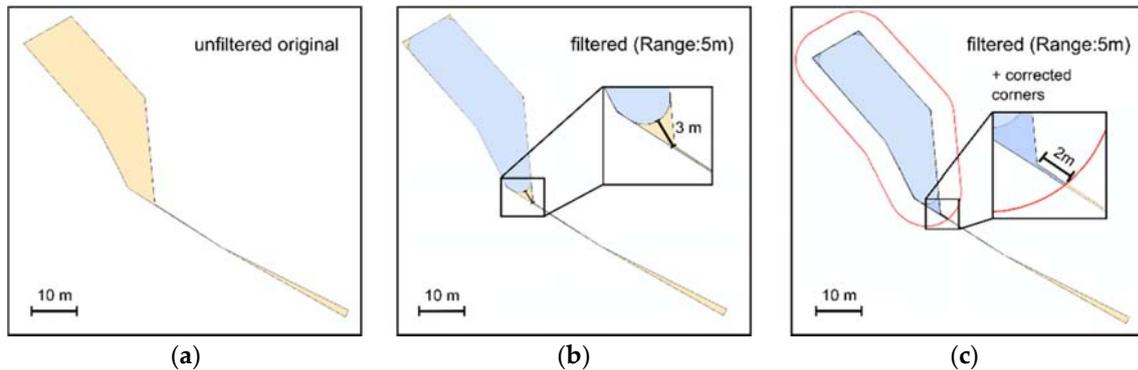


Abb. 5 Korrigiertes Opening: (a) originales Objekt; (b) Objekt nach Opening; (c) Objekt nach Opening und Korrektur der 'runden Ecken' (= korrigiertes Opening) (Quelle: eigene Darstellung)

In Abb. 5 werden ein ungefiltertes (a) und ein mit 5m mittels Opening gefiltertes Objekt (b) verglichen. Zu erkennen ist, dass zwar die schmale Fläche verschwunden ist, jedoch die Ecken des gefilterten Objekts (blau) abgerundet sind. Abb. 5c zeigt den Korrekturschritt. Wird dieses Objekt nochmals durch eine Dilatation erweitert (rote Linie) und anschließend mit dem originalen Objekt wiederum eine Schnittmenge (violett und blau) gebildet, entspricht diese bis auf einen zu vernachlässigbaren Fortsatz der gewünschten Fläche. Formal lässt sich dieses korrigierte Objekt wie folgendermaßen beschreiben:

Korrigiertes Opening:

$$uk = ((u \circ B) \oplus B) \cap u$$

Wird ein Objekt gefiltert, bedeutet das, dass in den entfernten Bereichen keine Änderung mehr stattfindet, weder eine Zu- noch Abnahme. So ist das korrigierte Opening ein geeignetes Verfahren zum Entfernen der Splitterflächen der gebildeten Schnittmengen. Damit können stabilere Ergebnisse von Wanderbewegungen erreicht werden:

3.2.2 Behandlung von Linienobjektverschiebungen

Bei Linienobjekten kann keine morphologische Filterung angewendet werden, da diese häufig schmäler als die zu filternden Splitterpolygone selbst sind. Hier wurde die von GOODCHILD & HUNTER (1997) vorgestellte Methode "A simple positional accuracy measure for linear features" zur Prüfung der Lageübereinstimmung von Linienobjekten herangezogen (Abb. 6). Dabei werden die zu vergleichenden Linienobjekte jeweils als Referenz betrachtet und gepuffert. Die Pufferbreite entspricht dabei der maximalen Lageabweichung, die zwischen zwei Linienpaaren toleriert werden soll. Die Linienobjekte des zweiten Datensatzes werden über die gepufferten Referenzobjekte gelegt und geprüft, welche sich innerhalb dieses Toleranzbereiches befinden.

Diese Methode ist auf die Bilanzierung des Flächennutzungswandels übertragen worden. Somit kann eine innerhalb der Lagegenauigkeit liegende Straßen- bzw. Gewässerachse von einer tatsächlichen neuen oder nicht mehr vorhandenen Achse unterschieden werden. Um jeweils eine separate auswertbare Menge der zu- und abgehenden Linienobjekte zu erhalten, muss diese Pufferung jeweils um die Alt- und Neudaten erfolgen:

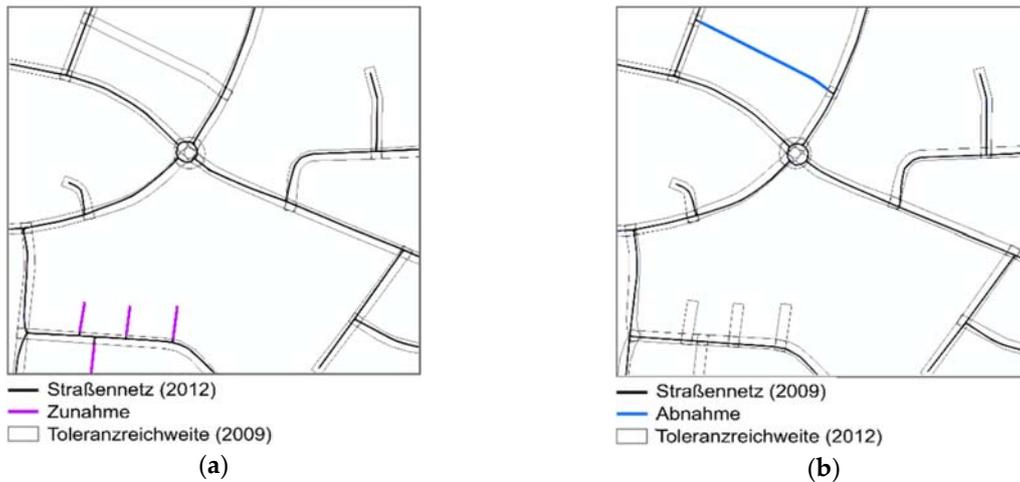


Abb. 6: Anwendung des Verfahrens von GOODCHILD & HUNTER auf die Überprüfung der Veränderung linienhaft modellierter Straßennetze mit einer Toleranzreichweite von 5 m: (a) Zunahme des Straßennetzes; (b) Abnahme des Straßennetzes (Quelle: eigene Darstellung)

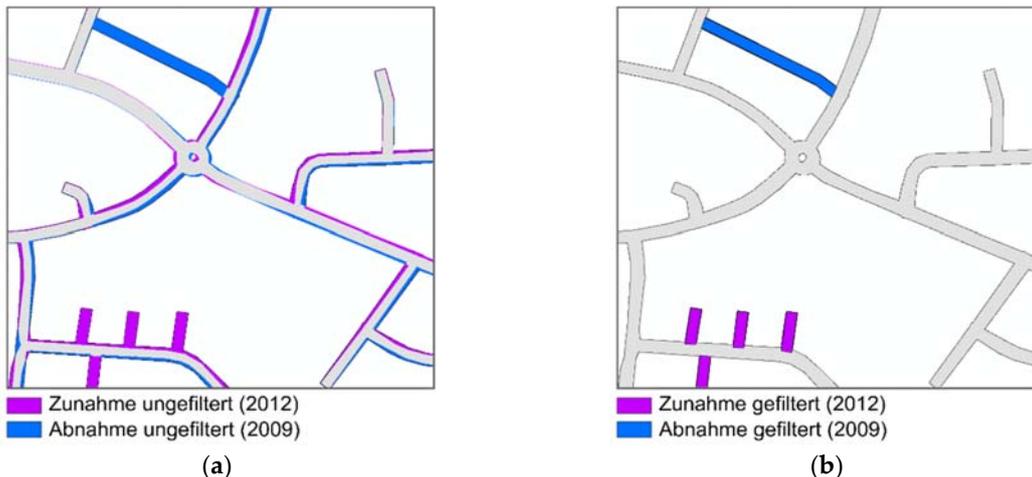


Abb. 7: Vergleich der ungefilterten (a) und gefilterten (b) Änderungen des Straßennetzes (Quelle: eigene Darstellung)

Als Input für die Bilanzierung des Flächenwandels werden hier die noch nicht gepufferten Linienobjekte des ATKIS Basis-DLM verwendet. Nur die außerhalb des Toleranzbereiches liegenden Liniensegmente werden anhand ihrer Objektbreitenangaben gepuffert. Die dadurch gebildete Fläche wird der Verschneidung zur Änderungsdetektion zugeführt. In Abb. 7 werden exemplarisch die Änderungen des Straßennetzes der ungefilterten Variante mit denen der gefilterten verglichen. Da nur Straßen außerhalb des Toleranzbereiches mit anderen Objektarten bei der Schnittmengenbildung verarbeitet werden, ergeben sich in den gefilterten Bereichen keine Änderungen. Die eben vorgestellten Verfahren zur Filterung von geringen Verschiebungen von Flächen- und Linienobjekten sind durch ein Regelwerk miteinander kombiniert. Das Verfahren zur Behandlung von Linienobjekten ist zudem erweitert worden, indem weitere Eigenschaften untersucht werden. Neben minimal verschobenen Objekten werden auch sich in der Art und Breite ändernde Linienobjekte beachtet (z. B. Umwidmung von Straßen oder nachträglich eingetragene Breitenangaben).

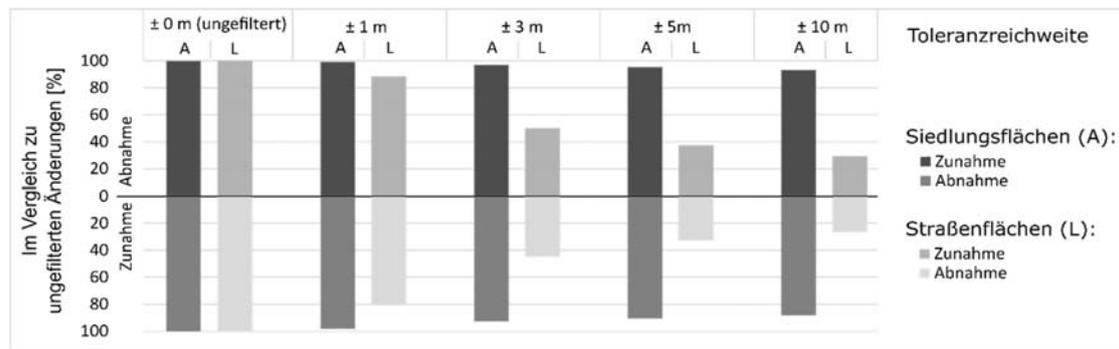


Abb. 8: Vergleich der Zu- und Abgänge von ungefilterten und gefilterten Linien (L)- und Flächenobjekten (A) in Abhängigkeit von der Filterbreite am Beispiel Schleswig-Holsteins (Quelle: eigene Berechnungen)

In Abb. 8 ist die Wirkung einer Filterung am Beispiel einer Änderungsanalyse zwischen 2009 und 2012 für Schleswig-Holstein dargestellt. Verglichen werden Siedlungsflächen mit den linienhaft modellierten Straßenflächen (gepuffert). Je schmaler ein Objekt, umso größer ist die Abweichung im Falle einer geringen Verschiebung. Bei einer Filterweite von ± 3 m wird bereits über die Hälfte der zu- und abgehenden Straßenflächen aus der Bilanz des Flächennutzungswandels ausgeschlossen. Die hohe Filterwirkung bei Linienobjekten bis ± 3 m lässt sich mit deren Lagegenauigkeit (± 3 m) erklären. Dem hingegen nehmen die Änderungen von Objekten der Siedlungsfläche nur um 10 % ab. Im Vergleich der Bundesländer ist zu beobachten, dass jeweils unterschiedlich häufig Lagekorrekturen durchgeführt werden, wobei das obige Beispiel eher als Ausnahmefall mit besonders häufigen Anpassungen betrachtet werden muss. Ist die gefilterte abnehmende Fläche gleich der hinzugehenden, ändert sich das Gesamtwachstum (Netto) nicht, jedoch sind Aussagen zu Vor- und Nachnutzungen dadurch präziser.

3.3 Beachtung modellbedingter Einflüsse

Eine Weiterentwicklung des Datenmodells kann im Falle einer Objektneuzuordnung zu Änderungen des Datenbestandes führen, welche nicht den tatsächlichen Änderungen entsprechen. Auch andere nicht nur auf das Datenmodell beruhende Einflüsse können die Modellierung bzw. den Datenbestand beeinflussen. Daher ist es besonders wichtig, die Effekte der Änderungen des Datenmodells und zusätzlicher Einflüsse der Modellierung zu beachten, um zuverlässigere Aussagen über Flächennutzungsänderungen ermitteln zu können. Die Ursache einer Objektänderung ist in den Daten des ATKIS Basis-DLM nicht enthalten. Die Migration in das neue ATKIS-Modell (*GeoInfoDok 6*) erfolgte in den Bundesländern zu unterschiedlichen Zeitpunkten zwischen 2008 und 2013 (ADV 2014). Um abzuschätzen, wie groß in diesem Zeitraum der Einfluss der Modelländerung auf die SuV-Entwicklung ist, wurden bereits bekannte Nomenklaturänderungen (KRÜGER et al. 2015) untersucht: Eine bekannte Neuzuordnung betrifft Industrie- bzw. Gewerbeflächen zu Halden. Im alten ATKIS-Modell sind Halden noch nicht als Grundflächenart geführt gewesen, sodass sich jene Flächen häufig mit Industrie- und Gewerbeflächen überlagerten. Im neuen Modell hingegen gehören Halden zu den Grundflächenarten, was wiederum eine Überlagerung mit anderen ausschließt. Dadurch sind viele Nutzungsänderungen von Industrie- bzw. Gewerbeflächen zu Halden detektiert worden, welche keiner tatsächlichen Änderung entsprechen. Des Weiteren ist bekannt, dass im zukünftigen

Datenmodell keine unbestimmten Flächen mehr erlaubt sein werden (KURSTEDT 2014), sodass schon jetzt eine deutliche Abnahme jener unbestimmten Flächen festzustellen ist. Eine weitere Modelländerung betrifft die Verkehrsbegleitflächen. So sind im alten Modell keine Verkehrsbegleitflächen definiert gewesen, während diese im neuen Modell durch ein Attribut beschrieben sind. Auffällig häufig sind daher Änderungen von Grünland zu Verkehrsbegleitflächen. Kleingärten waren ebenfalls im alten Modell noch nicht definiert und wurden dort als Gartenland geführt. Seitdem im neuen Modell Kleingärten definiert sind, werden zwischen diesen Nutzungsarten Neuzuordnungen durchgeführt.

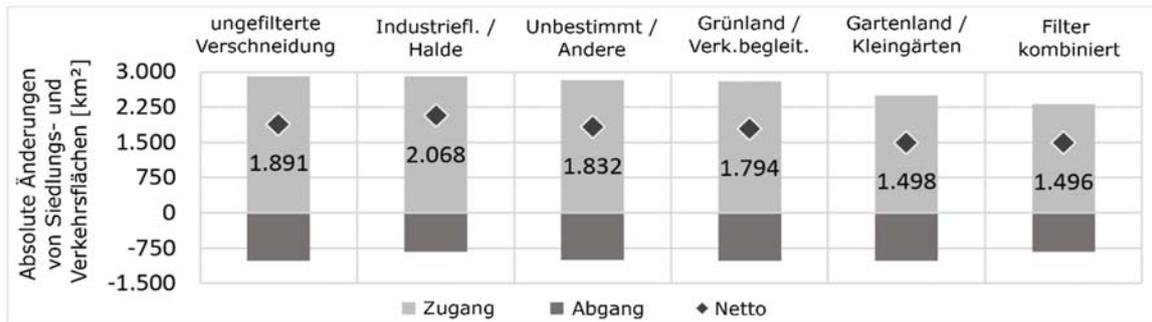


Abb. 9: Einfluss verschiedener Migrationseffekte auf die Entwicklung der SuV-Fläche Deutschlands während der ATKIS-Migrationsphase zwischen 2008 – 2013 (Quelle: eigene Berechnungen)

In Abb. 9 sind die Auswirkungen der eben beschriebenen modellbedingten Änderungen auf die Bilanz der SuV-Flächen veranschaulicht. Den größten Einfluss hat die Modelländerung von Gartenland zu Kleingärten. Allein 20 % des Gesamtwachstums der SuV-Flächen beruht auf derartigen Modelländerungen, was im Vergleich zu Untersuchungen von älteren nicht belasteten Änderungsraten ein extrem unwahrscheinlich hoher Wert ist. Um eine deutliche Überschätzung der SuV-Entwicklung im Migrationszeitraum zu vermeiden, sollten diese Flächen bei der Wechselbilanzierung nicht einbezogen werden. Nach Abschluss der Migration, d.h. mit den ATKIS-Daten nach 2013, sind diese Effekte nicht mehr in diesem Umfang zu erwarten.

4 Fazit

Das Verfahren zur Bilanzierung des Flächennutzungswandels ermöglicht Aussagen zu Vor- und Nachnutzungen einzelner Nutzungsarten. Dies hilft Rückschlüsse auf die Triebkräfte des Flächennutzungswandels zu identifizieren. Bei der Bilanzierung werden verschiedene Verfahren verwendet, um Effekte durch geometrisch leicht verschobene Objekte und modellbedingte Objektart- und Attributänderungen auszufiltern. Ohne Berücksichtigung von Modelländerungen und von immer wieder auftretenden geringen Lagekorrekturen der Geobasisdaten werden die Bilanzen des Flächenwandels teilweise erheblich verfälscht. Insbesondere betrifft dies Wechselbilanzen im Zeitraum des ATKIS-Modellwechsels. So liegt die tägliche Flächenneuanspruchnahme von nicht bereinigten Bilanzen der SuV-Flächen nach dem IÖR-Monitor bei 105 ha/d. Unter Beachtung der Migrationseffekte liegt diese immer noch bei 84 ha/d. Das Verfahren wird wegen zukünftig noch erfolgenden Veränderungen des Basis-DLM im Zuge der Harmonisierung von ATKIS und ALKIS ein großes Einsatzpotenzial im Flächennutzungsmonitoring haben.

5 Literaturverzeichnis

- ADV, 2008a: 'Hauptdokument GeoInfoDok 6.0: Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens', GeoInfoDok 6.0, Arbeitsgemeinschaft der Vermessungsverwaltungen
- ADV, 2008b: ATKIS-Objektartenkatalog Basis-DLM: Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland, 129.
- ADV, 2014: Sachstand der Migration: AAA-Umfrage zum Sachstand der Migration in den Ländern. www.adv-online.de/AAA-Modell/Sachstand-der-Migration, Zugriff 28.09.2014.
- BBSR, 2015: Datengrundlagen zur Beobachtung der Flächenentwicklungen in Deutschland. <http://www.bbsr.bund.de/BBSR/DE/Raumentwicklung/Flaechenpolitik/Projekte/Archiv/FlaechenmonitoringDetail/FlaechenmonitoringDetail.html?nn=391978>, Zugriff 15.11.2014.
- GONZALEZ, R.C. & WOODS, R.E., 2008: Digital image processing. 3rd ed., Pearson/Prentice Hall, Upper Saddle River, NJ, ISBN: 978-0-13-168728-8.
- GOODCHILD, M. & HUNTER, G., 1997: A simple positional accuracy measure for linear features. *International Journal of Geographical Information Science*, Taylor and Francis Ltd., 299-306.
- HAGEDORN, M. & VELTKAMP, R.C., 2001: State of the Art in Shape Matching. *Principles of Visual Information Retrieval*, Springer, London, 87-119.
- IÖR-MONITOR, 2014: Monitor der Siedlungs- und Freiraumentwicklung des Leibniz-Instituts für ökologische Raumentwicklung (IÖR). www.ioer-monitor.de, Zugriff 12.08.2014.
- KRÜGER, T., 2011: Aktuelle Ergebnisse des IÖR-Monitors zur Flächennutzung in Deutschland: Flächennutzungsmonitoring III. Erhebung – Analyse – Bewertung. Publikation des Instituts für ökologische Raumentwicklung, Meinel, G. & Schumacher, U. (Hrsg.), Berlin: Rhombos, IÖR Schriften Band **58**, 23-35.
- KRÜGER, T., HENNERSDORF, J., MEINEL, G. & BEHNISCH, M., 2015: Migration des ATKIS-Basis-DLM – Auswirkungen auf die Nutzung für das Flächenmonitoring: Kartographische Nachrichten **65** (2), 59-66.
- KURSTEDT, R., 2014: TOP Neues in der GeoInfoDok 7.0. 51. Tagung AG: Automation in Kartographie, Photogrammetrie und GIS, Hannover, 8.-9.9.2014.
- MEINEL, G., 2013: Aktuelle Ergebnisse und Entwicklungen im IÖR-Monitor -Anforderungen an ein Flächennutzungsmonitoring: Flächennutzungsmonitoring V. Methodik – Analyseergebnisse – Flächenmanagement, Meinel, G., Schumacher, U. & Behnisch, M. (Hrsg.), Berlin, Rhombos, IÖR Schriften **61**, 117-118.
- MEINEL, G. & KRÜGER, T., 2014: Methodik eines Flächennutzungsmonitorings auf Grundlage des ATKIS-Basis-DLM: Kartographische Nachrichten **64** (6), 324-331.
- RÖBER, B., HEINRICH, U. & ZÖLITZ, R., 2009: Über die Eignung von ATKIS als topographischer Basisdatensatz für numerische Modelle. *GIS Science* 2009, 12-18.
- UBA, 2014: Flächensparen – Böden und Landschaften erhalten. www.umweltbundesamt.de/themen/boden-landwirtschaft/flaechensparen-boedenlandschaften-erhalten, Zugriff 24.08.2014.

Array-Datenbanken für semantische inhaltsbasierte Suche und Analyse in Satellitenbildarchiven

MARTIN SUDMANN¹, DIRK TIEDE¹, NIKOLAUS AUGSTEN²,
ANDREA BARALDI^{1/3}, MARIANA BELGIU¹ & STEFAN LANG¹

Zusammenfassung: Die steigende räumliche und zeitliche Auflösung von Satellitenbildern birgt Herausforderungen beim Durchsuchen und Analysieren in Satellitenbildarchiven dar. Semantische inhaltsbasierte Suchmöglichkeiten und web-basierte Analysen in der Datenbank unterstützen den Umgang mit der steigenden Datenmenge. Dieser Beitrag stellt eine Methode und Implementierung vor, wie Bilddatenbanken durch semantische Abfragen besser in Wert gesetzt werden. Dies basiert auf der Anreicherung der Satellitenbilder mit vollautomatisch vorverarbeiteten semantischen Layern. Dies beinhaltet (i) eine Basisarchitektur mit einer Array-Datenbank und einem web-basierten grafischen Abfragesystem, (ii) ein Vergleich von Datenbanken zur Speicherung von Rasterdaten und (iii) die Implementierung in einem Prototyp (IQ – ImageQuerying). Die Idee und der Prototyp, die mit dem Copernicus Masters 2015 - T-Systems Big Data Challenge ausgezeichnet wurden, sollen dazu beitragen, Satellitenbilder effektiver nutzen zu können.

1 Einleitung

Aufgrund der stetig steigenden räumlichen und zeitlichen Auflösungen von Satellitenbildern und der damit einhergehenden Datenmengen ist die Erdbeobachtung mit zunehmenden Herausforderungen bei der Auswertung dieser Daten konfrontiert. Schon heute können die existierenden Methoden und technischen Möglichkeiten mit dieser Entwicklung kaum Schritt halten, z.B. wird nur ein geringer Prozentsatz aller Satellitenbilder der ESA-Archive heruntergeladen und ausgewertet (BARALDI & BOSCHETTI 2012). In diesem Zusammenhang wird auch der Begriff *dark data* (PURSS et al. 2015) für Daten verwendet, die zwar erhoben und gespeichert, aber niemals genutzt werden. Ein Grund, der hierfür genannt wird ist, dass keine geeigneten Werkzeuge existieren, um mit der Datenmenge intelligent umzugehen und deshalb die für bestimmte Analysezwecke richtigen Satellitenbilder oft gar nicht gezielt gesucht und gefunden werden können, um sie auszuwerten und Redundanzen zu vermeiden (BARALDI & BOSCHETTI 2012).

Ein Beispiel hierfür sind die Datenmengen der neuen Sentinel Satelliten, die im Jahr 2016 geschätzt 5.000 Terabyte an neuen Daten produzieren werden - mit steigender Tendenz (RECK et al. 2016). Die Auswertung dieser Datenmengen, die Selektion von geeigneten Bildern für bestimmte Anwendungszwecke und die Transformation in zuverlässige Informationsprodukte

¹ Paris-Lodron Universität Salzburg, IFFB Geoinformatik, Schillerstraße 30, A-5020 Salzburg, E-Mail: [martin.sudmanns, dirk.tiede, mariana.belgiu, stefan.lang]@sbg.ac.at

² Paris-Lodron Universität Salzburg, Fachbereich Computerwissenschaften, Jacob-Haringer Straße 2, A-5020 Salzburg, E-Mail: nikolaus.augsten@sbg.ac.at

³ Department of Agricultural and Food Sciences, University of Naples Federico II, Italy, E-Mail: andrea6311@gmail.com

stellen ein mittlerweile anerkanntes Problem dar (RECK et al. 2016; SOILLE et al. 2016). Betrachtet man Kosten und Aufwand für Satellitenmissionen und die Tatsache, dass derzeit viel Potential ungenutzt bleibt, ist es aus wissenschaftlicher, gesellschaftlicher und ökonomischer Sicht wichtig, Satellitenbilder effizienter zu nutzen.

Dieser Beitrag beschreibt eine Möglichkeit zur Aufwertung existierender Fernerkundungs-Bilddatenbanken, welche um semantische, inhaltsbasierte Abfrage- und Analysemöglichkeiten erweitert werden. Die erste Herausforderung hierbei ist, Computeralgorithmen so intelligent zu gestalten, damit diese die Inhalte der Bilder erkennen und interpretieren können. Jedoch bleiben computerbasierte Möglichkeiten nach wie vor hinter den Fähigkeiten eines menschlichen Betrachters zurück. Diese Diskrepanz wird als Semantische Lücke (*semantic gap*) bezeichnet. Dadurch sind die Erwartungen der Nutzer/-innen an ein Bildabfragesystem schwer zu erfüllen. Um diese Lücke zu verkleinern, basiert das vorgestellte System auf einer hybriden Methode, die sowohl eine vollautomatische Vorklassifikation der Daten basierend auf physikalischen Modellen (a-priori Wissen) als auch das von Benutzer/-innen als Abfrage eingebrachte Wissen nutzt. Das externe Wissen wird über eine web-basierte grafische Benutzeroberfläche (graphical user interface, GUI) in das System eingebracht, in Datenbankabfragen transformiert und abgespeichert. Dadurch können die Abfragen mit anderen Benutzer/-inne/n geteilt werden. Es ist daher ein inkrementell wachsendes System, das durch die Benutzung zunehmend mehr Möglichkeiten bietet. Die zweite Herausforderung ist verbunden mit den großen Datenmengen, da die klassische Herangehensweisen zur Speicherung von Satellitenbildern für eine datenbankbasierte semantische Suche und Auswertung von Satellitenbildern ungeeignet und ineffizient ist. Alternativ dazu werden die Daten in dem vorgestellten System in einem dreidimensionalen Datenmodell in einer Array-Datenbank gespeichert, um Abfragen durch Raum und Zeit effizient durchführen zu können.

2 Motivation und Stand der Technik

Zur Illustration soll ein gängiges Beispiel dienen: Ein/e Benutzer/in ist auf der Suche nach allen für ihre/seine Anwendung qualifizierten Satellitenbildern für ein ausgewähltes Gebiet (area-of-interest, AOI). Die Bedingung soll sein, dass die Aufnahmen (über diesem Gebiet) wolkenfrei sind. Hierfür muss der/die Benutzer/in in den meisten Fällen zunächst alle Bilder herunterladen (wenn die Datenvorschau nicht ausreicht) und anschließend die unbrauchbaren Aufnahmen manuell aussortieren. Selbst wenn eine Wolkenbedeckung als Qualitäts- und Suchkriterium angegeben werden kann, ist diese für das gesamte Bild global ausgewiesen. Es ist hingegen durchaus möglich, dass ein Bild insgesamt viele Wolken hat, aber das von dem/der Benutzer/in ausgewählte Gebiet wolkenfrei ist – oder umgekehrt. Ein System, das während der Abfrage und basierend auf dem ausgewählten Gebiet die Wolkenbedeckung berechnet und damit die qualifizierten Satellitenbilder automatisch findet, würde dem/der Benutzer/in in diesem Fall viel Arbeit ersparen können oder gewisse Abfragen überhaupt erst möglich machen. Dieses Beispiel lässt sich beliebig komplex erweitern, in dem anstatt Wolken eine Landnutzungs-kategorie oder eine Kombination davon abgefragt wird.

Diese Art der Fragestellung wird allgemein als inhaltsbasierte Suche (content-based image retrieval, CBIR) bezeichnet und gewinnt im Zeitalter von big data, in dem eine manuelle

Auswertung fehlerhaft, langwierig oder gar ganz unmöglich ist, immer mehr an Bedeutung. Derzeitige Bilddatenbanken erlauben jedoch keine inhaltsbasierte Abfragen für benutzerdefinierte Gebiete. Sowohl beispielsweise im USGS Earth Explorer Portal (<http://earthexplorer.usgs.gov/>) als auch im Sentinels Scientific Data Hub (<https://scihub.copernicus.eu/>) sind Abfragen über die geographische Position der Bilder, die Aufnahmezeit und textuelle Metadaten vorgesehen. Es wäre zwar in dem Beispiel möglich, ein spezielles Computerprogramm zu entwickeln, das die heruntergeladenen Bilder nach den gewünschten Kriterien aussortiert. Allerdings ist die Entwicklung eines in diesem Falle sensor- und abfrageabhängigen Programms umständlich, nicht übertragbar und auch nicht für große Datenmengen geeignet. Der derzeitige Ansatz, die Daten herunterzuladen und lokal zu verarbeiten ist nur zu einer begrenzten Datenmenge technisch überhaupt möglich. Für große Untersuchungsgebiete und multitemporale Fragestellungen können bald Terrabytes oder möglicherweise sogar Petabytes an Daten anfallen, die über das Internet gesendet werden müssten. Auch mit steigenden Übertragungsgeschwindigkeiten bleibt es effizienter, nicht die Daten zu den Benutzer/-innen zu transportieren, sondern den Algorithmus zu den Daten. Bestehenden Ansätzen für Online-Analysen auf der Datenbank-Ebene wie dem EarthServer (BAUMANN et al. 2015) oder dem Australian Geoscience Data Cube (AGDC) (PURSS et. al 2015) fehlt die Möglichkeit, beliebige benutzerdefinierte Abfragen und Analysen auf einem höheren semantischen Level durchzuführen.

Ein System, das semantische, inhaltsbasierte Abfragen und Analysen über das Internet möglich macht ist ein neuer Ansatz, um Bilddatenbanken zu verbessern und Satellitenbilder effizienter zu nutzen. Dafür gibt es mindestens zwei Anwendungsgebiete, die im Zusammenhang stehen: (i) inhaltsbasierte Suche als Erweiterung herkömmlicher Suchen und (ii) eine web-basierte Online-Analyse, bei der das Ergebnis entweder ein neuer, georeferenzierter Datensatz (Analyseergebnis) ist mit räumlich-zeitlicher Auswertung (z.B. *Wo hat sich eine Landnutzung in den letzten 5 Jahren verändert?*) oder einer zeitlichen Auswertung (*In welchem Jahr gab es die meisten Änderungen?*).

3 Methodologische Grundlagen

Um Satellitenbilder inhaltlich abfragbar zu machen, müssen diese semantisch durch einen oder mehrere semantische Layer angereichert werden, die den Inhalt auf einem sub-symbolischen oder symbolisch-kategorischen Level beschreiben. Diese können beispielsweise eine Landnutzungs-klassifikation oder auch ein NDVI beinhalten. Um den Ansatz für big data Fragestellungen nutzbar zu machen ist allerdings eine wesentliche Bedingung, die semantischen Layer parameterfrei und vollautomatisch zu erstellen. Für das vorgestellte System wurde die vollautomatische Vorklassifikation nach BARALDI et al. (2010) verwendet. Da die semantischen Layer direkt mit den Sensordaten verknüpft sind, sind inhaltsbasierte Abfragen darauf möglich. Die semantischen Layer bilden dann die Grundbausteine für diese Abfragen. Die Begriffe ‚semantischer Layer‘ und ‚Informationslayer bzw. –ebene‘ werden in diesem Beitrag synonym verwendet.

Die Satellitenbilder und deren assoziierte semantische Layer müssen an einer zentralen Stelle in einem Datenmodell gespeichert werden, das beliebig komplexe Abfragen in akzeptablen Zeiten ermöglicht. Auch wenn derzeit Fernerkundungsdaten als zweidimensionale Bilder direkt im Dateisystem verwaltet werden, bringt die Speicherung in Datenbanken mehrere Vorteile. Die Abfragen erfolgen über eine deklarative Abfragesprache, mit der die Datenbank für jede Abfrage

über die beste Zugriffsstrategie entscheiden und interne Optimierungen durchführen kann (BAUMANN & HOLSTEN 2011). Zusätzlich können ein optimiertes Datenmodell und verschiedene Indexierungsmethoden bestimmte Zugriffsmuster auf die Daten unterstützen. Beispiele sind Zeitreihenanalysen oder wenn das von den Benutzer/-innen gewählte Gebiet mehr als ein Bild oder nur Teile eines Bildes umfasst. Klassische Vorteile von Datenbanken wie erhöhte Sicherheit, Transaktionsmanagement oder bessere Backupstrategien können ebenfalls von Relevanz sein (BRINKHOFF & KRESSE 2011). Die strikte Trennung von logischem Modell (Abfragesprache) und physischem Modell (z.B. Dateiformate, Indexstrukturen, Algorithmen für den Datenzugriff) in Datenbanksystemen erleichtert die Wartbarkeit der Anwendungen, und ermöglicht die Optimierung von Anfragen getrennt von Anwendungsfällen zu betrachten. Verbesserungen im Datenbanksystem führen zu effizienteren Anfragen, ohne dass die Anwendung selbst geändert werden muss. Aus diesen Gründen wird für das vorgeschlagene System eine Array-Datenbank zur Speicherung der Bilder verwendet.

Eine web-basierte grafische Benutzeroberfläche dient als Schnittstelle zwischen den Benutzer/-innen und der Datenbank. Diese ist dabei so konzipiert, dass das Wissen der Benutzer/-innen in maschinenlesbarer Weise in die Wissensbasis und damit in das System codiert und eingebracht werden kann. Hierbei muss berücksichtigt werden, dass semantische, raum-zeitliche Abfragen über große Komplexität verfügen können und daher oft schwierig zu erstellen sind. Die GUI muss daher Möglichkeiten bieten, die Abfragen unterstützt und übersichtlich zu erstellen, diese zu speichern und zu teilen und damit auch anderen Benutzer/-innen zur Verfügung zu stellen.

Für diesen Anwendungszweck wurden vier mögliche Datenbanken zur Speicherung von Rasterdaten untersucht: (1) Oracle Spatial and Graph als Erweiterung der Oracle Datenbank (ORACLE 2016), (2) PostGIS als Erweiterung von PostgreSQL (POSTGIS 2016), (3) Rasdaman (RASDAMAN 2016) und (4) SciDB (SCIDB 2016). Diese lassen sich weiter unterteilen in Datenbanken, die auf dem relationalen Modell aufbauen (1,2) und Datenbanken, die über ein eigenes Array-Datenmodell verfügen (3,4).

Basierend auf den Besonderheiten von Satellitenbildern sowie in der Literatur vorgeschlagenen Datenbankvergleichen (BAUMANN & STRAMERJOHANNIS 2014), wurde eine zugeschnittene Testumgebung geschaffen, die sowohl eine qualitative Analyse als auch einen Performance-Test umfasst. Das Ziel der qualitativen Analyse soll als Entscheidungsgrundlage für oder gegen die Benutzung einer jeweiligen Datenbank für den vorgestellten Anwendungszweck dienen, die Performance-Tests sollen ein Eindruck vermitteln, wie sich die Datenbanken insgesamt verhalten und wo Stärken oder Schwächen liegen.

Auf Grundlage der Datenbankanalyse wurde schließlich eine Kombination der beschriebenen Komponenten in einer neuen, generischen Basisarchitektur und als Nachweis der Machbarkeit in einem Prototyp implementiert.

4 Ergebnisse

Für die Fragestellung sind vier Ergebnisse gesondert zu betrachten. Diese sind (i) ein Vergleich der vier Datenbanken, (ii) die Basisarchitektur des neuen Systems, (iii) das web-basierte Abfragesystem und (iv) der Prototyp des Systems zur semantischen, inhaltsbasierten Suche und Analyse in Bilddatenbanken (IQ – ImageQuerying).

4.1 Evaluation der Datenbanken

Für den vorliegenden Anwendungsfall haben die Datenbanken mit Array-Datenmodell vor allem bei Abfragen durch die Zeit besser abgeschnitten. Das Array-Datenmodell unterscheidet sich von Bildern als Dateien hauptsächlich dadurch, dass die Daten nicht nur auf Basis der räumlichen, sondern auch der zeitlichen Dimension als Würfel angeordnet sind (Abb. 1). Durch die Gleichberechtigung der Dimensionen sind Abfragen durch die Zeit signifikant beschleunigt. Zusätzliche wichtige Elemente sind die deklarative Abfragesprache ähnlich SQL und das intelligente Partitionieren und Indexieren, das es der Datenbank ermöglicht, nur auf die Daten zuzugreifen, die für eine bestimmte Abfrage wirklich benötigt werden.

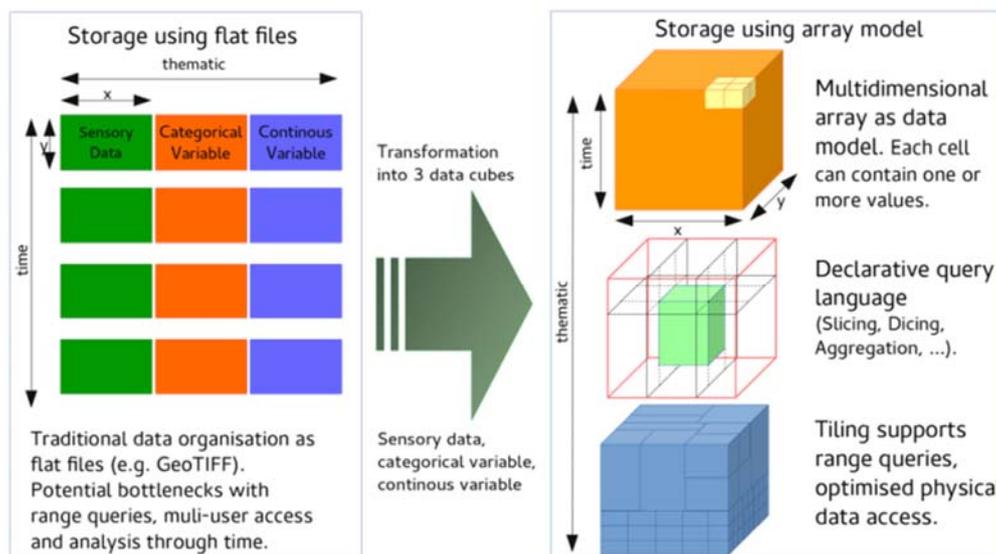


Abb. 1: Eigenschaften des Array-Datenmodells. Eigene Grafik, 2016.

Die experimentellen Ergebnisse haben gezeigt, dass vor allem die Analysen - je nach Art - wesentlich effizienter sein können als mit herkömmlichen Speichermethoden. Hingegen könnten sich die Zeiten für Import und Export in die Datenwürfel negativ auswirken. Es kann daher ratsam sein, die Daten im Vorfeld zu klassifizieren und entweder in import/export-optimierte oder analyse-optimierte Datenmodelle zu speichern. Für den hier genannten Anwendungsfall der inhaltsbasierten Suche ist dies machbar: Für die Analysen können die semantischen Layer in Datenwürfel genutzt, für den Download können die originalen Satellitenbilder genutzt werden. Darüber hinaus konnten Entwicklungs- und Erweiterungsgebiete im Datenbankbereich identifiziert werden, die zwar in klassischen Datenbanken vorhanden sind, aber im betrachteten Anwendungsfall von Array-Datenbanken nicht oder nur mangelhaft unterstützt werden. Dazu gehören erweiterte Elemente für Abfragesprachen, die auch räumlich-zeitliche Referenzierungen innerhalb des Datenwürfels erlauben (z.B. durch benutzerdefinierte Funktionen (user-defined functions, UDF) und benutzerdefinierte Datentypen), Indexierung der semantischen Layer (z.B. B+ tree, Hash Index, R-tree index, Bitmap index) oder Hinweise zur allgemeinen Optimierung (z.B. der Tiling-Strategie oder Pyramid-Layers) und Abfrageplanung.

4.2 Web-basiertes Abfragesystem

Die GUI des Abfragesystems ist eine in JavaScript programmierte Web-Applikation, die sich mit einem Standard-Browser bedienen lässt. Zentrales Element ist das Abfragepanel. In diesem wird die Abfrage sequentiell aufgebaut durch die Kombination von Untergruppen wie die Auswahl des Gebiets (AOI) und der Zeit sowie Kriterien für inhaltsbasierte Abfragen.

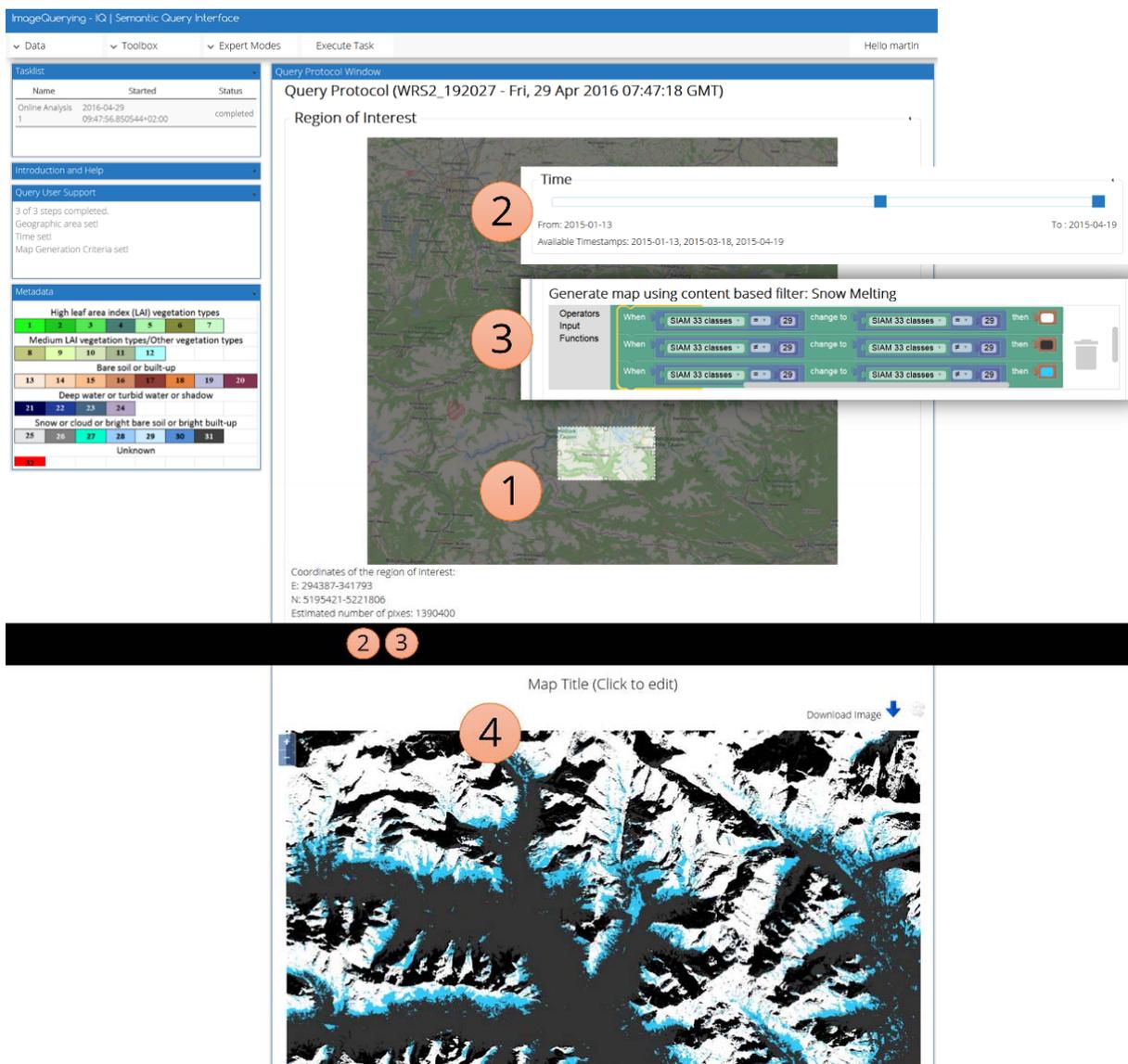


Abb. 2: Das web-basierte Abfragesystem (IQ-Frontend, Stand 29.04.2016) erlaubt es dem Benutzer, die Datenbank grafisch abzufragen. Hierfür selektiert der Benutzer im Abfragepanel zunächst die Area-of-Interest (1), den Zeitbereich (2) und erstellt die Regel zur inhaltsbasierten Abfrage mit grafischen Elementen (3). Nach der Ausführung erscheint das Ergebnis unterhalb der Eingabeparameter. Das vorliegende Beispiel zeigt im Nationalpark Hohe Tauern die Bereiche (blau) mit einer Abnahme der Schneebedeckung zwischen Januar und April 2015 und Bereiche (weiß) mit einer konstanten Schneebedeckung (verändert nach TIEDE et al. 2016).

Das Ergebnis der Abfrage oder der Analyse wird ebenfalls im Abfragepanel angezeigt. Im Gegensatz zu anderen Ansätzen erspart dies den Benutzer/-innen das vor- und zurück-navigieren und erleichtert so den Überblick und das Revidieren von komplexen Abfragen. Seitlich sind Bereiche mit Kontextinformationen wie beispielsweise Metadaten oder Hilfen, die immer im Blickfeld der Benutzer/-innen sind. Die Erstellung des Regelsatzes für die inhaltsbasierte Suche oder Analyse erfolgt ebenfalls grafisch. Abb. 2 zeigt ein Screenshot des operationellen Systems mit einer Beispielabfrage.

4.3 Architektur

Die neue Architektur (Abb. 3) kombiniert die Komponenten, erlaubt das Speichern der Daten in einer Array-Datenbank und den Zugriff über eine web-basierte Benutzeroberfläche, die mit einem Standardbrowser bedient werden kann.

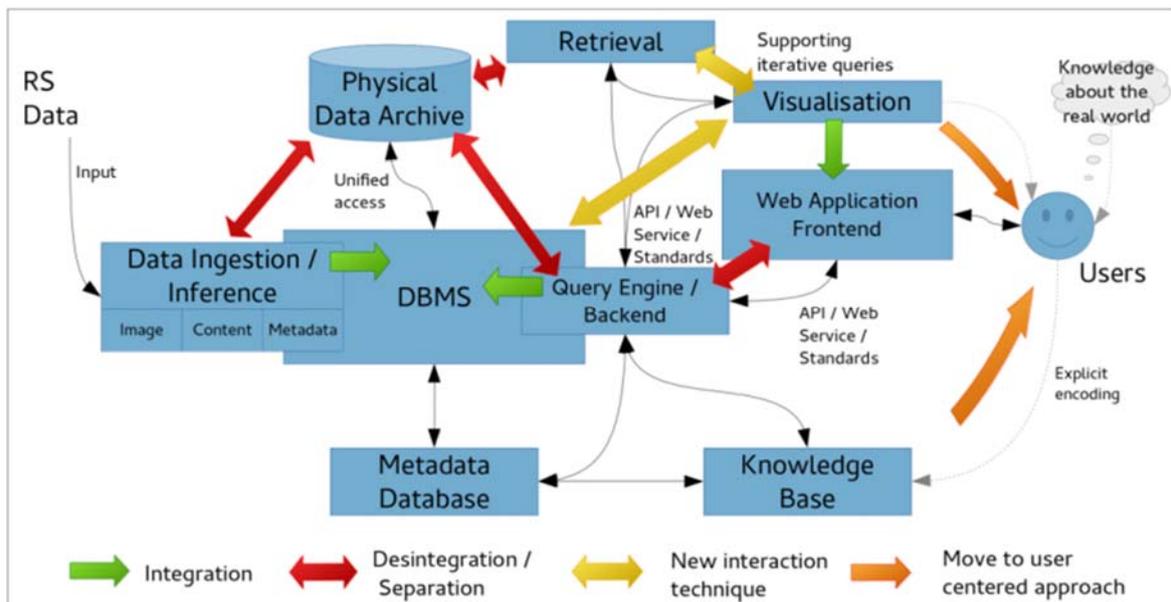


Abb. 3: Basisarchitektur mit Array-Datenbank. Eigene Grafik, 2016.

4.4 IQ – ImageQuerying

IQ ist ein Prototyp, der die drei Komponenten (i) automatisch vorverarbeitete semantische Layer (siehe Nummer 1), (ii) eine Array-Datenbank – in diesem Fall Rasdaman (2) und (iii) eine web-basierte Abfrage beinhaltet (3) (Abb. 4).

Das Datenmodell ist in einer Rasdaman Installation implementiert. Dabei werden die Daten in drei Raum-Zeit-Würfel transformiert, jeweils einen für die Sensordaten, für semantische Layer mit kategorischen Variablen und semantische Layer mit kontinuierlichen Variablen. Die kategorialen Variablen bestehen in dieser Implementierung aus: einer Vorklassifikation mit 33 Klassen, einer mit 96 Klassen und einer bitemporale Klassifikation, die aus dem Vergleich zwischen zwei aufeinanderfolgenden Zeitschnitten berechnet wurde. Räumlich benachbarte Pixel, die die gleiche Klassifikation aufweisen, wurden zu Objekten aggregiert. Darauf basierend wurden die Größe und

Formparameter berechnet, die im Datenwürfel für kontinuierliche Variablen gespeichert sind. Eine dritte kontinuierliche Variable ist der *Greenness*-Index für Vegetationsanalysen.

Benutzer/-innen können die Datenbank mit vorgefertigten grafischen Elementen abfragen, die sich wie Puzzle-Teile zu beliebig komplexen Regelsätzen zusammenfassen und kombinieren lassen. Alle Komponenten des Systems sind auf eine Skalierung in den big data Bereich ausgerichtet.

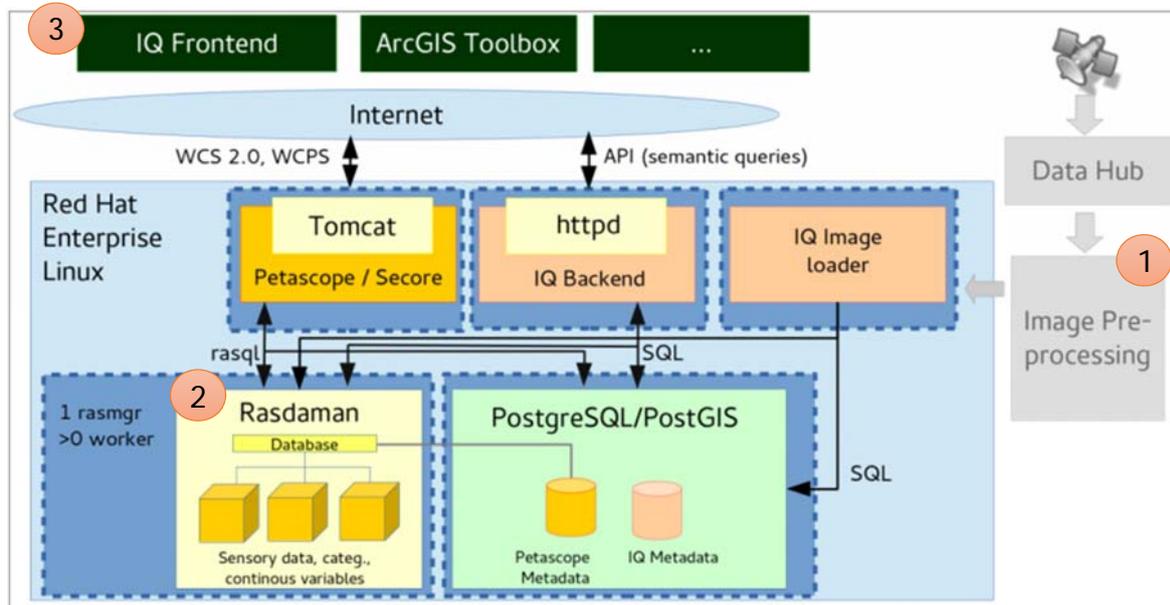


Abb. 4: Implementierung als ImageQuerying (IQ). Eigene Grafik, 2016.

5 Schlussfolgerungen und Ausblick

Die vorgestellte Methode und Implementierung soll Satellitendaten effizienter für Benutzer nutzbar machen. Die Frage, wie man mit den großen Datenmengen, die von den derzeitigen und geplanten Satelliten aufgenommen werden, effizient umgeht, ist relevant für die verbesserte Nutzung und das Ausschöpfen des Potentials von Satellitenbildern. Dieser Beitrag hat eine mögliche Lösung aufgezeigt, die daraus besteht, existierende Bilddatenbanken durch zusätzliche Layer semantisch anzureichern. Diese sind automatisch vorverarbeitet und Grundlage für semantische, inhaltsbasierte Abfragen. Dadurch ist es möglich, zielgerichtete und effizientere Abfragemöglichkeiten anzubieten und darüber hinaus Analysen direkt web-basiert auf der Datenbankebene durchzuführen. Da es derzeit kaum operationelle vergleichbare Systeme oder Alternativen gibt, ist dies derzeit ein innovatives Alleinstellungsmerkmal.

Es hat sich bezüglich Datenmenge und Performance als zielführend herausgestellt, die semantischen Layer mit den Eigenschaften von Array-Datenbanken zu kombinieren. Über eine web-basierte Schnittstelle lässt sich die Datenbank abfragen bzw. Analysen auf den Daten durchführen. Hierfür wurde eine Webapplikation entwickelt, in der die Benutzer/-innen auf eine einfache Art und Weise beliebig komplexe Abfragen erstellen, teilen und benutzen können. Dies ermöglicht die Transformation von Wissen in maschinenlesbare Elemente und verkleinert die semantische Lücke, indem das System inkrementell intelligenter wird.

Als nächster Schritt ist eine Optimierung hinsichtlich der Performance und Verwaltung von größeren Datenmengen angedacht. Für die Entwicklung oder Implementierung von verbesserten Methoden ist eine enge Zusammenarbeit mit der Datenbankforschung notwendig. Darüber hinaus wird die Wissensbasis durch weitere semantische, räumlich-zeitliche Abfragen erweitert. Hierfür wird der Prototyp in einer realen Anwendung getestet, wobei auch der Input von verschiedenen Domain-Experten (für unterschiedlichste Fragestellungen) von Bedeutung sein wird. Weitere Forschung an der Schnittstelle zwischen Fernerkundung und Computer Vision kann zur Verbesserung der automatisch erstellten semantischen Layer beitragen. Beispielsweise haben Neuronale Netze (speziell Bildkontext-sensitive convolutional neural networks, CNN) in den letzten Jahren viel Potential gezeigt (LÄNGKVIST et al. 2016), sodass auch diese ergänzend für die Interpretation der Satellitenbilder und bei der Erstellung der semantischen Layer genutzt werden könnten.

6 Danksagung

Die Implementierung des IQ Systems als Prototyp erfolgte innerhalb des FFG (österreichische Forschungsgemeinschaft) geförderten Projektes AutoSentinel-2/3 (Vertragsnummer: 848009) im Rahmen des Austrian Space Application Programme (ASAP) und wurde mit dem *Copernicus Masters Award 2015 – Big Data Challenge* ausgezeichnet (http://www.copernicus-masters.com/index.php?kat=winners.html&anzeige=winner_t-systems2015.html).

7 Literaturverzeichnis

- BARALDI, A. & BOSCHETTI, L., 2012 Operational automatic remote sensing image understanding systems: Beyond geographic object-based and object-oriented image analysis (GEOBIA/GEOOIA). Part 1: Introduction. *Remote Sensing* **4** (9), 2694-2735.
- BARALDI, A., DURIEUX, L., SIMONETTI, D., CONCHEDDA, G., HOLECZ, F. & BLONDA, P., 2010: Automatic spectral rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS, AVHRR/MSG, AATSR, IKONOS/QuickBird/OrbView/GeoEye and DMC/SPOT-1/-2 imagery – Part I: System design and implementation. *IEEE Trans. Geoscience Remote Sensing* **48** (3), 1299-1325.
- BAUMANN, P., MAZETTI, P., UNGAR, J., BARBERA, R., BARBONI, D., BECCATI, A., BIGAGLI, L., BOLDRINI, E., BRUNO, R., CALANUCCI, A., CAMPALANIA, P., CLEMENTSI, O., DUMITRUA, A., GRANTI, M., HERZIGJ, P., KAKALETRISK, G., LAXTONL, J., KOLTSIDAK, P., LIPSKOCHA, K., MAHDIRAJA, A.R., MANTOVANI, S., MERTICARIUA, V., MESSINAM, A., MISEVA, D., NATALIH, S., NATIVIC, S., OOSTHOEKA, J., PAPPALARDOM, M., PASSMOREN, J., ROSSIA, A.P., RUNDUE, F., SENN, M., SORBERAE, V., SULLIVANO, D., TORRISIF, M., TROVATOM, L., VERATELLIH, M.G. & WAGNERJ, S., 2015: Big data analytics for earth sciences: the EarthServer approach. *International Journal of Digital Earth*, 1-27.
- BAUMANN, P. & HOLSTEN, S., 2011. A comparative analysis of array models for databases. *Database Theory and Application, Bio-Science and Bio-Technology*. Springer Berlin Heidelberg. 80-89.

- BAUMANN, P. & STRAMERJOHANN, H., 2014. Towards a systematic benchmark for array database systems. In: *Specifying Big Data Benchmarks*, Springer, 94-102.
- BRINKHOFF, T. & KRESSE, W., 2011. Databases. *Springer Handbook of Geographic Information*, Springer Berlin Heidelberg, 11-34.
- ORACLE, 2016. Oracle Spatial and Graph Website. <http://www.oracle.com/technetwork/database/options/spatialandgraph/overview/index.html>, letzter Zugriff 29.04.2016.
- LÄNGKVIST, M., KISLEV, A., ALIREZAIE, M. & LOUTFI, A., 2016: Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks. *Remote Sensing* **8** (4).
- POSTGIS, 2016. PostGIS Website. <http://postgis.org/>, letzter Zugriff 29.04.2016.
- PURSS, M.B.J., LEVIS, A., SIMON, O., IP, A., SIXSMITH, J., EVANS, B., EDBERG, R., FRANKISH, G., HURST, L. & CHAN, T., 2015: Unlocking the Australian Landsat Archive – From dark data to High Performance Data infrastructures. *GeoResJ* **6**, 135-140.
- RASDAMAN, 2016. Rasdaman Website. <http://www.rasdaman.org/>, letzter Zugriff 29.04.2016.
- RECK, C., CAMPUZANO, G., DENGLER, K., HEINEN, T. & WINKLER, M., 2016: German Copernicus Data Access and Exploitation Collaborative Infrastructure. *Proceedings of the 2016 conference on Big Data from Space*, Soille, P. & Marchetti, P.G. (Hrsg.), 52-55.
- SCI DB, 2016. SciDB Website. scidb.org/, letzter Zugriff 29.04.2016.
- SOILLE, P., BURGER, A., RODRIGUEZ, D., SYRRIS, V. & VASILEV, V., 2016: Towards a JRC Earth Observation Data and Processing Platform. *Proceedings of the 2016 conference on Big Data from Space*, Soille, P. & Marchetti, P.G. (Hrsg.), 65-68.
- TIEDE, D., SUDMANN, M., BARALDI, A., BELGIU, M. & LANG, S., 2016: ImageQuerying – Prototypische Implementierung eines Systems zur raum-zeitlichen und inhaltsbasierten Satellitenbildsuche und Satellitenbildanalyse. *AGIT - Journal für Angewandte Geoinformatik* [im Druck].

Automated Quality Control Procedures for the International Soil Moisture Network

ANGELIKA XAVER¹, WOUTER DORIGO¹ & WOLFGANG WAGNER¹

Zusammenfassung: In situ Bodenfeuchtebeobachtungen sind essentiell, um modellierte und fernerkundungsbasierte Bodenfeuchteprodukte zu evaluieren und zu kalibrieren. Daher sind aussagekräftige Qualitätsmaße für in situ Bodenfeuchtemessungen von höchster Relevanz. Basierend auf den Daten des International Soil Moisture Network (ISMN), werden komplexe automatisierte Verfahren der Qualitätskontrolle präsentiert, um Ausreißer, Sprünge und Plateaus mithilfe der Form von Bodenfeuchte-Zeitreihen zu detektieren. Mehrere Bedingungen werden definiert, um diese fehlerhaften Ereignisse durch Untersuchung der ersten und zweiten Ableitungen, berechnet mit dem bekannten Savitzky-Golay Filter, zu identifizieren. Die Performance der vorgestellten Qualitätskontrolle wird durch einen Vergleich der automatisiert detektierten Resultate mit manuell markierten Daten, basierend auf einer Auswahl von 40 Bodenfeuchtedatensätzen des ISMN, evaluiert.

1 Introduction

Satellite-derived soil moisture products represent the most promising source of global and long-term soil moisture data availability. However, in situ soil moisture measurements are still crucial for evaluating and calibrating satellite-derived and model-based soil moisture products. Despite being measured directly in the soil, in situ measurement time series typically contain a large number of artefacts such as outliers and breakpoints. Thus, knowledge about the quality of in situ soil moisture measurements is fundamental when using them to assess the reliability of any satellite or model soil moisture product (e.g. DORIGO et al. 2015; ALBERGEL et al. 2012).

This paper presents novel automated quality control procedures for in situ soil moisture time series. The plausibility of soil moisture measurements is evaluated by examining the shape of soil moisture time series by means of its first and second derivative, calculated by the widely-known Savitzky-Golay filter. The advanced automated time series-based methods have been developed and tested using soil moisture data from the International Soil Moisture Network (ISMN) (DORIGO et al. 2011a,b), the largest data bank of in situ soil moisture data worldwide.

2 Data

2.1 International Soil Moisture Network

The International Soil Moisture Network (ISMN; <http://ismn.geo.tuwien.ac.at/>; DORIGO et al. 2011a,b), has been developed and operated by the Department of Geodesy and Geoinformation (GEO) of the TU Wien since 2010. The ISMN acts as a data repository for ground-based soil

¹ Technische Universität Wien, Department Geodäsie und Geoinformation, Forschungsgruppen Photogrammetrie und Fernerkundung, Gußhausstraße 27-29, A-1040 Wien, E-Mail: [Angelika.Xaver, Wouter.Dorigo, Wolfgang.Wagner]@geo.tuwien.ac.at

moisture measurements, with the primary goal of providing reference data for calibrating and validating remotely-sensed soil moisture missions (e.g. SMOS, SMAP). The ISMN collects soil moisture observations from various networks distributed all over the globe. After running through a fully automated processing chain, where the observations are harmonized in terms of unit, temporal resolution and data format, the data is stored in a database and becomes available to the public through a web portal.

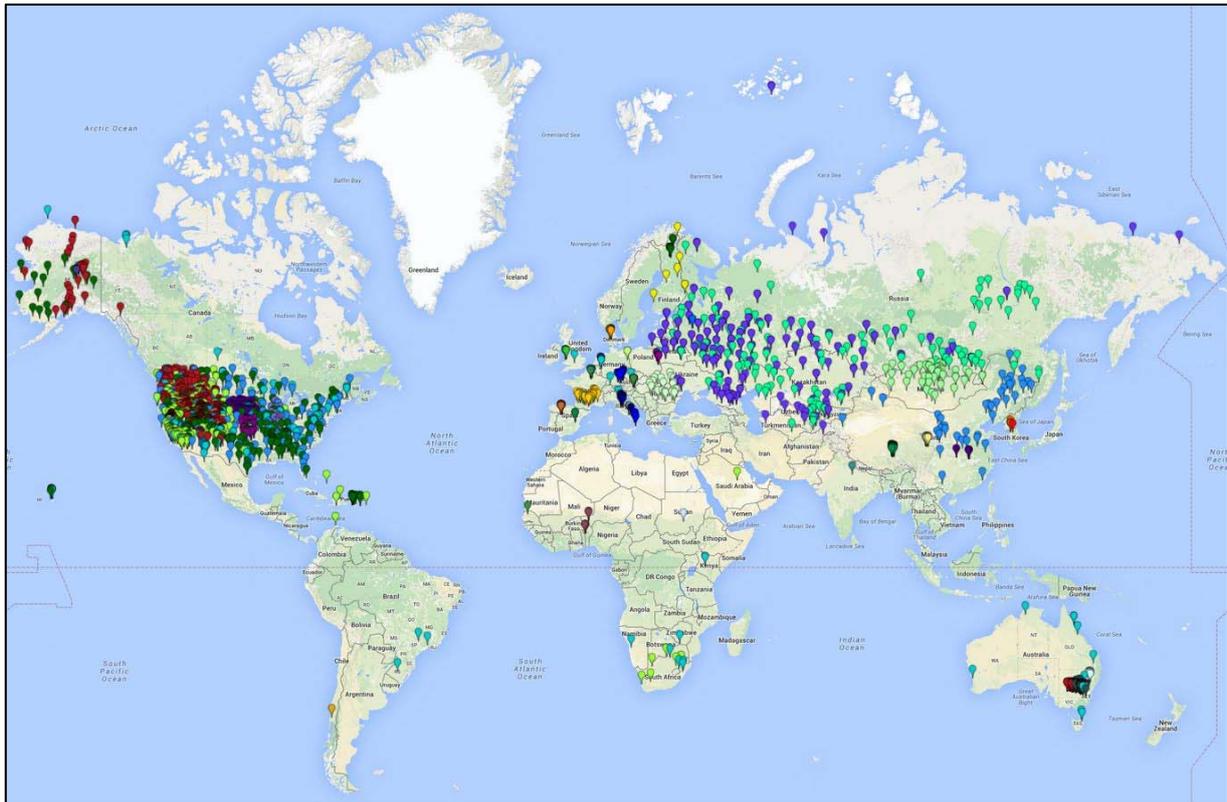


Fig. 1: Stations available at the online platform of the ISMN (status April 2016)

Currently, the ISMN stores the data of 53 networks consisting of 2144 stations (Figure 1). While most datasets are updated on an irregular basis, a few networks provide their data in near-real-time (NRT). These datasets are automatically downloaded, processed and written into the database. Due to the large amount of data processed with each update it is infeasible to perform an operator-based manual quality control of the data. Thus, automated quality control procedures within the ISMN are important in order to provide reliable soil moisture measurements to its users. Furthermore, not every network applies quality control procedures to their data and if quality control mechanisms exist, they are not consistent. Thus, the importance of harmonized quality control procedures, which can be applied to all networks consistently, is evident.

2.2 Soil moisture characteristics

The difficulty in defining quality principles for in situ soil moisture measurements lies in the characteristic shape of a soil moisture time series, which is represented by alternating phases of

wetting and drying (HILLEL 1998). Precipitation events may result in a severe rise of soil moisture within only one or a few hours, while the drying process of soil proceeds slowly, resulting in an inverse exponential shape (Figure 2). These specific properties inhibit the use of typical outlier detection algorithms, which would flag most of the natural rises of soil moisture caused by rain events as erroneous.

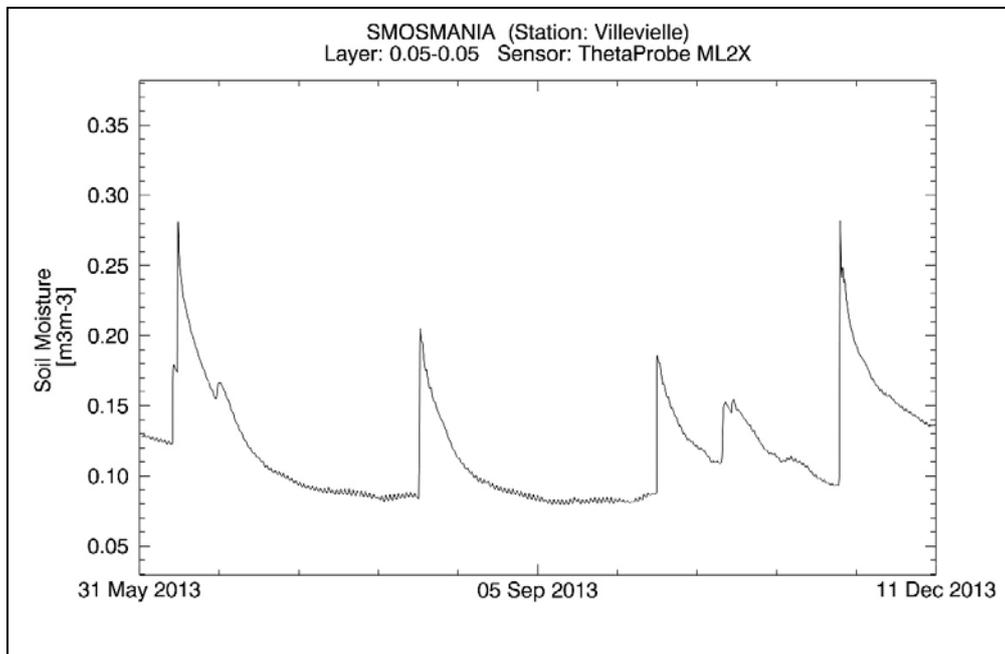


Fig. 2: Example of characteristic wetting and drying phases of an in situ soil moisture time series

Furthermore, the identification of outliers is complicated by the complexity of the signal, e.g. consecutive precipitation events may result in overlapping rises of soil moisture, variations induced by daily temperature cycles may exist (DORIGO et al. 2011b; ROBINSON et al. 2008), or random noise can be present within the soil moisture observations.

In fact, soil moisture regimes may vary strongly depending on the prevalent climate, vegetation and soil composition. Ideally, quality control algorithms should be able to cope with these unavoidable natural and artificial phenomena. In reality, a trade-off has to be made to identify suspicious measurements without over-flagging natural events.

2.3 Potential erroneous events within soil moisture readings

A variety of suspicious measurements can occur caused by malfunction or irresponsiveness of the sensor, or a connection problem while writing on the data logger. The resulting erroneous shapes within the soil moisture time series can be generalized into three categories:

- Spikes: A “spike” is as an outlier of a single measurement which lies significantly above or beneath the preceding and succeeding measured soil moisture values (Figure 3, top left). Thus, a “spike” lasts only for a single unit of time and can appear in both positive and negative directions.

- Breaks: A “break” is a sudden, from one timestamp to the next, rise (“positive break”) or fall (“negative break”) in the measurement time series (Figure 3, bottom left). A sudden drop of the level of soil moisture cannot occur in reality, whereas an instant rise in the amount of soil moisture can either have a natural cause, i.e. as the result of a precipitation event, or be an artefact. Problematic is the fact that breaks may lead to a general offset (bias) of the soil moisture signal.
- Constant Values: Two different kinds of constant values may occur. First, constant values can occur if consecutive precipitation events result in high soil moisture readings, where the sensor is not able to represent any further variation of water in the soil beyond that level. This is referred to as “saturated plateau” (Figure 3, top right). Second, constant values may occur after a sudden drop (“negative break”) of soil moisture, for instance caused by an energy supply problem or when the soil is frozen. This event will be referred to as “low level plateau” (Figure 3, bottom right).

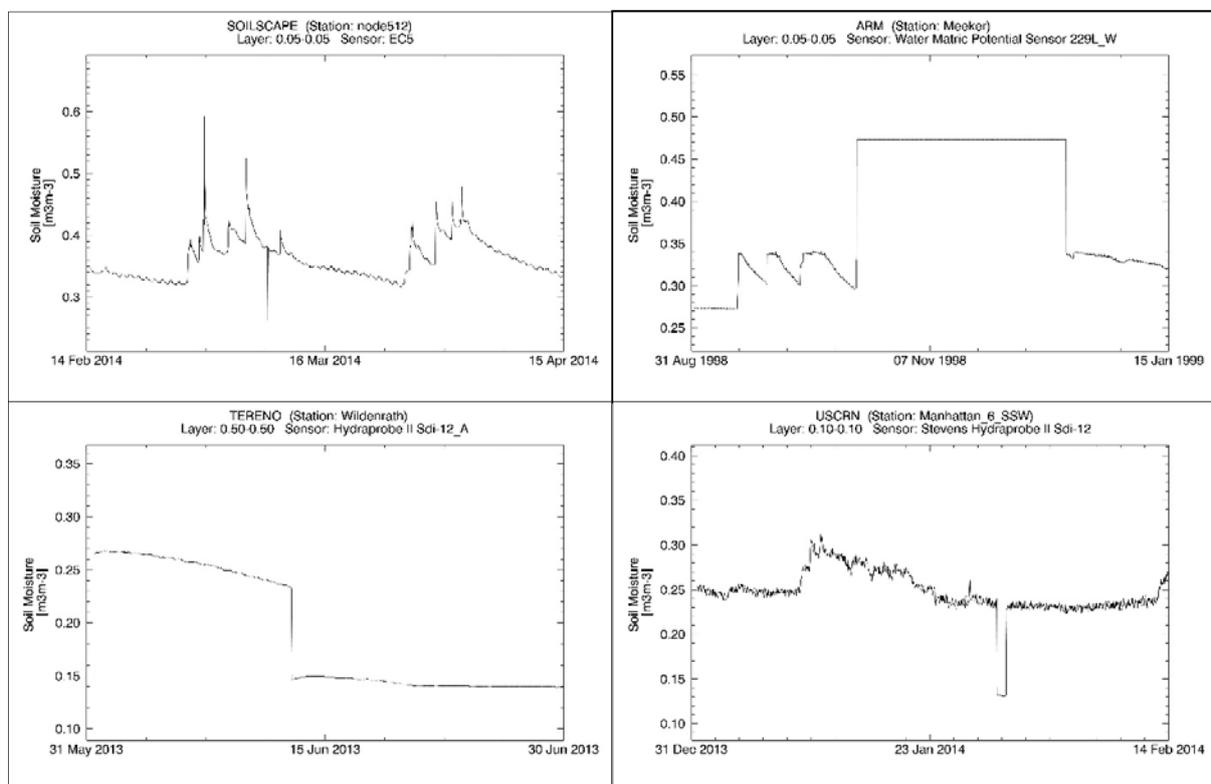


Fig. 3: Examples of erroneous events within soil moisture time series. From top left to bottom right: (negative) spike, saturated plateaus, negative break and low level plateaus.

3 Methods

3.1 Savitzky-Golay Filter

The detection of spurious events is based on the first and second derivative of soil moisture time series, which are calculated by using a Savitzky-Golay filter (SAVITZKY & GOLAY 1964). The advantage of this method is not only the ability to get a smoothing filter and derivatives with the

same equation, but also to preserve higher moments (FLANNERY et al. 1992). Short periods of missing values can be handled well by the Savitzky-Golay filter (EILERS 2003). In addition, the computation is simple, as the ‘smoothing coefficients’ are applied to the data by a convolution, and therefore fast, i.e., suited for NRT processing.

The Savitzky-Golay filter was parametrized with a small symmetrical filter width of three data points and a second degree polynomial.

3.2 Detection Algorithms

Based on the characteristic appearance of each type of suspicious events within the first and second derivative the following conditions were developed through empirical investigation. More details, i.e. all equations, can be found in DORIGO et al. (2013) and XAVER (2015).

3.2.1 Spikes

A spike is detected if the following conditions are fulfilled:

1. The soil moisture signal increases or decreases by at least 15%, which corresponds to approximately three times a typical maximum sensor uncertainty.
2. A spike within the soil moisture time series results in a spike in the second derivative surrounded by smaller peaks in opposite sign (Figure 4, bottom). These surrounding peaks have to be of similar size, thus their ratio is approximately one.

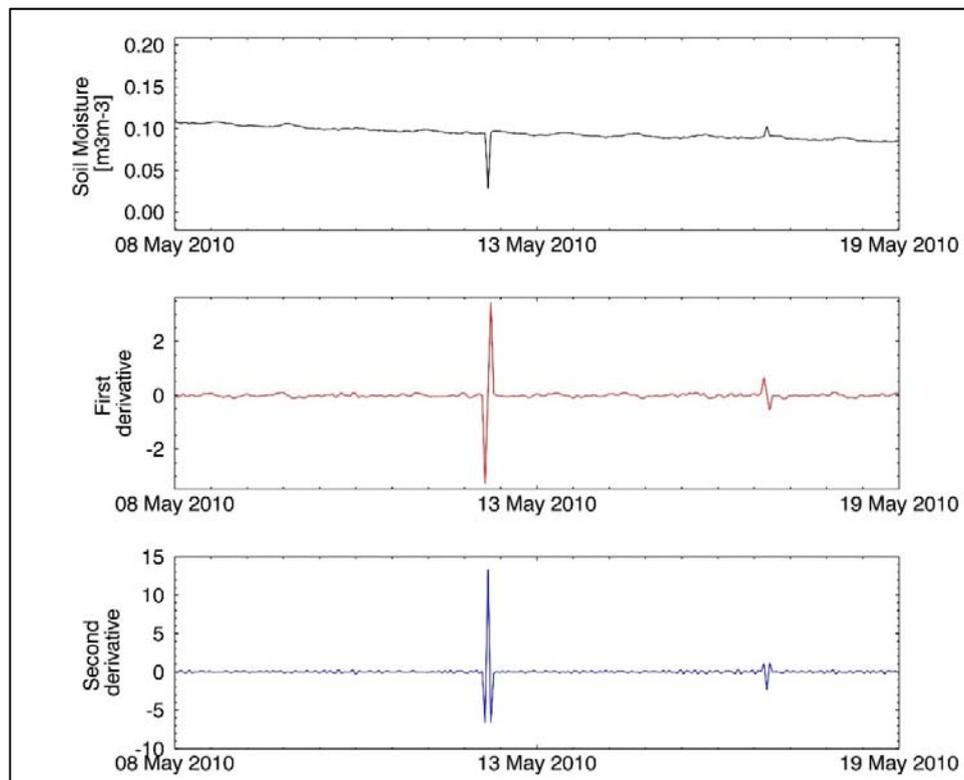


Fig. 4: Shape of first (middle) and second derivative (bottom) for a spike within the soil moisture time series (top) for Station Eulo of network OZNET measured with a Stevens Hydra Probe in a depth from 0.00 to 0.05m

3. The coefficient of variation is applied to the soil moisture time series to an interval of 12 hours before and 12 hours after the potential spike, but excluding the measurement of the spike itself. Its value has to be below one. With this condition an over-flagging in the case of noisy data is avoided.

3.2.2 Breaks

A break is detected if the following conditions are fulfilled:

1. The relative change of soil moisture with respect to the previous time step has to be at least 10%, whereas the absolute amount of change has to be at least $0.01\text{m}^3\text{m}^{-3}$ to avoid over-flagging.
2. A characteristic strong negative (positive) change within the first derivative can be observed when a negative (positive) break appears in the soil moisture time series. This first derivative jump has to be at least ten times smaller (larger) than the arithmetic mean of a 25-hours period of the first derivative centered at the potential break.
3. A negative (positive) break leads to a pronounced negative (positive) second derivative value at the time step before the break, followed by a large positive (negative) value a single time step after the break. These two values have approximately the same magnitude, thus their ratio must be close to unity.

3.2.3 Plateaus

Soil moisture observations are flagged as saturated plateaus if the following conditions are fulfilled:

1. The variation of soil moisture readings over a period of at least 12 hours does not exceed 1% of a typical average sensor uncertainty of $0.05\text{m}^3\text{m}^{-3}$.
2. The characteristic strong rise of soil moisture at the beginning and the sudden drop at the end of the saturated plateau are represented in the first derivative through a sharp positive peak at the beginning, and a strong negative peak at the end of the plateau. Those peaks have to exceed certain threshold values (DORIGO et al. 2013, XAVER 2015).
3. Saturated plateaus consist of the highest recorded soil moisture values of the time series, thus, the level of soil moisture during the plateau has to be at least 95% of the maximum value observed during the whole measurement period.

Soil moisture observations are flagged as low level plateau if the following conditions are fulfilled:

1. A negative break, as described above, is detected.
2. As long as the coefficient of variation stays below 0.01 the measurements following the negative break are identified as low level plateau.

4 Results and Discussion

The performance of the automated quality control procedures was evaluated by comparing the flags obtained from the proposed quality control procedures to a visual classification of occurring erroneous measurements. For this purpose, a selection of 40 datasets from 19 different networks

was drawn from the ISMN and inspected visually. The result of the flagging performance is presented in Table 1.

The overall percentage of flagged data is less than 10% of all readings. As some of the erroneous events represent only one time step (e.g. spikes and breaks) this low outcome does not surprise. Depending on the event, the detection accuracy varies between 42 and 92%. The percentage of correct observations detected as erroneous measurements, i.e. the number of false alarms, lies under 2%.

Tab. 1: Results of the flagging performance (all values given in percentage)

Flagging results	Erroneous measurements		Correct measurements		Flagged observations
	„Erroneous“	„Correct“	„Erroneous“	„Correct“	
Spikes	80.0	20.0	0.0	100.0	0.0
Positive breaks	41.5	58.5	0.0	100.0	0.0
Negative breaks	57.4	42.5	0.0	100.0	0.0
Low level plateaus	59.6	40.4	1.4	98.6	2.8
Saturated plateaus	92.2	7.7	1.6	98.4	6.3

In general, the quality control algorithms are working very well (see Figure 5 for examples). Reasons for the failure of the flagging procedures can be summarized as follows:

- Existence of missing values: Single missing values can be overcome by the introduced algorithms, but multiple missing values or whole periods of missing data lead to a disturbance of the Savitzky-Golay filter and the characteristic shapes in the first and second derivatives.
- Imperfect definition: Some thresholds had to be defined for the different quality control procedures, e.g. for identifying saturated plateaus, to detect as many erroneous events as possible while avoiding over-flagging of correct data. Of course, real data does not always comply with these thresholds and thus some erroneous measurements cannot be detected.
- Noisy data: Noisy data causes also noisy first and second derivatives, resulting from the applied parameterization of the Savitzky-Golay filter. Therefore, the characteristics of the two derivatives which are used to identify the different suspicious measurements are superimposed by the random noise.
- Consecutive erroneous events: The effect of consecutive erroneous events is similar to that of noisy data. The characteristic shape of the derivatives is disturbed and the defined methods fail to detect erroneous measurements.

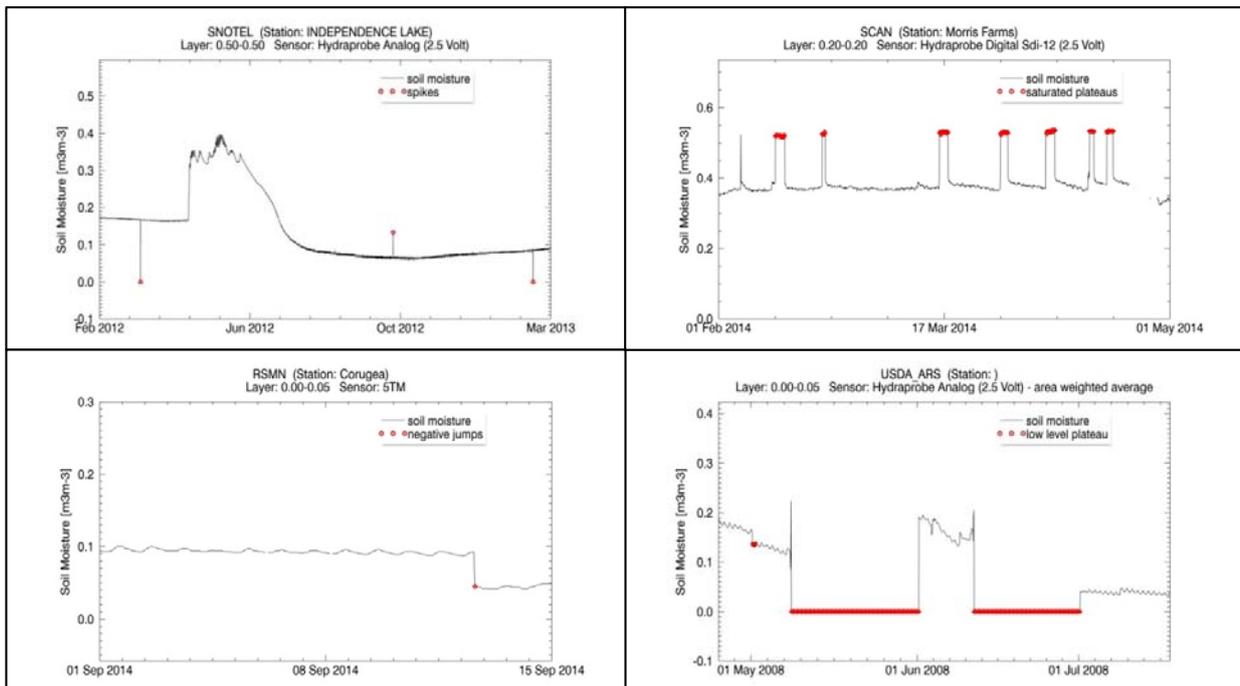


Fig. 5: Examples of flagging results obtained by the automated quality control procedures. From top left to bottom right: spikes, saturated plateaus, negative break and low level plateaus.

5 Conclusions

Automated quality control procedures cannot be expected to work perfectly. Nevertheless, an overall good percentage of erroneous measurements could be identified by the introduced methods. Moreover, the resulting number of false alarms was extremely low. Even if not all erroneous observations can be detected, the spectrum-based quality control algorithms act as a clear indicator for the quality of a soil moisture time series.

The herein described quality control procedures are implemented within the ISMN processing chain since spring 2014. For each case of the described quality detection methods (i.e., spikes, breaks, and plateaus) separate flags are defined. The flags are attached to the soil moisture measurements and provided to all users of the ISMN in addition to the actual measurements, which themselves remain unchanged. Thus, the quality control procedures will help to improve the reliability of validation studies based on in situ soil moisture observations on the one hand, and will help to identify problems which may occur at the measurement sites on the other hand.

6 References

This paper is based on the diploma thesis 'Automated Quality Control Procedures for the International Soil Moisture Network' (XAVER 2015).

- ALBERGEL, C., DE ROSNAY, P., GRUIER, C., MUNOZ-SABATER, J., HASENAUER, S., ISAKSEN, L., KERR, Y. & WAGNER, W., 2012: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. *Remote Sensing of Environment* **118**, 215-226.
- DORIGO, W., GRUBER, A., DE JEU, R., WAGNER, W., STACKE, T., LOEW, A., ALBERGEL, C., BROCCA, L., CHUNG, D., PARINUSSA, R. & KIDD, R., 2015: Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sensing of Environment* **162**, 380-395.
- DORIGO, W., OEVELEN, P. v., WAGNER, W., DRUSCH, M., MECKLENBURG, S., ROBOCK, A. & JACKSON, T., 2011a: A new international network for in situ soil moisture data. *Eos Transactions AGU* **92** (17), 141-142.
- DORIGO, W., WAGNER, W., HOHENSINN, R., HAHN, S., PAULIK, C., XAVER, A., GRUBER, A., DRUSCH, M., MECKLENBURG, S., VAN OEVELEN, P., ROBOCK, A. & JACKSON, R., 2011b: The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences* **15** (5), 1675-1698.
- DORIGO, W., XAVER, A., VREUGDENHIL, M., GRUBER, A., HEGYIOVÁ, A., SANCHIS-DUFAU, A.D., ZAMOJSKI, D., CORDES, C., WAGNER, W. & DRUSCH, M., 2013: Global Automated Quality Control of In situ Soil Moisture data from the International Soil Moisture Network. *Vadose Zone Journal* **12**, 3.
- EILERS, P., 2003: A Perfect Smoother. *Analytical Chemistry* (75), 3631-3636.
- HILLEL, D., 1998: *Environmental soil physics*. Academic Press.
- FLANNERY, B., TEUKOLSKY, S. & VETTERLING, W., 1992: *Savitzky-Golay Smoothing Filters. Numerical Recipes in C: The art of scientific computing*. Cambridge University Press.
- ROBINSON, D., CAMPBELL, C., HOPMANS, J., HORNBUCKLE, B., JONES, S., KNIGHT, R., OGDEN, F., SELKER, J. & WENDROTH, O., 2008: Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review. *Vadose Zone Journal* **7**, 358-389.
- SAVITZKY, A. & GOLAY, M.J.E., 1964: Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* **36**, 1627-1639.
- XAVER, A. 2015: *Automated Quality Control Procedures for the International Soil Moisture Network*. Diploma thesis, TU Wien, Austria.

Superweitwinkel- und Fisheye-Kameras in photogrammetrischen Anwendungen

MUYAN XIAO¹ & MICHAEL CRAMER¹

Zusammenfassung: Im Rahmen der Arbeit soll untersucht werden, in wie weit sich das klassische perspektivische Kalibriermodell (mit physikalischer Parameter von D. Brown) auf superweitwinklige bzw. Fisheye-Kameras übertragen lässt bzw. modifizierte Parametersätze notwendig sind. Dafür wird das Kalibriermodell mathematisch simuliert und der Einfluss der Kalibrierungen an empirischen Datensätzen untersucht. Aus dem Ergebnis der Simulation lässt sich feststellen, dass sich der Unterschied zwischen dem erweiterten perspektivischen Modell und einer idealen Fisheye-Projektion mit steigendem Einfallswinkel des Abbildungsstrahls vom Objektpunkt vergrößert. Bis zu einem maximalen Einfallswinkel von etwa 60° zeigt die durchgeführte Simulation eine Abweichung beider Modelle kleiner 1 pix. Das erweiterte klassische perspektivische Kalibriermodell reicht also nicht aus um die Bilder mit großen Bildwinkeln wie bei Fisheye-Objektiven komplett auszuwerten. In der empirischen Untersuchung werden drei Fisheye-Kamerasysteme und eine normale weitwinklige Kamera verwendet. Der Unterschied von äquidistantem bzw. erweitertem perspektivischem Kalibriermodell wird anhand verschiedener Datensätze einer Laborszene untersucht.

1 Einleitung

Fisheye-Optiken sind Objektive, die extrem kurze Brennweite besitzen und sehr große Öffnungswinkel bis 180° oder sogar mehr realisieren können. Heutzutage finden sich solche Kamerasysteme im Massenmarkt überwiegend in Freizeitanwendungen. Aufgrund des großen Sichtfelds und der Realisierung von kurzen Aufnahmedistanzen sind Fisheye-Objektive im Sportbereich anzutreffen. Als Beispiel soll die „Action Cam“ Hero GoPro genannt werden. Fisheye-Kameras finden sich aber auch zunehmend im photogrammetrischen Bereich. Zunächst waren sie eher im Bereich der Innenraumvermessung bzw. Robotik zu finden. Mittlerweile gibt es erste UAV-Systeme mit Fisheye-Kameras z.B. für die Erzeugung von 3D-Punktwolken oder zur (Bauwerks-)Überwachung. Ein weiterer Anwendungsbereich ist das fahrzeuggestützte Mobile-Mapping (BEERS 2011), vor allem auch für die Modellierung von Straßenszenen. Bei Verwendung von Fisheye-Kameras werden im Vergleich zu normalen Kameras viel weniger Bilder benötigt, um den gleichen Anteil der Überlappung zwischen den Bildern zu bekommen (KEDZIERSKI et al. 2009).

Die starke Verzeichnung derartiger Optiken erlaubt in der Regel nicht die Verwendung von herkömmlichen Parametermodellen zur Modellierung der Kamerageometrie. Viele Forschungsarbeiten bieten daher verschiedene Methode an, um Fisheye-Optiken zu kalibrieren. In einigen Untersuchungen werden bestimmte Eigenschaften aus dem Lochkamera-Modell als Bedingungen zur Rektifizierung der Bilder von Fisheye-Kameras eingesetzt (ABRAHAM &

¹ Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, E-Mail: xiaomy92@gmail.com, michael.cramer@ifp.uni-stuttgart.de

FÖRSTNER 2005). Manche Arbeiten versuchen ein generelles Kalibriermodell für verschiedene Abbildungsgeometrien einschließlich Fisheye-Linsen zu erstellen. KANNALA & BRANDT (2006) empfehlen ein Modell basierend auf einem Polynom neunten Grades für alle Arten von Fisheye-Kameras. GENNERY (2006) formuliert ein allgemeines Linsenmodell, das unterschiedliche Abbildungsgeometrie zusammen kombiniert, wobei ein Parameter das spezielle Kameramodell bestimmt. Nach GROSSBERG & NAYAR (2005) werden Abbildungsgeometrien einfach durch eine Gruppe von Elementen beschrieben, die geometrische, radiometrische und optische Eigenschaften bezeichnet.

2 Abbildungsgeometrie der Fisheye-Objektive

2.1 Projektionsart der Fisheye-Optiken

Je nach Projektionsart können Fisheye-Objektive mit äquidistanter, flächentreuer, orthographischer und stereo-graphischer Projektion klassifiziert werden (nach PRENZEL 1986; RAY 2002). Die meisten Fisheye-Optiken sind entsprechend der äquidistanten Projektionsart konstruiert. Im Folgenden wird diese Abbildungsgeometrie im Vergleich zur perspektivischen Abbildung vorgestellt (nach SCHNEIDER et al. 2009; ABRAHAM & FÖRSTNER 2005).

Der Abbildungsvorgang der äquidistanten Projektion wird in Abb. 1 dargestellt, wobei die Volllinie den entsprechenden äquidistanten Abbildungsstrahl zeigt und die Strichlinie die Zentralprojektion bezeichnet. Die Gleichungen (1) und (3) beschreiben die jeweils zugehörige Beziehung zwischen dem Einfallswinkel (α) eines Objekts und der radialen Distanz (r) des abgebildeten Punkts, wobei c die Kamerakonstante ist. Aus der Gleichung (3) wird der Einfallswinkel eines Abbildungsstrahls zu radialen Distanz linear transformiert. Deswegen wird ein Objektpunkt mit einem Einfallswinkel von 90° auf dem Rand des Bildkreises ($R = c \cdot \pi/2$) abgebildet, während der gleiche Punkt in der perspektivischen Abbildung unendlich weit projiziert wird. Die Gleichungen (2) und (4) sind die Kameramodelle beider Abbildungsarten, mit denen der Bezug zwischen Bildkoordinaten (x, y) und Objektkoordinaten (X, Y, Z) hergestellt wird.

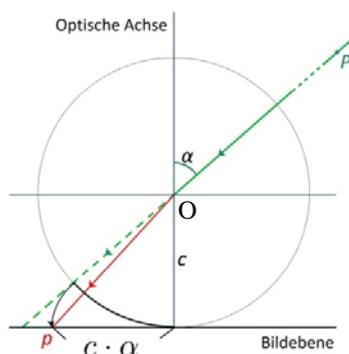


Abb. 1: Äquidistante und perspektive Projektion

Perspektivische Projektion:

$$r = c \cdot \tan \alpha \quad (1)$$

$$x = c \cdot \frac{X}{Z} \quad y = c \cdot \frac{Y}{Z} \quad (2)$$

Äquidistante Projektion:

$$r = c \cdot \alpha \quad (3)$$

$$x = c \cdot \frac{X}{\sqrt{X^2+Y^2}} \cdot \arctan \frac{\sqrt{X^2+Y^2}}{Z}$$

$$y = c \cdot \frac{Y}{\sqrt{X^2+Y^2}} \cdot \arctan \frac{\sqrt{X^2+Y^2}}{Z} \quad (4)$$

2.2 Physikalisches Verzeichnungsmodell

Das Verzeichnungsmodell von Brown definiert physikalisch relevante Parameter, um die Abweichung zwischen Abbildungsrealität und theoretischem Kameramodell zu kompensieren.

Diese Parameter beschreiben sowohl radial-symmetrische- und tangentielle Verzeichnung als auch Unebenheit und in-plane Verzerrung (BROWN 1971). Die Anwendung dieser Parameter in der geometrischen Kamerakalibrierung ist bekannt. In dieser Arbeit wird das perspektivische Kameramodell mit den Brown-Verzeichnungsparametern erweitert und mit dem äquidistanten Modell verglichen. Parallel kann das Verzeichnungsmodell von Brown auch für die Erweiterung von geometrischen Modellen von Fisheye-Optiken verwendet werden (SCHNEIDER et al. 2009). Entsprechend wird das physikalische Verzeichnungsmodell sowohl für die Zentralprojektion als auch die Fisheye-Abbildung in den später in den empirischen Untersuchungen eingesetzten Auswerteprogrammen implementiert.

3 Simulation

Um die Kameramodelle mathematisch darzustellen und das Potenzial des erweiterten perspektivischen Modells für die Fisheye-Geometrie ohne Einfluss von möglichen realen Fehlerquellen zu simulieren, wird folgende Simulation vorgenommen: Eine simulierte Fisheye-Kamera macht aus unterschiedlichem Abstand Nadir-Aufnahmen einer mit bekannten Punkten versehenen Fläche (Abb. 2).

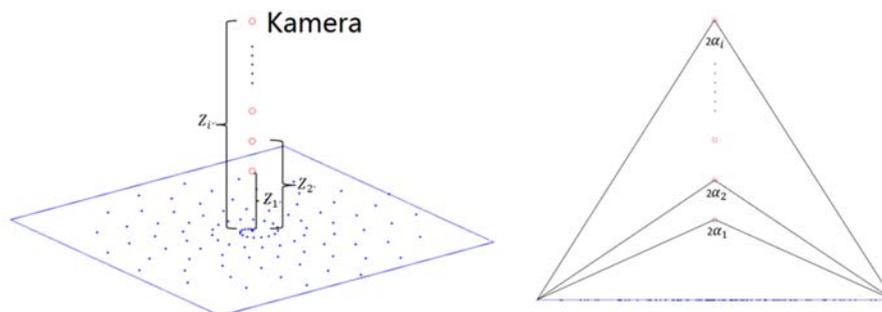


Abb. 2: Simulationsszene

Die Fläche hat eine Ausdehnung von $60\text{m} \times 60\text{m}$ und enthält 100 bekannte Punkte. Die Koordinaten (X, Y, Z) der Punkte im Objektraum werden mit Ursprung im Projektionszentrum des Bildes simuliert. Wenn sich die Distanz (Z) zwischen dem Aufnahmestandpunkt und der Fläche vergrößert, wird der zugehörige Einfallswinkel (α) repräsentiert durch die Objektpunkte kleiner (siehe Abb. 2, rechts). Für die Simulation werden zwei unterschiedliche Kameras angenommen, die in ihrer Geometrie ähnlich den später in den empirischen Untersuchungen eingesetzten Kameras sind: Kamera 1 hat eine Kamerakonstante von 1.25 mm , eine Pixelgröße von $3.75\text{ }\mu\text{m}$, und Kamera 2 entsprechend 5 mm und $1.41\text{ }\mu\text{m}$.

Mit den vorgegebenen Aufnahmestandpunkten lassen sich nun die entsprechenden Bildkoordinaten der Objektpunkte mit dem äquidistanten Modell (Gl. 4) berechnen. Parallel wird in der Simulation versucht, die Objektpunkte auch mit dem um die radiale Verzeichnung erweiterten perspektivischen Modell zu berechnen. Dabei wurden sechs Verzeichnungsparameter $(k_1, k_2, k_3, k_4, k_5, k_6)$ verwendet – i.d.R. werden nur die Parameter $k_1 - k_4$ eingesetzt. Die

mittlere Abweichung der simulierten Bildkoordinaten in Abhängigkeit von den unterschiedlichen maximalen Einfallswinkeln ist für beide Abbildungsmodelle logarithmisch in Abb. 3 dargestellt.

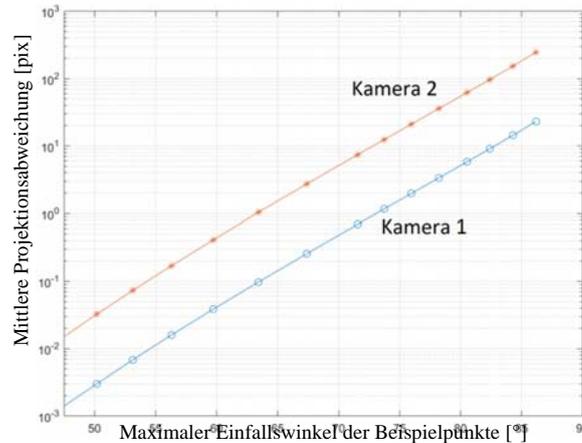


Abb. 3: Abweichung der Bildkoordinaten beider Abbildungsmodelle (in [pix])

Die Simulation zeigt, dass die Abweichung zwischen dem äquidistanten Modell und dem erweiterten perspektivischen Modell mit steigendem Einfallswinkel immer größer wird. Außerdem ist die Größe der Abweichung abhängig von der gewählten Kamerageometrie. Für die Kamera 1 liegt die Abweichung bei einem maximalen Einfallswinkel von etwa 70° noch im sub-pixel Bereich. Für die Kamera 2 ist die Abweichung kleiner als 1 pix, wenn der max. Einfallswinkel kleiner gleich etwa 63° ist. Damit können Fisheye-Kameras (je nach Geometrie) bis zu diesen Winkeln noch mit einem erweiterten perspektivischen Abbildungsmodell sub-pixel genau ausgewertet werden. Für größere Bildwinkel reicht die Anwendung des erweiterten perspektivischen Modells nicht aus. Deswegen ist es wichtig, das korrekte Abbildungsmodell für Fisheye-Kameras für genaue photogrammetrische Anwendungen zu benutzen.

4 Empirische Untersuchung

Für die empirische Untersuchung werden Bildverbände einer Laborraumszene erfasst. Zentrum dieser Szene ist eine Kalibrierwand (Schrankfläche), die mit gleichmäßig verteilten kodierten Zielmarken beklebt ist. Die Ausdehnung dieser Kalibrierwand beträgt etwa 3 m×2.45 m (B×H). Zur Anwendungen kommen 3 unterschiedliche Fisheye-Optiken um die Bandbreite vorhandener Systeme gut abzudecken. Es handelt sich dabei um eine (Industrie-)Kamera mit einem Zirkular-Fisheye (Chameleon 1.3 MP Mono USB 2.0 (Sony ICK445) mit Lensagon BF2M12520 Objektiv). Außerdem wird das UAV-System DJI Phantom Vision mit der Kamera FC 200 untersucht. Die dritte Konfiguration ist eine vollformatige Spiegelreflexkamera (Nikon D800E) mit einem 16 mm Fisheye-Objektiv. Diese Kamera wird auch in Kombination mit einem klassischen 20 mm Weitwinkel-Objektiv untersucht. Dieses repräsentiert die eher klassische Aufnahme Konfiguration, die dann als Referenzlösung für die Fisheye-Objektive dient.

Ähnlich der Simulation werden Bildverbände aus unterschiedlicher Distanz (ca. 1 m, 2 m, 3 m Aufnahmeabstand) erfasst. Damit ergeben sich unterschiedliche maximale Einfallswinkel.

Jeder Bildverband von einem Fisheye wird mit sowohl dem perspektivischen als auch dem äquidistanten Kalibriermodell mit den bekannten Softwarepaketen Agisoft PhotoScan bzw. Pix4D Pix4Dmapper ausgewertet. Nach der Bildorientierung wird eine dichte Punktwolke für jeden Bildverband generiert. Im Folgenden werden die Ergebnisse der Auswertungen dargestellt und diskutiert.

Tab. 1 zeigt die mittlere Verbesserung (Rückprojektionsfehler) an den kodierten Marken nach der Orientierung jedes Bildverbands. Tab. 2 zeigt das Rauschen der Punktwolke.

Tab. 1: Mittlere Verbesserung (Rückprojektionsfehler) der kodierten Marken

Bildverband	Einfallswinkel (max.)	Fisheye-Modell [pix]		Perspektivisches Modell [pix]	
		PhotoScan	Pix4Dmapper	PhotoScan	Pix4Dmapper
Chameleon 1m	70°	0.389	0.468	7.420	3.209
Chameleon 2m	47°	0.148	0.292	2.511	1.504
Chameleon 3m	32.5°	0.162	0.231	1.347	0.743
Nikon-Fisheye 1m	62.5°	0.136	—	58.173	—
Nikon-Fisheye 2m	42.5°	0.151	—	5.553	—
DJI 1m	52°	1.128	0.660	2.042	0.962
DJI 2m	33°	0.453	0.353	0.922	0.640
Nikon-Weitwinkel	47°	—	—	0.110	—

Tab. 2: Rauschen (Abweichung von ausgleichender Ebene) der 3D-Punktwolke (relative Genauigkeit)

Bildverband	Einfallswinkel (max.)	Fisheye-Modell [cm]		Perspektivisches Modell [cm]	
		PhotoScan	Pix4Dmapper	PhotoScan	Pix4Dmapper
Chameleon 1m	70°	0.92	1.70	12.57	3.63
Chameleon 2m	47°	0.64	1.51	2.47	3.58
Chameleon 3m	32.5°	0.73	1.67	0.90	2.20
Nikon-Fisheye 1m	62.5°	0.34	—	9.38	—
Nikon-Fisheye 2m	42.5°	0.42	—	0.93	—
DJI 1m	52°	1.46	1.70	2.44	1.98
DJI 2m	33°	1.26	1.51	1.36	1.55
Nikon-Weitwinkel	47°	—	—	0.42	—

Die Genauigkeit der Punktwolke wird durch die Analyse des Rauschens bzw. der Abweichung von einer die Kalibrierungswand ausgleichenden Ebene vorgenommen (relative Genauigkeit). Ein Beispiel dieser relativen Genauigkeitsanalyse wird in Abb. 4 gezeigt, wobei die Histogramme der Differenzbildung für den Datensatz Chameleon 1 m dargestellt sind. Oben und unten sind jeweils die Ergebnisse der Punktwolken ausgewertet mit dem Fisheye- und perspektivischen Kalibriermodell. Tab. 2 zeigt die statistische Auswertung (Standardabweichung) der Differenzen der Punktwolke von der vermittelnden Ebene.

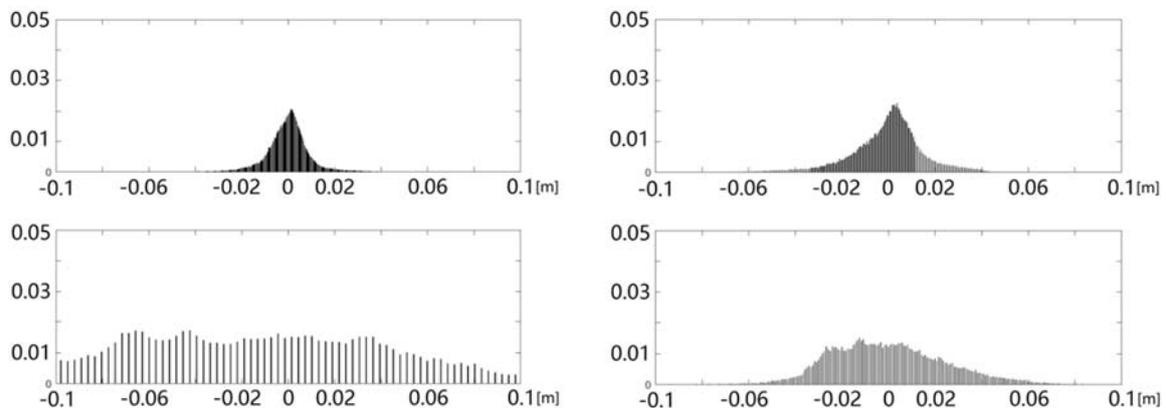


Abb. 4: Häufigkeitsverteilung des Rauschens der 3D-Punktwolke: äquidistantes Modell (oben) bzw. perspektives Modell (unten), Datensatz Chameleon $1m$, Auswertung PhotoScan (links) bzw. Pix4Dmapper (rechts)

Die Anwendbarkeit des klassisch perspektivischen Kameramodells ist eine Funktion vom maximalen Einfallswinkel, wie auch schon in der Simulation gezeigt. Die Auswertungen zeigen, dass der Rückprojektionsfehler der Bündelausgleichung mit dem erweiterten perspektivischen Modell in Abhängigkeit vom max. Einfallswinkel abnimmt. Das gilt auch für die Genauigkeit der Punktwolke. Wenn man das Rauschen der abgeleiteten Punktwolke vergleicht, so scheinen die Differenzen zwischen beiden Abbildungsmodellen für die kleineren max. Einfallswinkel geringer. Der Rückprojektionsfehler für den Nikon-Fisheye $1m$ Datensatz (Perspektivmodell) ist auffallend groß. Das Objektiv hat eine viel größere Kamerakonstante als die anderen Systeme und damit ist die Abweichung größer (siehe auch Simulation).

In allen Fällen liefert das Fisheye-Abbildungsmodell bessere Resultate als das erweiterte Perspektivmodell, wie in Abb. 4 dargestellt: die Histogramme der Differenz der Punktwolke mit dem perspektivischen Kalibriermodell von der vermittelnden Ebene sind viel breiter gestreut als mit dem Fisheye-Kalibriermodell, mit dem fast alle Bildverbände mit einem mittleren Rückprojektionsfehler besser 1 pix ausgewertet werden. Die besten Genauigkeiten werden mit der Nikon-Fisheye-Konfiguration erreicht, was an der deutlich besseren Abbildungsqualität liegt. Der Rückprojektionsfehler (Fisheye) bleibt für beide Nikon-Datensätze praktisch gleich und erreicht die Qualität der Weitwinkel-Referenzlösung. Außerdem wird für die Genauigkeit der Punktwolke mit einer Standardabweichung von ca. 0.4 cm die Qualität der Referenzlösung erreicht.

Wenn man die Ergebnisse der Auswertungen beider Softwarepakete vergleicht, so ist die Genauigkeit der Punktwolke mit dem Fisheye-Kalibriermodell von Pix4Dmapper schlechter als das Ergebnis von PhotoScan. Es könnte daran liegen, dass für das Fisheye-Kalibriermodell in Pix4Dmapper ein anderes Verzeichnungsmodell als die physikalischen Parameter von Brown implementiert sind (HARTLEY & ZISSERMAN 2003). Hinzu kommen Unterschiede in den Auswerteprozessen der beiden Softwarepakete. PhotoScan erlaubt es, die Bündelblockausgleichung zuerst nur mit automatisch extrahierten Verknüpfungspunkten (z.B. Feature-Punkte) zu berechnen und erst danach signalisierter Kontrollpunkte (kodierte Zielmarken) in den Bildern manuell zu messen. So wird die Triangulation nicht von den

manuellen Messungen beeinflusst und die Projektionsfehler an den Zielmarken können bezüglich der schon existierenden Orientierung ermittelt werden. In Pix4Dmapper werden die Rückprojektionsfehler erst dann geliefert, wenn die Triangulation gleichzeitig mit den von der Software automatisch erkannten Verknüpfungspunkten und den manuell gemessenen Zielmarken durchgeführt wird. Aufgrund der Berücksichtigung der manuell gemessenen Punkte im Rahmen der Ausgleichung werden die deren Residuen entsprechend minimiert, sodass die Auswertung mit dem perspektivischen Modell von Pix4Dmapper kleinere Rückprojektionsfehler an den Zielmarken und geringere Abweichung zwischen den Punktwolken und der ausgleichenden Ebene im Vergleich zu den Ergebnissen von PhotoScan liefert.

5 Zusammenfassung

Die theoretischen und empirischen Auswertungen dieser Studie zeigen, dass Fisheye-Bilder bei Verwendung des richtigen Abbildungsmodells die sub-pixel Genauigkeitsanforderung photogrammetrischer Anwendungen erfüllen. Die Untersuchung zeigt aber auch, dass das äquidistante Kalibriermodell die Fisheye-Kameras besser als das (erweiterte) perspektivische Modell mit Verzeichnungsparametern von Brown modelliert. Die Genauigkeit des erweiterten perspektivischen Modells wird mit steigendem max. Einfallswinkel immer schlechter. Da die Realisierung eines großen Sichtfelds die eigentliche Motivation der Fisheye-Objektive ist, kann das klassische Modell daher in den meisten Fällen nicht verwendet werden.

Bezüglich der Simulationen in dieser Arbeit sind beide Abbildungsmodelle jedoch ineinander überführbar, sofern der max. Einfallswinkel kleiner etwa 60° bleibt. Die meisten normalwinkligen Kameras können einen Bildwinkel von 120° nicht erreichen. Deswegen könnte das äquidistante Modell mit entsprechenden Verzeichnungsparametern als ein generelles Kalibriermodell für beiden Kameratypen eingesetzt werden.

Die Unterschiede der mathematischen Modellierung der Fisheye-Geometrie in unterschiedlichen Softwarepaketen werden ebenfalls thematisiert. Pix4Dmapper implementiert dabei ein anderes Verzeichnungsmodell als die physikalische Parametersätze von Brown. Das Abbildungsmodell beeinflusst die Genauigkeit, was der Vergleich mit den Ergebnissen aus PhotoScan zeigt.

6 Literaturverzeichnis

- ABRAHAM, S. & FÖRSTNER, W., 2005: Fish-eye-stereo calibration and epipolar rectification. *ISPRS Journal of Photogrammetry and Remote Sensing* **59**, 278-288.
- BEERS, B., 2011: Collection and application of 2d and 3d panoramic imagery. *Photogrammetric Week '11*, Fritsch, D. (Ed.), Wichmann Verlag, 277-287.
- BROWN, D., 1971: Close-range camera calibration. *Photogram. Eng. Remote Sens* **37**, 855-866.
- GENNERY, D., 2006: Generalized camera calibration including fish-eye lenses. *International Journal of Computer Vision* **68** (3), 239-266.
- GROSSBERG, M. & NAYAR, S., 2005: The raxel imaging model and ray-based calibration. *International Journal of Computer Vision* **61** (2), 119-137.

- HARTLEY, R. & ZISSERMAN, A., 2003: Multiple view geometry in computer vision. Cambridge university press.
- KANNALA, J. & BRANDT, S., 2006: A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28** (8), 1335-1340.
- KEDZIERSKI, M., WALCZYKOWSKI, P. & FRYSKOWSKA, A., 2009: Application of fisheye lens and terrestrial laser scanning in architectonic documentation of hard-to-reach cultural heritage objects. *Proceedings of the ASPRS 2009 Annual Conference*.
- PRENZEL, W., 1986: Entwicklungstendenzen der fotografischen Optik. *Bild und Ton - wissenschaftliche Zeitschrift für visuelle und auditive Medien*.
- RAY, S., 2002: *Applied photographic optics: Lenses and optical systems for photography, film, video, electronic and digital imaging*. Focal Press.
- SCHNEIDER, D., SCHWALBE, E. & MAAS, H.-G., 2009: Validation of geometric models for fisheye lenses. *ISPRS Journal of Photogrammetry and Remote Sensing* **64** (3), 259-266.

Detection of Star-shaped Objects in Aerial Images: Use Case Oil Palm Tree Detection

ABHISHEK MANANDHAR¹, LUDWIG HOEGNER¹ & UWE STILLA¹

Abstract: Object detection is one of the major methods of remote sensing. Generally in object detection features are extracted from the image to represent the object in the image numerically with lesser number of elements than the image itself. Out of multiple features, shape feature is one of the widely used features. In this study, new rotation invariant shape features are introduced and comparative studies are made among them by using them to detect a naturally occurring star-shaped object, palm trees, in aerial images. Detection of the palm trees helps in creating inventory which is an initial step towards monitoring.

This study proposes shape features, the circular autocorrelation of polar shape matrix (CAPS), Mean of CAPS and Entropy of CAPS to be used with object detection framework. The framework used in the study to evaluate features uses a well-known machine learning algorithm, support vector machine (SVM), to detect palm trees in aerial images.

The implementation on aerial images, taken with unmanned air vehicle (UAV) in five plantation regions in Indonesia, Malaysia, and Thailand, shows promising results. The CAPS produced the best results among the features studied, with an average accuracy of 84% over 8 images, chosen considering different challenges from the five plantation region.

1 Introduction

There are multiple examples of star-shaped objects in nature. Especially, in aerial images, objects like palm tree canopy, banana tree canopy, etc. have star shapes. The shape information of those objects can be used in detection and monitoring of those objects from aerial imageries.

Palm trees produce a wide variety of products like vegetable oil, biofuel, furniture etc. Palm vegetable oil is currently the most traded oil. In 2011, 53 million tons of palm vegetable oil was produced, with trees cultivated in 15 million hectares. The demand for the product is increasing and hence, the production is estimated to rise by 30% in Malaysia and Indonesia, the countries that contribute to more than 90% of trade worldwide (POTTS et al. 2014). This demand will lead to larger cultivation and, hence, the necessity for smart monitoring of the plantations. Detection of the palm trees in aerial imageries and keeping inventory can serve as a start.

Object detection and localization in aerial image is one of the major methods of remote sensing. Researchers, government, and non-government institutions use this method to detect and localize objects of interest like buildings, cars, ships, trees etc., which further help in creating inventories and monitoring the status of the objects. The approach has also been used for identifying hazards like wildfires in inaccessible areas.

Generally, in object detection, different features are extracted from an image to represent the object present in the image and those features are used with detection framework. Features are the numerical representation of the objects with relatively fewer elements than in the image itself.

¹ Technische Universität München, Photogrammetry and Remote Sensing, Arcisstrasse 21, D-80333 München, Germany, E-Mail: [Abhishek.Manandhar, Ludwig.Hoegner, stilla]@tum.de

There have been myriad of studies to derive different types of features which can be classified into different categories based on their use and characteristics.

Shape features are among the widely used features in object detection. As the name suggests, shape features are a representation of the shape of the object. As all other features, shape features also have to cope with different challenges, especially induced by different saliencies, rotation, scale, shift, illumination etc. Different features try to address different challenges according to need.

In this study, we derive a new rotation invariant shape feature, as we consider to detect the star-shaped object. Besides, scale and shift invariance can be achieved by using the moving window detection scheme at different scales, therefore, rotation invariance will give feature extra strength. The shape features namely, Circular Autocorrelation of Polar Shape Matrix (CAPS), Mean of CAPS, Entropy of CAPS, etc. will be derived from the star-shaped model and will be used with object detection framework to detect and count palm trees in different plantation regions.

2 Method

The primary objective of the study is to identify plausible shape feature which can be used in object detection framework for detection of the star-shaped object. The framework used in the study (Figure 1) consists of two major processes: Standard feature extraction and the Object detection. Standard feature extraction further comprises of two sub-processes, namely Generic shape feature extraction and Dimensionality reduction and standardization.



Fig. 1: Object detection framework

In this framework, the process of generic shape feature extraction is kept as a varying process while keeping all other components fixed, in order to provide an identical platform for the evaluation of different generic features. The process of generic shape feature extraction is described in more detail in Section 2.1.

The generic shape features used in this study are of different dimensions, which pose a challenge in comparing them. Therefore, all features are standardized and the dimensions are reduced to one-dimensional scalar value using Support Vector Machine (SVM).

SVM is a standard tool for machine learning technology and widely used for solving problems in classification, regression and novelty detection (CHANG & LIN 2011; BISHOP 2006) because of its good generalization property for even small samples (LIU et. al 2012; VAPNIK & VAPNIK 1998; CORTES & VAPNIK 1995). In this method, the data is mapped into a higher dimensional input space and an optimal separating hyperplane, a straight line in case of two-dimensional input space, is constructed in this space (SUYKENS & VANDEWALLE 1999; CORTES, C. & VAPNIK 1995).

SVM uses training examples of objects and non-objects to determine the decision boundary between them. Generally, the trained SVM are used with sliding window detector where a window translates over the image, calculates the features and based on that feature decide directly the class for the window. But, in this study, it is used not to detect class but to calculate the distance from the decision boundary in feature space and assign it to the central pixel of the window. The process

is illustrated in Figure 2. The function $f(\mathbf{x})$, distance from the decision boundary in feature space, calculated at all pixels of the image result in surface, mentioned as the *feature map*, with peaks at the objects.

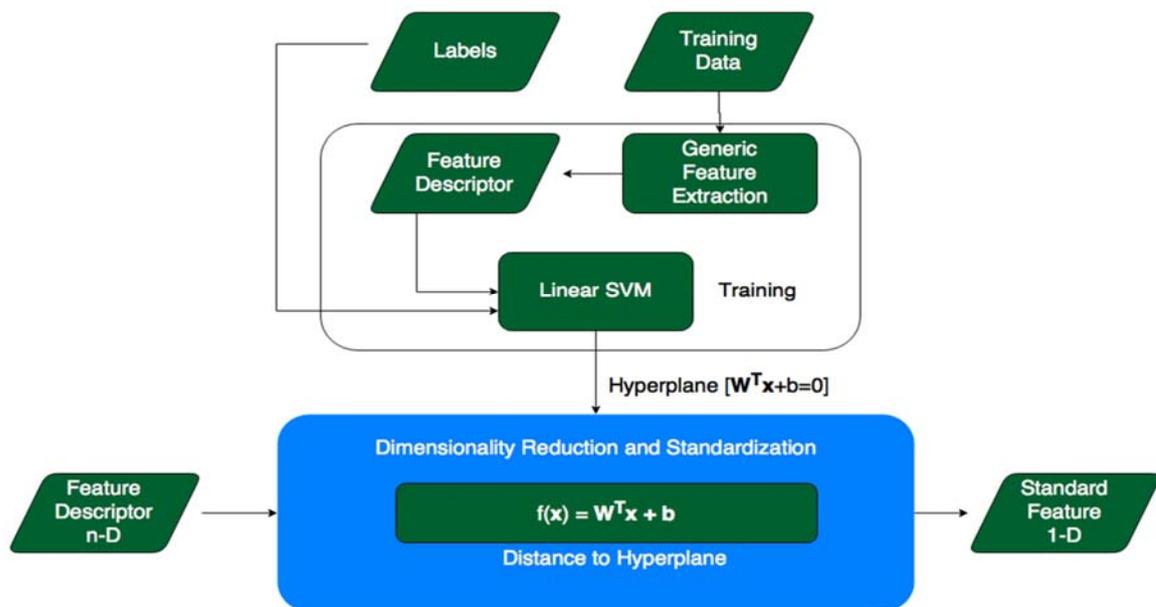


Fig. 2: Dimensionality reduction and standardization of generic feature

A local maximum detection algorithm is applied to the feature map to achieve the spatial distribution of object centers in the image, called *object map*, as the final result. The results for different features are evaluated by comparing the results with the ground truths. The average of precision and recall are used as the accuracy measure. Precision or user’s accuracy is the proportion of detections that are true in reality while recall or producer’s accuracy is the proportion of the true objects detected.

2.1 Generic Feature Extraction

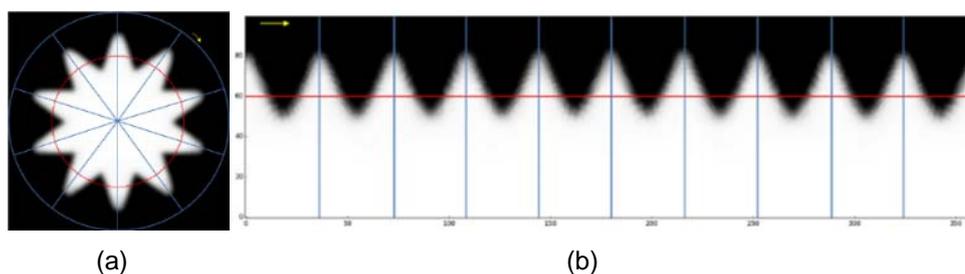


Fig. 3: (a) Model of star-shaped object. (b) Polar shape matrix representation of model

Ideally, star-shaped objects (Figure 3a) are radially symmetric object having constant angular differences between consecutive leaves. If the star-shaped object is rotated about its center with the angle equivalent to the angle difference between two consecutive leaves, the resulting object is identical to the original non-rotated object.

In the polar shape matrix representation, the star-shaped object form triangular waveform like structure (Figure 3b). The polar shape matrix is calculated by sampling the object along the circumference of different concentric circles, centered at the center of the object, in a constant angular interval α (Equation 1).

$$\mathbf{Z}_{r\theta} = \mathbf{Z}_{xy}(x_m + r(\cos \theta)^T, y_m + r(\sin \theta)^T) \quad (1)$$

where \mathbf{Z}_{xy} is the image, $\theta = (0, \alpha, 2, \dots, 360 - \alpha)^T$, $\mathbf{r} = (0, \beta, 2\beta, \dots, \frac{d_m}{2})^T$. d_m is minimum diameter that encloses the shape and x_m and y_m are the coordinates of the center of the shape.

In the case of star-shaped objects, the circular shift along the angular axis of the polar shape matrix results in the repetition of the same image when the shift is equivalent to the angular difference between leaves. The similarity decreases until it reaches the least when the shift is half the angular difference between leaves, and gradually increases to maximum on the other half. This scenario repeats periodically with the period equivalent to the number of leaves. In order to measure the similarity between original polar shape matrix and the circularly shifted version of it, correlation coefficients are used. Therefore, comes the name of the feature: circular autocorrelation of polar shape matrix (CAPS).

Circularly shifting along the angular axis and calculating the correlation for each shift takes long computation time. However, using Wiener-Khinchin Theorem (COHEN 1998) to compute it in Fourier domain makes the process efficient. The theorem is given by,

$$ACR = real(F^{-1}\{F\{\mathbf{Z}_{r\theta}\} \cdot F^*\{\mathbf{Z}_{r\theta}\}\}) \quad (2)$$

where F is the Fourier Transform.

The output of the Wiener-Khinchin function is a circular correlation on both dimensions of the matrix, and the values depend on the distribution of gray level in the matrix i.e. overall lighting condition. As the range of the gray level distribution varies for different images, the output of the Wiener-Khinchin function also varies between images. This is standardized by converting the output to the Pearson correlation coefficient, where the values range from -1 for exact inverse correlation to +1 for exact positive correlation. The conversion can be achieved as following:

$$\mathbf{R} = \frac{ACR - N\mu_z^2}{N\sigma_z^2} \quad (3)$$

where N is the total number of elements in the polar shape matrix. μ_z and σ_z are mean and standard deviation of the polar shape matrix respectively.

As the interest of the study is only along angular shift, the vector with the values at no radial shift from \mathbf{R} is considered. Furthermore, the circular autocorrelation function is symmetric, therefore, only half of this vector i.e. $\frac{360}{2\alpha}$ elements are enough to describe the object and is used as the feature vector.

CAPS, as shown in Figure 4, is a standardized feature vector where the elements range from -1 to +1 in value. The maximum value can appear multiple times for rotationally periodic objects while it only appears at no angular shift for objects that are not rotationally periodic. Circular autocorrelation by nature is unaffected by the linear shift. But, in the case of CAPS, the shift is

carried out along the angular axis. Therefore, CAPS is unaffected by rotation.

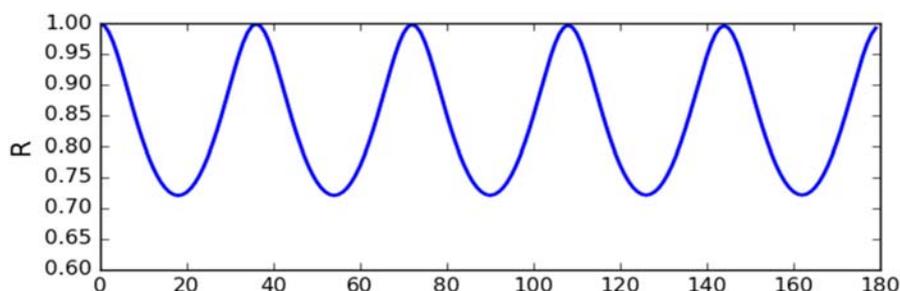


Fig. 4: CAPS of the model

Besides the multiple maximum values and fluctuating curve, CAPS of star-shaped objects also have on average high values of correlation throughout all elements. Taking the mean, as the feature, can serve as the minimal representation of the shape. The average of CAPS elements is highest for the circular object. Therefore, considering it as the feature will be a measure of the object's similarity with the circular object. The average of the CAPS is given by

$$\mu_{caps} = \frac{1}{n} \sum \mathbf{CAPS} \quad (4)$$

The randomness of CAPS is low due to its periodic nature. Entropy gives the measure of randomness (TSAI et. al 2008). Entropy is small for the less random signal and large for the more random signal. Therefore, the CAPS of star-shaped objects should have lower entropy compared to other objects. However, the entropy of CAPS of circular shape has the lowest, i.e. negative infinity. We consider entropy of CAPS as other feature of study. The entropy of a signal is computed with a probability of the signal. If S_1, S_2, S_3, \dots are the signal that occurs with probabilities $P(S_1), P(S_2), P(S_3), \dots$ respectively then, the entropy $H(\mathbf{S})$ of the signal \mathbf{S} with N_s as the total number of unique signal occurrences is given by (GRAY 2011, TSAI et. al 2008):

$$H(\mathbf{S}) = - \sum_{t=1}^{N_s} P(S_i) \log P(S_i) \quad (5)$$

Both mean of CAPS (M-CAPS) and entropy of CAPS (E-CAPS) are rotation invariant and invariant to the overall illumination change. Besides that, both of them are a scalar value. However, for higher object likelihood, M-CAPS tends to be large while E-CAPS tends to be small.

2.2 Results and Discussion

The framework was implemented for detection of a naturally occurring star-shaped object, palm trees. It was applied to 8 test images from 5 cultivation regions in Indonesia, Malaysia, and Thailand. The images were acquired with Trimble UX5 unmanned air system (UAS). The ground sampling distances of image ranged from 3.2 cm to 10 cm. However, for the study, the pixel sizes were resampled to approximately 5cm. The images were projected to Universal Transverse Mercator projection system with WGS84 as an ellipsoid of reference.

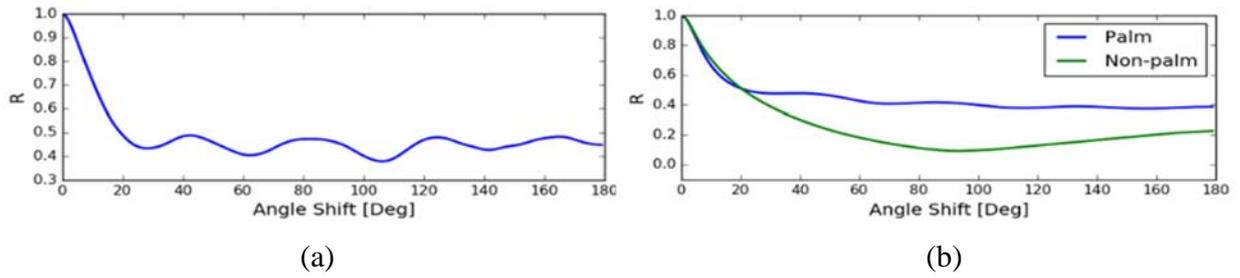


Fig. 5: (a) CAPS of a palm tree training sample. (b) Mean of all elements across all training samples for palm trees and non-palm objects

2.3 Feature characteristics on training data

644 samples of positive and 597 samples of negative training data were randomly selected from one cultivation region only. The inspection of the palm tree canopies found that a window covering 201×201 pixels, in general, covers the entire canopy. Therefore, all the training samples were of the size 201×201 pixels. Same size of the window was also used in the moving window scheme

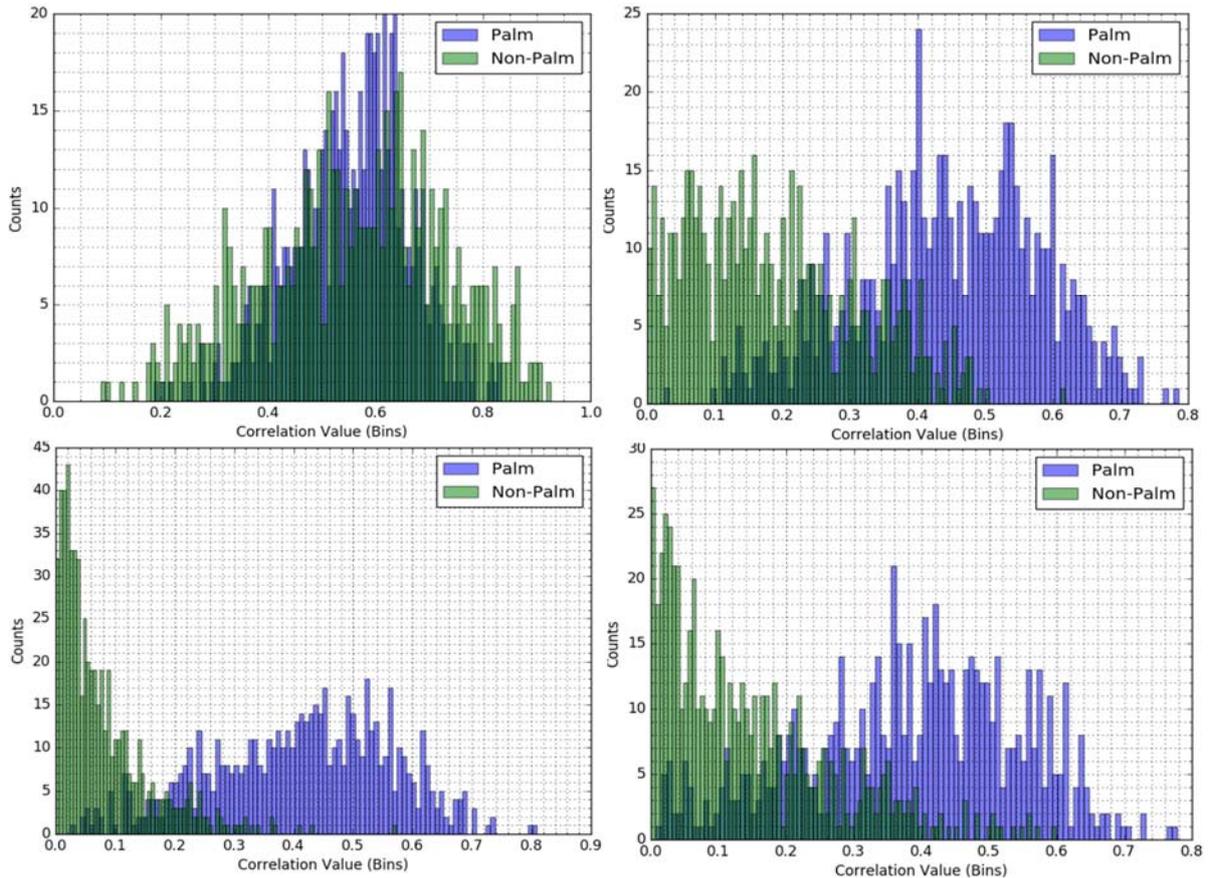


Fig. 6: Histogram of few elements of CAPS calculate across all training samples.

for feature extraction. During the transformation of the sample to the polar shape matrix, $\alpha = 1px$ and $\beta = 1px$ were used, which resulted in CAPS with 180 elements.

As can be seen in Figure 5a, CAPS of the training samples are not a periodic curve implying that the palm trees are not perfectly star-shaped. Furthermore, the presence of background noise and non-palm objects also contribute to such a characteristics. However, the CAPS is still a fluctuating curve with relatively higher values at the position of leaves of palm trees in the polar shape matrix representation. This is, however, in general, not the case for the non-palm objects. The average of each element across all palm trees and non-palm object shows the significant elementwise

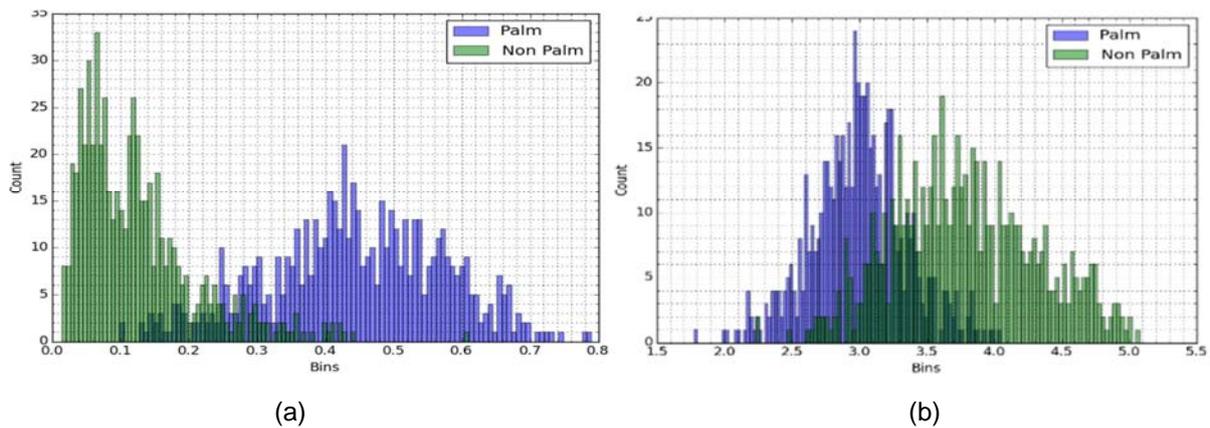


Fig. 7: (a) Histogram of M-CAPS across all training samples. (b) Histogram of E-CAPS across all training samples.

difference between the palm-trees and non-palm objects in Figure 5b.

Histograms of each CAPS elements derived from all positive and negative examples give the preliminary evaluation of the contribution of the element in distinguishing the palm trees and non-palm objects. From the histogram of elements calculate from all positive and negative training samples (Figure 6), it can be perceived that some elements play a significant role in distinguishing the palm trees and non-palm object while some contribute insignificantly, which is demonstrated by a large overlap between the distributions of the palm trees and the non-palm objects. As all elements of CAPS are used, a combined effect of all the elements can expect to be significant.

The histogram of M-CAPS, Figure 7a, shows small overlap between the palm trees and the non-palm objects. As expected, the means are greater for palm trees compared to non-palm objects. It signifies that the palm tree samples tends to be more circular compared to the non-palm training objects. On the contrary, E-CAPS (Figure 7b) have higher overlap. However, the values for palm trees are lower compared to non-palm objects as expected.

The combination of both M-CAPS and E-CAPS, Figure 8, shows highly separable clusters for palm trees and non-palm objects.

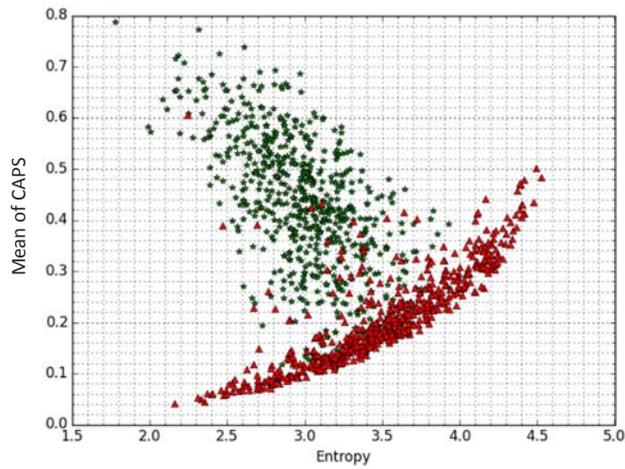


Fig. 8: Distribution of M- & E-CAPS across all training samples



Fig. 9: Object map of Image 01 using CAPS

2.4 Implementation on test images

A moving window of 201×201 pixels was shifted along the rows and columns of the test images at an interval of 5 pixels on both sides, in order to achieve higher efficiency. This would, however, decrease the locational accuracy of detection.

At each window, the features were calculated and trained SVM was used to standardize the feature, which results in the feature maps as shown in Figure 10. The feature maps have peaks at the center of palm trees. But there are also multiple false peaks in non-palm objects, especially, in the artificial objects like houses and the roads. In the case of the feature map derived with M-CAPS, high values can be seen in straight lines like structures such as road and plantation line while in the case of E-CAPS, the extensive regions have high values. The combination of M-CAPS and E-CAPS shows a very clean feature map, however, it also consists the peaks on the houses. Using local maximum detection algorithm, the object map was derived from the feature map as in Figure 9. Figure 9 is the object map derived using CAPS as the feature. The detections are mostly at the centers of the palm trees. There are few false detections on the house and misses on the palm trees along the road. Similar object maps are derived from all test images and all features and the accuracy measures are calculated and are shown in Figure 11.

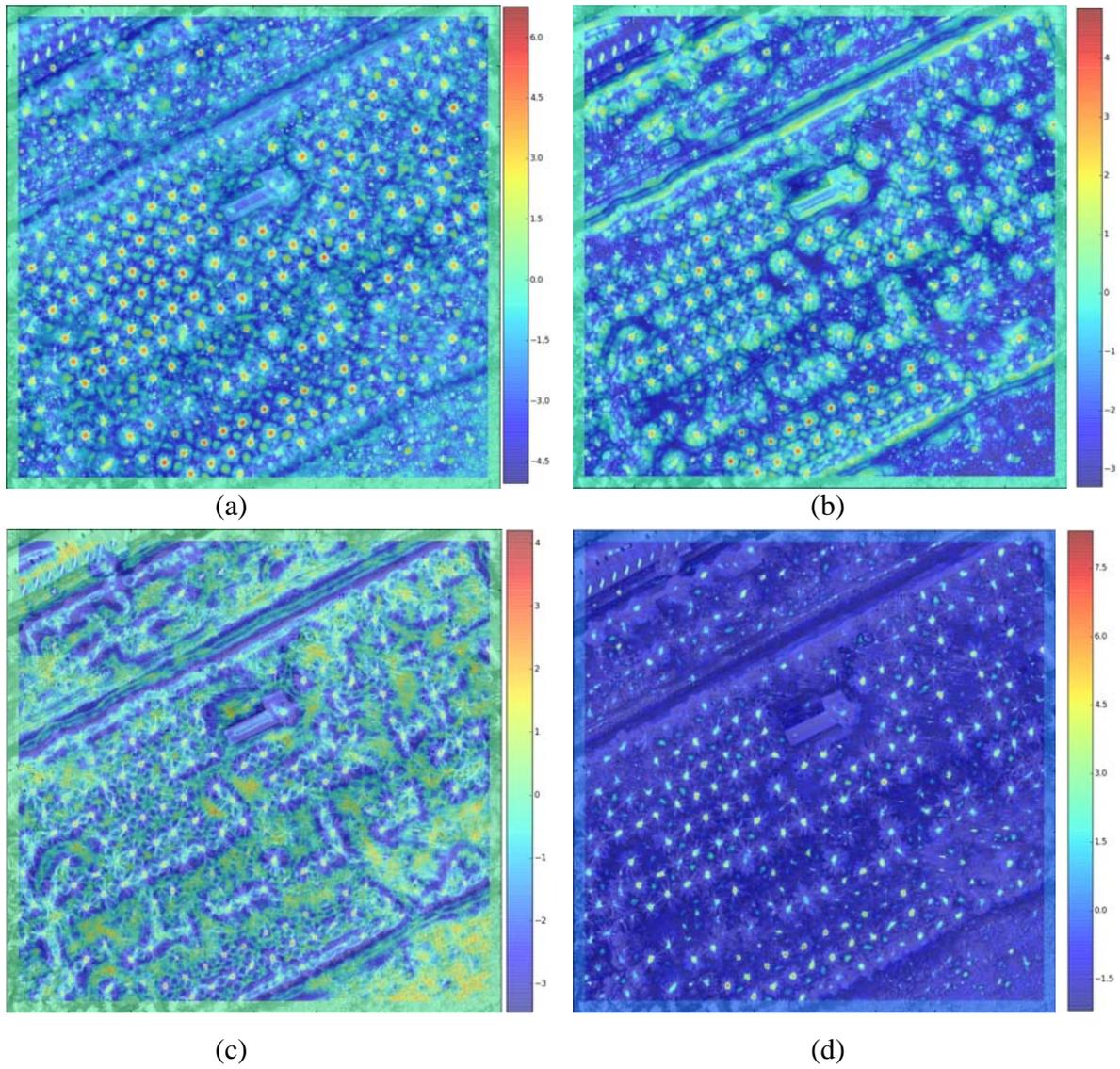


Fig. 10: Feature maps extracted from Image 01 using (a) CAPS (b) M-CAPS (c) E-CAPS (d) Combination of M- & E-CAPS

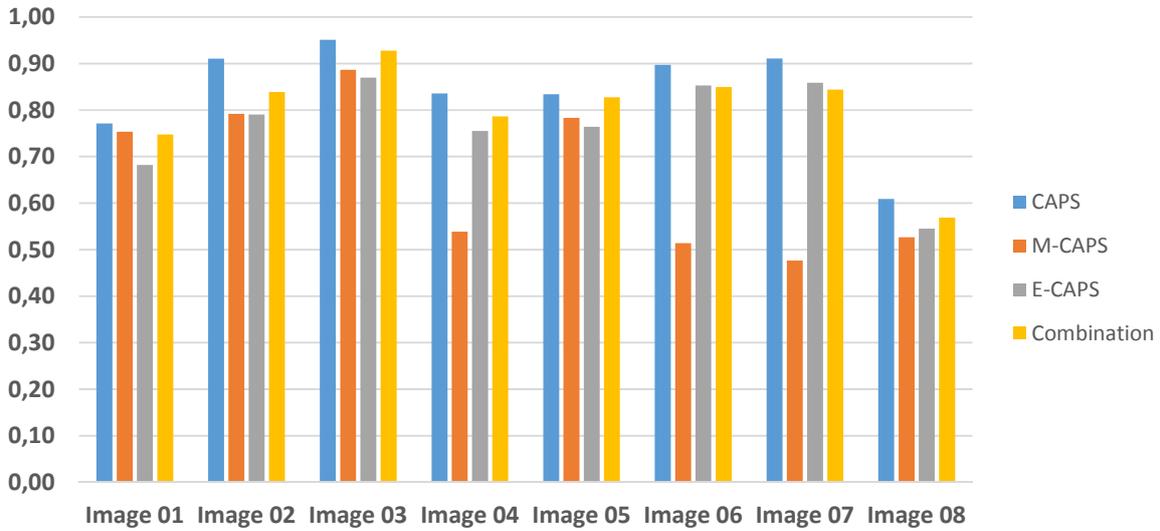


Fig. 11: Accuracies of each feature for each images

The detection using CAPS as the feature produced a good result in detecting palm trees. For six out of eight images, the results were achieved with an accuracy higher than 80% with the highest accuracy up to 95%. The detection using M-CAPS produced on average worst results. The results were especially poor for image 04, 06, and 07, where the palm trees were more resembling the star-shaped objects. The detection using E-CAPS produced better results than M-CAPS. However, the maximum accuracy, it was able to achieve, was 87% and only three out of eight images were detected with accuracy higher than 80%. The combination of M-CAPS and E-CAPS produced better result compared to M- and E-CAPS independently, with the maximum accuracy of 92%, and five out of eight images were detected with accuracy over 80%. The average accuracies of the detection using CAPS, M-CAPS, E-CAPS and the combination of M- & E-CAPS are, as shown in Figure 12, 84%, 66%, 76%, and 80% respectively.

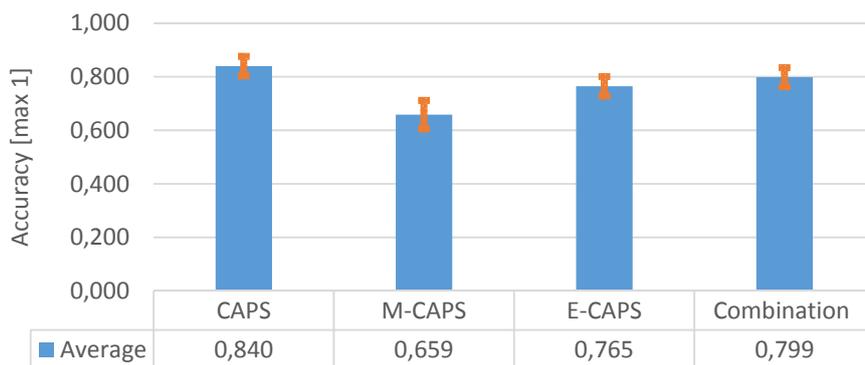


Fig. 12: Average accuracies of each feature across all test images

In all the images, there were few false detections on the houses, roads, and trees with circular canopies, which is driven by the fact that the objects having regular shapes like square and circle also have high CAPS and thus corresponding M- and E-CAPS.

In general, the framework performed quite well in the scenes with multiple challenges like varying sizes, contrasts and cultivation densities. However, artificial objects like houses had several false detections. The false detections were generally at the corners and the centers of buildings. Such false detections can be eliminated using masks or filters with normalized difference vegetation indices. The detection of other regular shapes like square, circles etc. can be reduced by using a learning algorithm, which determines the non-linear decision boundary.

3 Acknowledgements

We would like to acknowledge Trimble Navigation Ltd. for the generous help in the study. The organization supported the project by contributing the UAV images and technical resources for the experiment. We would like to especially thank Mr. Roland Winkler and Dr. Barbara Zenger-Landolt for advising, supporting and giving feedbacks throughout the study.

4 References

- BEN-HUR, A. & WESTON, J., 2010: A user's guide to support vector machines. *Data mining techniques for the life sciences*, Springer, 223-239.
- BISHOP, C.M., 2006: *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- CHANG, C.-C. & LIN, C.-J., 2011: Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, **2** (3), 27.
- COHEN, L., 1998. Generalization of the Wiener-Khinchin theorem. *Signal Processing Letters, IEEE*, 5(11), pp.292-294.
- CORTES, C. & VAPNIK, V., 1995: Support-vector networks. *Machine learning*, **20** (3), 273-297.
- GRAY, R.M., 2011: *Entropy and information theory*. Springer Science & Business Media.
- LIU, X., GAO, C. & LI, P., 2012: A comparative analysis of support vector machines and extreme learning machines. *Neural Networks*, **33**, 58-66.
- POTTS, J., LYNCH, M., WILKINGS, A., HUPPE, G., CUNNINGHAM, M. & VOORA, V., 2014: The state of sustainability initiatives review 2014: Standards and the green economy. *International Institute for Sustainable Development (IISD) and the International Institute for Environment and Development (IIED)*, 332.
- SUYKENS, J.A., & VANDEWALLE, J., 1999: Least squares support vector machine classifiers. *Neural processing letters*, **9** (3), 293-300.
- TSAI, D.Y., LEE, Y. & MATSUYAMA, E., 2008: Information entropy measure for evaluation of image quality. *Journal of digital imaging*, **21** (3), 338-347.
- VAPNIK, V.N. & VAPNIK, V. 1998: *Statistical learning theory*, Volume 1. Wiley, New York.

Unbeschwertes und sicheres Skivergnügen dank Geodaten

STEPHAN LANDTWING¹ & THOMAS GÖTZ²

Zusammenfassung: Betreiber von Wintersportgebiete und Bergbahnanlagen bewegen sich topografisch, klimatisch und ökonomisch in einem rauen Umfeld. Steigende Kundenansprüche, Wettbewerbs- und Investitionsdruck zwingen die Betriebe mehr denn je zu einem maximal effizienten Mitteleinsatz. Zu diesem Zweck greifen moderne und innovative Wintersportgebiete auf eine erstaunliche Palette an Daten und Dienstleistungen aus den Bereichen Fernerkundung, Vermessung und GIS zurück.

Der folgende Beitrag zeigt anhand von drei aktuellen Praxisbeispielen, wie mithilfe von digitalen Geländemodellen und Echtzeit-GNSS der Kunstschnee-Einsatz optimiert wird, wie durch vollautomatischer Überwachungsmessungen die Sicherheit von kritischer Infrastruktur gewährleistet wird, und wie mittels 3D-Karten und Mobile Apps das Erlebnis der Feriengäste um eine Dimension erweitert wird.

1 Einleitung, Umfeld

Perfekt präparierte Skipisten von November bis April, modernste Transportanlagen ohne Wartezeiten mit WLAN und geheizten Sitzflächen, stylische Pistenrestaurants und Lounges an der Piste sowie Echtzeit-Informationen zu Betriebszustand und Events: Die Ansprüche der Wintersport-Gäste an ihre Skigebiete sind in den letzten Jahren rasant gestiegen!

Wintersportgebiete und ihre Betreiber sehen sich vor der Herausforderung, in anspruchsvollem bis unwirtlichem Gelände komplexe technische Anlagen zu planen, zu errichten und zu betreiben. Klimaerwärmung und internationale Konkurrenz potenzieren den Wettbewerbs- und Investitionsdruck und zwingen die Betriebe zu einem maximal effizienten Mitteleinsatz.

Darüber hinaus gilt es, das Produkt „Wintersporterlebnis“ bei den (potenziellen) Kunden attraktiv zu präsentieren und zu bewerben. Hier haben neben den traditionellen Kanälen in jüngster Zeit Plattformen wie Mobile Apps, Google Earth, TripAdvisor, YouTube etc. stark an Bedeutung gewonnen. Die Wintersportler möchten sich bereits zuhause ein Bild von ihrer Feriendestination machen.

Bei all diesen Aufgaben greifen moderne und innovative Wintersportgebiete und Bergbahnen heute auf eine erstaunliche Palette an Daten und Dienstleistungen aus den Bereichen Fernerkundung, Vermessung und GIS zurück:

- Gelände- und Oberflächenmodelle zur Planung von Bahnanlagen (Längs-, Querprofile) und Pisten (Massenbewegungen, Volumenberechnung)
- Gezielte Pistenpräparierung durch Maschinen-Steuerung auf der Basis von Geländemodellen („Snow Management“)

¹ BSF Swissphoto AG, Dorfstrasse 53, 8105 Regensdorf-Watt, Schweiz
E-Mail: stephan.landtwing@bsf-swissphoto.com

² Grünenfelder und Partner AG, Denter Tumas 6, 7013 Domat/Ems, Schweiz
E-Mail: thomas.goetz@gruenenfelder.ch

- 3D-Visualisierungen für Pistenpläne und interaktive Karten, aufgebaut aus Oberflächenmodellen und (Winter-)Orthophotos
- GIS und Leitungskataster zur Verwaltung der Anlagen, z.B. zur technischen Beschneigung mit Rohr- und Stromleitungssystemen
- Überwachung von potenziell gefährdeten Anlagen auf Bewegungen und Deformationen (z.B. im Permafrost)
- Geländemodelle zur Simulation von Naturgefahren (Lawinen, Murgänge, Steinschlag)

Die Erfassung dieser Fernerkundungs- und Vermessungsdaten im hochalpinen Gelände ist ebenfalls eine Herausforderung. Angesichts von Topographie und Witterung kommt der Wahl der effizienten Technologie eine besondere Bedeutung zu; oft sind Speziallösungen oder Kombinationen von Messverfahren gefragt, um den Vermessungsauftrag erfolgreich durchführen zu können: Luftgestützte (Laserscanning- oder Bild-basierte Sensoren auf Flugzeug, Helikopter oder Drohne) sowie terrestrische (Laserscanning, Monitoring, GNSS, Totalstationen) Sensoren und Verfahren ergänzen sich.

2 Projektbeispiel 1 – Snow Management-System Weisse Arena

2.1 Ausgangslage und Ziel

Die künstliche Beschneigung von Skipisten zur Erhöhung der Schneesicherheit hat sich in den letzten Jahren im Alpenraum fast flächendeckend durchgesetzt. Gemäss Schweizerischem Institut für Schnee- und Lawinenforschung (SLF 2016) werden in der Schweiz zur Zeit 36% der Skipistenfläche technisch beschneit. In Österreich sind es bereits 66% und in den Italienischen Alpen können einzelne Skigebiete sogar bis zu 100% beschneit werden.

Schweizer Skigebiete haben bis 2015 ca. 500 Millionen in Beschneiungsanlagen samt Speicherseen, Rohr- und Stromleitungssystemen und Schneekanonen investiert (LINER & ARN 2015). Die Betriebskosten sind ebenfalls hoch: Die Produktion von 1 m³ technischem Schnee („Kunstschnee“) kostet 4 bis 7 CHF. Grosse Skigebiete wie die Weisse Arena Flims Laax Falera produzieren pro Saison bis zu 1 Mio. m³ Schnee.

Hinzu kommen die Kosten für Pistenpräparierung und Schneeverteilung: Eine Pistenmaschine verbraucht ca. 25 – 30 l Diesel pro Betriebsstunde und steht rund 1200 Betriebsstunden pro Saison im Einsatz. Ein komplett ausgestattetes Pistenfahrzeug kostet rund 0.5 Mio Franken.

Die effiziente Bewirtschaftung der Schneereserven hat somit einen hohen Einfluss auf die Betriebskosten eines Wintersportgebietes. Die Weisse Arena Bergbahnen haben seit der Wintersaison 2013/14 ein Snow Management-System im Einsatz, welches genau dies ermöglicht: Mithilfe eines hochgenauen Geländemodells sowie präzisen Echtzeit-Positionsmessungen an den Pistenfahrzeugen ist die aktuelle, effektive Schneehöhe bei der Pisten-Präparierung jederzeit bekannt. Schnee-Depots können geschätzt und Schwachstellen mit dünner Schneedecke können gezielt verstärkt werden. Somit sinkt sowohl der Anteil des ungenutzt liegengelassenen Schnees als auch die Anzahl „brauner Flecken“ auf den Skipisten.

2.2 Datengrundlagen, Datenerfassung

Als Grundlage für das Snow Management wird ein Geländemodell (DGM) benötigt, welches in der schneefreien Zeit aus der Luft erfasst wird. Als effizienteste Methode zur Bereitstellung dieses DGM für ganze Skigebiete (Flächen von 10 bis 100 km²) mit der benötigten Genauigkeit von ca. 10 cm hat sich das Lufgestützte Laserscanning (Lidar) etabliert. Neben dem Geländemodell werden auch sämtliche Objekte der Erdoberfläche gemessen: Vegetation, Gebäude, Seilbahnen, Stromleitungen, usw. sowie optional ein Orthophoto erstellt. Diese 3D-Daten stehen für weitere Anwendungen auch ausserhalb des Snow Managements zur Verfügung: Anlagenplanung, Geländearbeiten, Raumplanung etc.

Im Herbst 2013 wurde das Gelände der Weissen Arena durch BSF Swissphoto mit einem flugzeuggestützten Laserscanner befliegen. Die Fläche umfasst rund 68 km² und liegt auf einer Höhe von 1100 – 3000 m ü. M. Die Erfassung von Laserscanning-Daten im hochalpinen Umfeld bietet einige spezielle Herausforderungen:

- Flugplanung: Einhalten möglichst konstanter Punktdichte und Streifenbreite innerhalb der Möglichkeiten des Sensors durch detaillierte Planung anhand eines bestehenden Geländemodells und geländeverfolgender Befliegung.
- Befliegung: Nutzung eines Flugzeugs mit hoher Steigrate, Sauerstoffversorgung und drei-Mann-Crew (Pilot, Kopilot, Sensor-Operateur) zur Bewältigung der Aufgaben Flugzeugsteuerung, Funk, Sichtflug/Verkehr, Einhalten Flugplanung (inkl. 3D-Profil), Kontrolle Sensor-Funktion und Echtzeit-Datenkontrolle.
- Datenprozessierung: Tightly Coupled GNSS/IMU-Berechnung zum robusteren Umgang mit reduziertem GNSS-Empfang durch Abschattung durch Berge, grosse Höhenunterschiede zu Basis-Stationen; Streifenanpassung mit sorgfältig ausgewählten Verknüpfungsflächen; Bodenpunkt-Klassifizierung im steilen Gelände mit angepassten Parametern und manueller Nacharbeit.

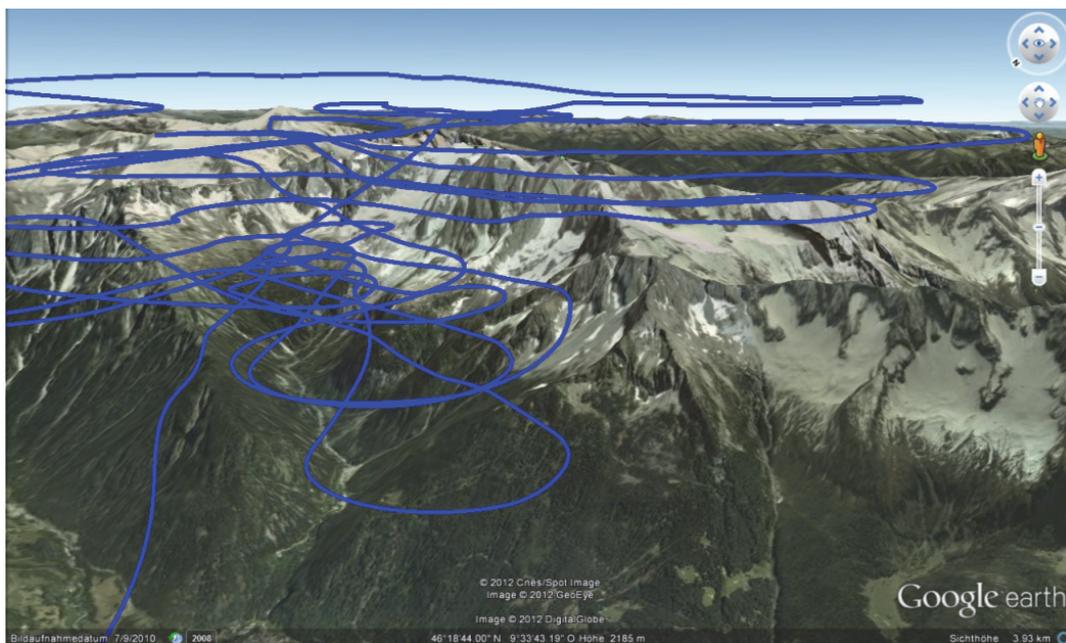


Abb. 1: Flugpfad einer Laserscanning-Befliegung im Hochgebirge (© BSF Swissphoto)

Als Resultat der Laserscanning-Befliegung liegt ein digitales Geländemodell mit einer Punktdichte von 6 Punkten pro Quadratmeter vor, welches als Dreiecksmodell oder 0.5 m-Raster verwendet werden kann.

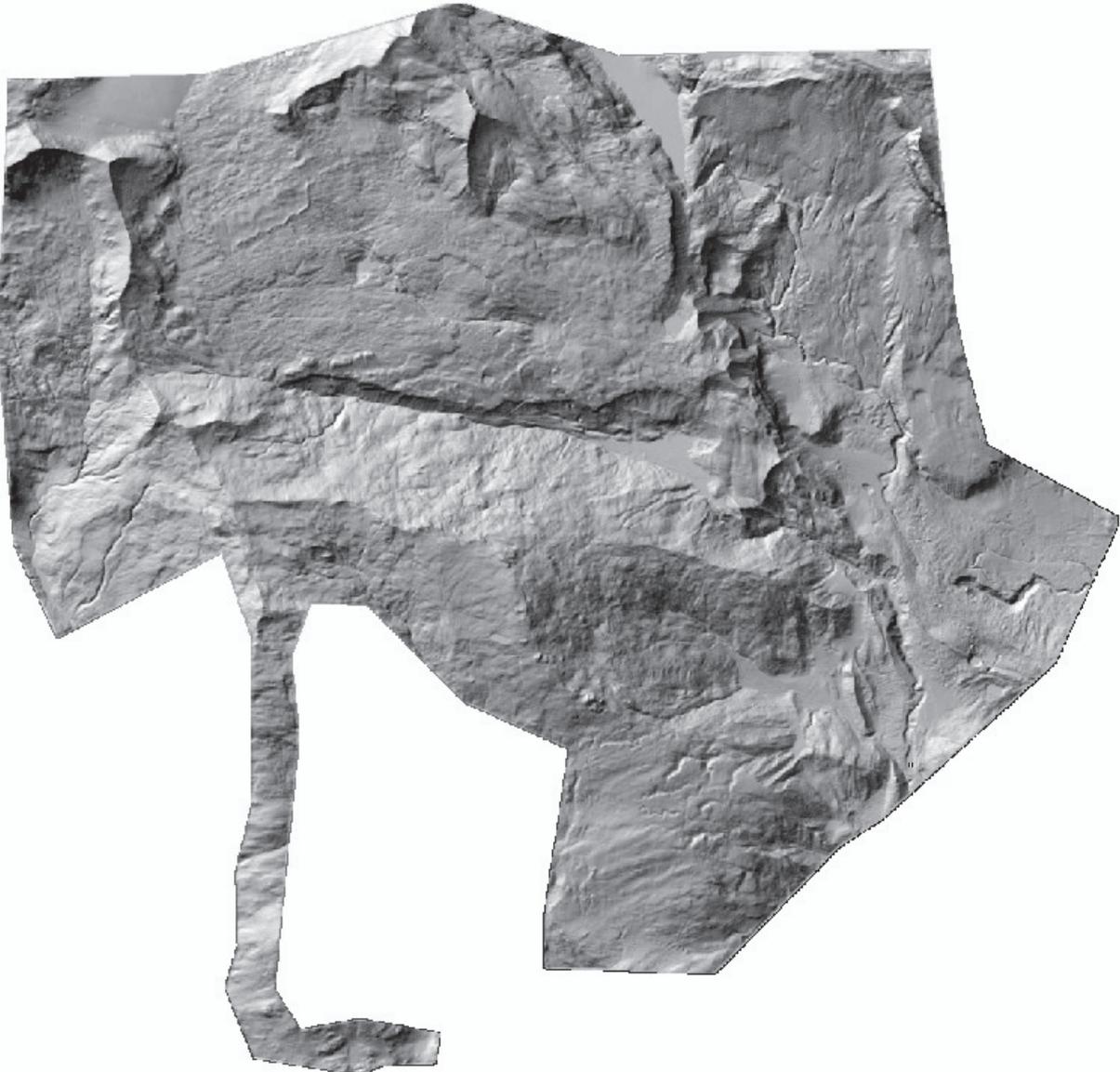


Abb. 2: Schummerungs-Darstellung des Geländemodells der Weissen Arena

Punktuelle Nachführungen des Modells (z.B. nach Geländeänderungen) können alternativ zum luftgestützten Laserscanning auch mit terrestrischen Methoden oder mittels UAV-gestützten Bandaufnahmen erfolgen.

2.3 Installation, Systemkomponenten

Mit dem DGM als Grundlage wurde auf den Pistenfahrzeugen der Weissen Arena das Snow Management-System iCONAlpine von Leica installiert. Dieses System – welches von der

Baumaschinensteuerung stammt – ist momentan das einzige auf dem Markt mit einer zuverlässigen und genauen Schildhöhenmessung. Die Installation am Pistenfahrzeug beinhaltet folgende Komponenten:

- Zweiphasen-GNSS-Antenne auf dem Dach inkl. Empfang von RTK-Korrekturen einer dedizierten Basisstation oder dem Swipos-Dienst
- Zwei Neigungssensoren zur Messung von Quer- und Längsneigung der Maschine
- Optional zwei zusätzliche Neigungssensoren zur Messung von Quer- und Längsneigung des Planierschildes
- Anzeigedisplay für den Fahrer
- Datenübertragung zum Betriebszentrum

Durch Abgleich der aktuellen Position des Pistenfahrzeugs mit dem Geländemodell lässt sich jederzeit die Schneehöhe unter dem Schild berechnen.

2.4 Betrieb und Resultate

Mit Hilfe des innovativen Snow Management-Systems sieht der Fahrer des Pistenfahrzeuges auf seinem Display jederzeit, wie viel Schnee sich unter seinem Schild befindet. Zugleich misst jedes Pistenfahrzeug auf seiner Fahrt laufend die aktuelle Schneehöhe und sendet diese Daten in die Betriebszentrale. Dort werden diese Daten in einem GIS-System aggregiert und erlauben so flächendeckende und tagesaktuelle Auswertungen über die Schneemenge und deren Verteilung. Dies wiederum erlaubt es, die Beschneigung optimal zu steuern und die Kosten für die Kunstschneeproduktion zu minimieren.



Abb. 3: Ansichts-Modi des Fahrer-Displays mit Schneehöhen

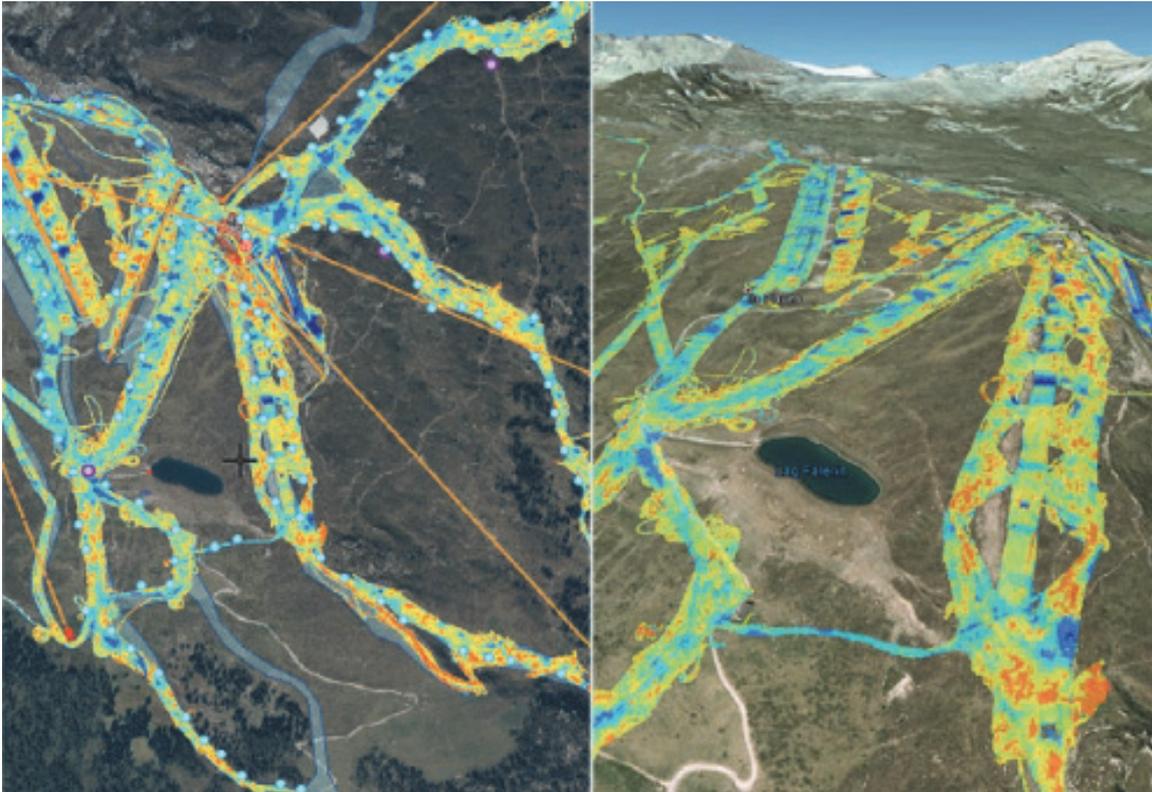


Abb. 4: Die Visualisierung der Schneehöhen im Betriebs-GIS der Weisse Arena-Bergbahnen (links 2D, rechts 3D) ermöglicht jederzeit einen schnellen und genauen Überblick über die Schneesituation im Skigebiet und erlaubt gezieltes Beschneien und Präparieren der Pisten. (© Weisse Arena/Leica)

Seit der Wintersaison 2015/2016 sind 10 Pistenfahrzeuge der insgesamt 17 Maschinen starken Flotte im Skigebiet Weisse Arena mit dem iCONalpine-System ausgerüstet und sorgen dafür, dass die Schneehöhenkarte des Skigebietes jede Nacht aktualisiert wird. Die Fahrer haben das neue Hilfsmittel nach anfänglicher Skepsis sehr positiv aufgenommen, als klar wurde, dass die Schneehöhenmessung über Erwarten genau und zuverlässig funktioniert. Die Bergbahnen schätzen, dass der Effizienzgewinn durch gezieltere Pistenpräparierung und effizientere Kunstschneeproduktion in der Grössenordnung von 10% liegt.

3 Projektbeispiel 2: Geo-Monitoring Talstation Heimberg-Motta

3.1 Ausgangslage und Ziel

Viele Bauten von Seilbahnanlagen befinden sich im Permafrost-Gebiet, welches ab ca. 2500 m ü.M. beginnt. Als Folge der Klimaerwärmung treten Stabilitätsprobleme bei der Infrastruktur und Verwerfungen im Gelände auf (siehe BOMMER et. al 2009). Auch die Jahreszeiten (Wasserhaushalt) haben einen entscheidenden Einfluss auf das Bewegungsverhalten der Bahnanlagen.

Beim Bau der Talstation der 6er-Sesselbahn Heimberg-Motta der Lenzerheide Bergbahnen im Jahr 2013 wurden unerwartete Setzungen festgestellt. Umfang, Verlauf und Entwicklung der

Deformationen mussten überwacht und aufgezeichnet werden, um bei kritischen Veränderungen sofort betriebliche und/oder bautechnische Massnahmen ergreifen zu können.

3.2 Installation, Systemkomponenten

Sofort nach Auftreten der ersten Anzeichen von Setzungen wurde die Talstation noch während der Bauphase durch die Grünenfelder und Partner AG mittels periodischer (3 mal pro Woche), manueller Deformationsmessungen (Tachymeter, Nivellement, Inklinometer) überwacht. Nach Fertigstellung der Bahnanlage erfolgte ein lückenloser Übergang zu einem automatischen GeoMonitoring-System Swissmonitor.

Dieses besteht aus folgenden Komponenten:

- 1 Leica Tachymeter
- 4 Zweiachs-Neigungssensoren (je 2 am Steher vorne und hinten)
- 3 Einphasen-GNSS-Empfänger (1 Referenz / je 1 Rover an Berg- und Talstation)
- Redundanten Datenkommunikation via RS485 und UMTS



Abb. 5: Von links: GNSS-Antenne (Vordergrund) und Seilbahnstation Heimberg (Hintergrund); Tachymeter-Station im Betriebsraum; Neigungssensoren-Box im Schnee (© Grünenfelder und Partner)

Um das zuverlässige Funktionieren des Systems bei hochalpinen Bedingungen an 365 Tagen zu gewährleisten, mussten einige besondere Massnahmen getroffen werden:

- Geheizte Box der Neigungssensoren
- Tachymeterstation in geheiztem Gebäude
- Höherlegung der GPS-Antennen (Schnee)
- Schutz-Vorrichtung für Prismen (Schnee und Vandalismus)
- Zweimal täglich Reinigung der Prismen bei Schneefall durch Betriebspersonal

3.3 Betrieb und Resultate

Das Monitoring-System erlaubt kontinuierliche Überwachungsmessungen (Messintervall 1 Stunde) der Talstation in Lage und Höhe sowie deren Verkippung. Diese Messwerte werden der Betriebsleitung der Lenzerheide Bergbahnen auf einem Internet-Portal in Form von übersichtlichen Grafiken, Zeitreihen und Tabellen zur Verfügung gestellt. Überschreiten

Verschiebung oder Verkippung spezifizierte Alarmwerte, erfolgt eine vollautomatische Alarmierung via SMS. Zur Vermeidung von Fehlalarmen erfolgen umfangreiche Plausibilitätsprüfungen mittels Vergleich zu Nachbarsensoren bzw. Messstellen.

VERSCHIEBUNGEN

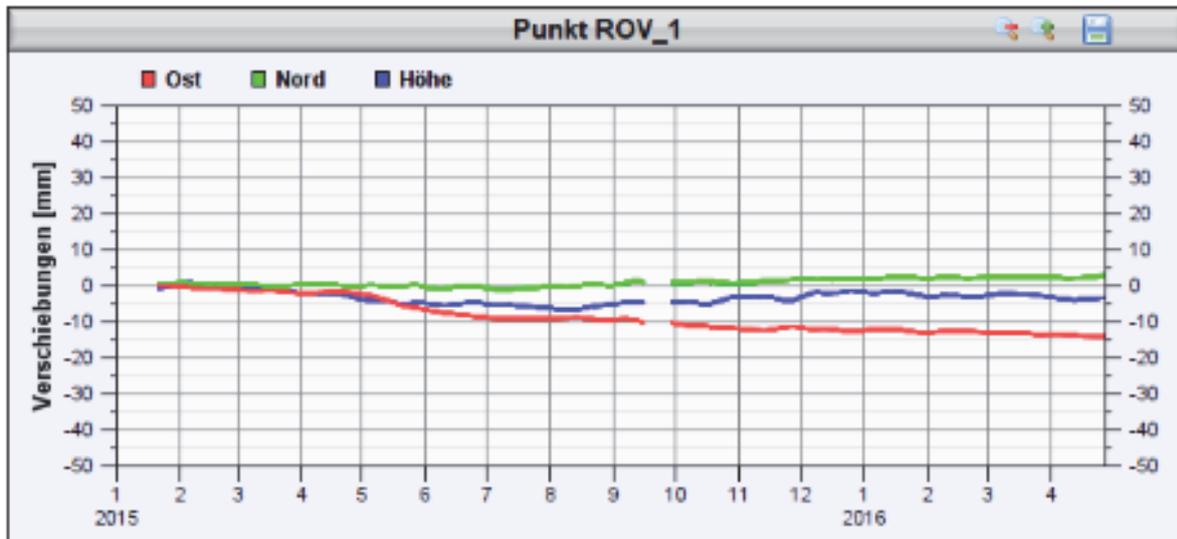


Abb. 6: Darstellung der Messresultate auf dem Internet-Kundenportal (© Grünenfelder und Partner)

Nur dank dem konsequenten und nachweislich zuverlässigen Einsatz des Monitoring-Systems konnte am Ende der Bauzeit trotz weiter auftretenden Setzungen rechtzeitig die Konzession erteilt und die Anlage Heimberg-Motte in Betrieb genommen werden. Seither funktioniert das System Swissmonitor mit einer Verfügbarkeit von über 95% sehr zuverlässig. Deformationen traten und treten im erwarteten Bereich auf und werden quartalsweise mit manuellen Inklinometer- und Nivellement-Messungen bestätigt. Dank diesen genauen und tagesaktuellen Zahlen konnten und können gezielte bauliche Massnahmen getroffen und die strukturelle Integrität der Anlagen jederzeit gewährleistet werden.

4 Projektbeispiel 3: 3D-Pistenplan Oberengadin

4.1 Ausgangslage und Ziel

Das Produkt „Wintersporterlebnis“ muss bei den (potenziellen) Kunden attraktiv präsentiert und beworben werden. Hier haben neben den traditionellen Kanälen in jüngster Zeit Plattformen wie Mobile Apps, Google Earth, TripAdvisor, YouTube etc. stark an Bedeutung gewonnen. Die Wintersportler möchten sich bereits zuhause ein Bild von ihrer Feriendestination machen.

Die Firma 3D RealityMaps entwickelt zu diesem Zweck für die Tourismusregion Oberengadin einen interaktiven 3D-Pistenplan als Internet- und Mobil-App, welche Pisten, Bahnen, Berghütten, Webcams und vieles mehr auf einer fotorealistischen 3D-Karte darstellt.

4.2 Datenerfassung und –aufbereitung

Basis für den 3D-Plan bilden hochauflösende Luftbilder, welche – im Gegensatz zu praktisch allen existierenden Bilddaten – gezielt im Winter aufgenommen werden. Um ein maximal attraktives Bild zu erhalten, muss der Messflug kurz nach Schneefällen, bei wolkenlosen Wetterbedingungen – oft über ein grosses Gebiet – sowie genau zur Mittagszeit erfolgen. Für das Oberengadin erfasste BSF Swissphoto im Winter 2014/15 an einem Tag 3'200 Luftbilder über eine Fläche von mehr als 1'700 km².

Aus den orientierten Luftbildern berechnet 3D RealityMaps über Bildkorrelation ein hochgenaues 3D-Oberflächenmodell. Das angewendete Verfahren wurde speziell weiterentwickelt, um auch auf texturarmen, weissen Schnee- und Firnflächen ein zuverlässiges Matching zu erzielen. Das mit dem Orthophoto texturierte Oberflächenmodell wird anschliessend mit den Zusatzinformationen (Vektordaten inkl. Links) überlagert.

4.3 Präsentation

Der Nutzer kann die 3D-Karte in einem Web-Browser ohne Plug-In oder via App auf Smartphones oder Tablets nutzen. Per Maus oder Touchscreen kann er das Oberengadin interaktiv erkunden und die Landschaft aus allen Perspektiven betrachten. Die Suche nach geeigneten Pisten und Bahnen ist in der fotorealistischen Karte effizienter und intuitiver, Echtzeit-Information (Pisten- und Anlagenzustand, Tagesmenü im Pistenrestaurant, Wartezeit an der Talstation, freie Parkplätze etc.) kann problemlos eingebunden werden.



Abb. 7: Interaktiver 3D-Pistenplan des Oberengadins (© 3D RealityMaps)

Die fotorealistischen Karten sind mit geographischen Koordinaten versehen und damit voll navigationsfähig. Der App-Nutzer sieht seine eigene Position in der Karte und kann sich so orientieren.

Die Umsetzung für das Oberengadin befindet sich zurzeit im Teststadium und wird zur Ski-WM 2016 in St. Moritz aufgeschaltet. Ähnliche Applikationen von 3D RealityMaps sind für Skigebiete in Österreich und Italien bereits erfolgreich im Einsatz.

5 Fazit & Ausblick

Die Bedeutung von genauen 3D-Geodaten nimmt im Umfeld von Wintersportgebieten genauso zu wie in anderen Bereichen von Wirtschaft und Gesellschaft. Wie die obenstehenden Beispiele zeigen, gibt es bereits zahlreiche praxisreife Anwendungen für state-of-the-art Messsysteme und nachgelagerte Auswertung und Präsentation, welche Betrieb und Benutzung von Wintersportgebieten effizienter, sicherer und erlebnisreicher machen. Weitere Anwendungen sind bereits in Vorbereitung oder Einführung:

- Übertragung des Konzepts des Schnee-Managements auf die Maschinensteuerung bei der Schneeräumung der Alpenpässe im Frühling.
- Integration weiterer für die Gäste relevanter Informationen in die 3D-App, z.B. Überlagerung des 3D-Modells mit der aktuellen, lokalen Lawinengefährdung.

An vielen Stellen werden von Bergbahnbetreibern oder ihren Subunternehmern 3D-Geodaten generiert und gespeichert. Der nächste entscheidende Schritt wird in der Aggregation und Überlagerung dieser Daten in einem zentralen, übersichtlichen GIS-System („Cockpit“) bestehen. So können bei geeigneter Datenintegration sämtliche für Betrieb und Unterhalt wichtigen Daten gemeinsam verwaltet und analysiert werden, z.B.:

- Infrastruktur: Leitungskataster, Anlagen-Komponenten, Überwachungsmessungen
- Schneehöhen, Beschneiungsanlage, Flottenmanagement Pistenfahrzeuge
- Gästeverhalten, Wartezeiten, Kapazitäten
- Raumplanung, Amtliche Vermessung

6 Literaturverzeichnis

- BOMMER, C., PHILLIPS, M., KEUSEN, H.-R. & TEYSSEIRE, P., 2009: Bauen im Permafrost: Ein Leitfaden für die Praxis, Birmensdorf, Eidg. Forschungsanstalt für Wald, Schnee und Landschaft WSL.
- LINER, M. & ARN, M., 2015: Künstliche Beschneiung in der Schweiz. ProNatura Hintergrund November 2015, 3.
- SLF, 2016: Technischer Schnee. http://www.slf.ch/forschung_entwicklung/schnee/kunstschnee (abgerufen am 22. April 2016).

AUTORENVERZEICHNIS

A

ADAMS, M.	306
ADLER, P.	405
AGOUB, A.	208
ARTUSO, R.	66
ATZBERGER, C.	417
AUGSTEN, N.	555

B

BACHMANN, M.	395
BAHR, T.	124
BALTRUSCH, S.	131
BALTSAVIAS, E.	506
BANASKIEWICZ, M.	288
BANNEHR, L.	34
BARALDI, A.	555
BART, R.	143
BAUMANN, A. G.	506
BAYER, S.	471
BÉDERT, N.	66
BELGIU, M.	555
BLAHÁ, M.	44
BLASER, S.	217
BÖCK, S.	417
BÖSCH, R.	306
BONGARTZ, J.	89
BRÄNDLI, M.	288
BRAUNISCH, V.	405
BRÉDIF, M.	351
BRIDGES, J.T.	288
BRIESE, C.	131
BROCKMANN, H.	131
BRODOVSKII, M.	153
BRUGGISSER, M.	251
BRUNN, A.	153, 436, 477
BUCHER, T.	471
BÜHLER, Y.	306
BYRNE, S.	288

C	
CANNELLE, B.	449, 487
CASCIELLO, A.	288
CASPER, E.	370
CAVEGN, S.	113
CHIZHOVA, M.	153
CHRISTEN, M.	217
CLAUSEN, P.	26
COENEN, M.	516
COORS, V.	370
CRAMER, M.	13, 574
CREMONESE, G.	288
CROMMELINCK, S.	524
D	
DA DEPPO, V.	288
DEBEI, S.	288
DE ROUX, N.	288
DORIGO, W.	565
DORN, C.	131
DRESCHLER-FISCHER, L.	165
E	
EBERHARDT, L.	97
EDER, K.	461
EINZMANN, K.	417
EKKEL, T.	259
EL-MAARRY, M.R.	288
ERISMANN, M.	288
F	
FICAI VELTRONI, I.	288
FÖRG, P.	13
G	
GÄNSRICH, A.	471
GALLIANI, S.	361
GANDOR, F.	66
GAMBICORTI, L.	288
GERBER, M.	288
GERBER, T.	288
GHOSE, K.	288
GINZLER, C.	306
GÖTZ, T.	593
GRAW, V.	428
GREIWE, A.	97
GRENZDÖRFFER, G.	1
GRIMM, P.	298
GRUBER, M.	288
GUBLER, P.	288
GUILLEN ESCRIBA, C.	251

H	
HAALA, N.	113
HANEL, A.	177
HANSEN, C.J.	288
HASTEDT, H.	259
HAUBER, E.	288
HAUSNER, T.	288
HEINEMANN, S.	428
HENSON, R.	288
HERMLE, D.	317
HEUSINGER, A.	477
HOEGNER, L.	177, 461, 582
HOLLAUS, M.	240
HOLLENSTEIN, D.	327
HOLZER, N.	124
I	
INSINNA, P.	240
IMMITZER, M.	417
IVANOV, A.	288
J	
JACOBSEN, K.	80
JENAL, A.	89
JEPPING, C.	259
JOHNSON, M.	288
JOST, J.	288
JUNG, M.	531
K	
KADA, M.	208
KAHMEN, O.	538
KAISER, M.	461
KASSER, M.	188
KERSTEN, T.	271
KESTAY, L.	288
KIRK, R.	288
KLÖDEN, H.,	177
KNEER, C.	89
KOLBE, T. H.	54, 382
KOROVIN, D.	153
KREMER, J.	298
KÜKENBRINK, D.	251
KUNDE, F.	208
KUTZNER, T.	382
KUZMIN, R.	288

L

LANDTWING, S.	593
LANG, S.	555
LASICA, R.	124
LASINGER, K.	361
LAUSCH, A.	34
LEITERER, R.	251
LINDSTAEDT, M.	271
LOCHMATTER, P.	288
LOESCH, B.	217
LUHMANN, T.	259, 538

M

MAAS, H.-G.	230
MALLET, C.	351
MANANDHAR, A.	582
MANDLBURGER, G.	131, 240
MANGOLD, N.	288
MARINANGELI, L.	288
MARKIEWICZ, W.J.	288
MASSIRONI, M.	288
MC EWEN, A.S.	288
MENZ, G.	428
MEYER, T.	436
MISCHLER, H.	288
MISGAISKI-HASS, M.	271
MORS DORF, F.	251
MÜCKE, W.	131

N

NEBIKER, S.	113, 197, 217, 327, 338
NG, W. T.	417
NOWOSIELSKI, W.	288

O

OKUBO, C.	288
ORLEANSKI, P.	288
OTEPKA, J.	131

P

PELÒ, E.	288
PETERSEN, M.	405
PFEIFER, N.	131, 240
PIAZZA, D.	288
POCK, T.	44
POMMEROL, A.	288
PRZYBILLA, H.-J.	271

Q

R

REHAK, M.	26
REZNICEK, J.	259
RICHARD, A.	44
RICHTER, K.	230
RIEDER, M.	288
RIEGER, W.	461
RIENOW, A.	428
RÖDER, M.	477
ROLOFF, V.	288
ROTTENSTEINER, F.	516
RUNKEL, I.	317

S

SABOW, F.	13
SCHAEPMAN, M.	251
SCHINDLER, K.	44, 361
SCHLIENKAMP, A.	97
SCHNEIDER, F.	251
SCHORCHT, M.	545
SCHULTZ, J.	428
SELG, F.	428
SEPPKE, B.	165
SERVONET, A.	288
SINDRAM, M.	54
SKALOUD, J.	26
SODOR, B.	288
SÖRGEL, U.	496
SPRECKELS, V.	97
STEIN, A.	395
STEINER, G.	487
STILLA, U.	153, 177, 461, 582
STOFFEL, A.	306
SUDMANN, M.	555
SUMMER, G.	131
SUTTER, G.	288
SZALAI, S.	288

T

TIEDE, D.	555
THOMAS, N.	288
TORABZADEH, H.	251
TORNABENE, L.L.	288
TROTTMANN, W.	288
TROZNAI, G.	288
TSCHIRSCHWITZ, F.	271

U

ULRICH, C.	34
UTHACHAROENPONG, T.	288

V	
VERNANI, D.	288
VIERTL, J.	288
VOGEL, C.	44
VUOLO, F.	417
W	
WAGNER, D.	370
WAGNER, W.	565
WAHBEH, W.	338
WAJER, P.	288
WEBER, I.	89
WEGNER, J. D.	44
WEIGEL, T.	288
WEINMANN, M.	351
WESTFELD, P.	230
WILLENBORG, B.	54
WILMS, C.	165
WEIß, R.	230
WRAY, J.J.	288
WÜEST, R.	197
X	
XAVER, A.	565
XIAO, M.	574
Y	
YONG, X.	496
Z	
ZAWISTOWSKI, T.	288
ZIELEWSKA-BÜTTNER, K.	405
ZIETHE, R.	288
ZIMMERMANN, C.	288
ZWICK, M.	197

PROGRAMMKOMITEE

Baltrusch, Sven
Baltsavias, Emmanuel
Bovet, Stéphane
Briese, Christian
Brunn, Ansgar
Cavegn, Stefan
Christen, Martin
Ernst, Julius
Franzen, Michael
Gervaix, Francois
Gold, Wolfgang
Grenzdörffer, Görres
Gülch, Eberhard
Haala, Norbert
Haunert, Jan-Henrik
Hochschild, Volker
Höfle, Bernhard
Hollaus, Markus
Jacobsen, Karsten
Jansa, Josef
Kaufmann, Viktor
Kersten, Thomas
Kneubühler, Mathias
Krzystek, Peter
Morsdorf, Felix
Navratil, Gerhard
Nebiker, Stephan
Petzold, Bettina
Ressler, Camillo
Sauerbier, Martin
Schneider, Danilo
Sörgel, Uwe
Weidner, Uwe
Wetzel, Hans-Ulrich